

GenomIUm: Un Método Basado en Patrones para el Diseño de Interfaces de Usuario de Acceso a Datos Genómicos



Diciembre 2019

Carlos Iñiguez Jarrin
PHD THESIS



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Directores:

Oscar Pastor Lopez
Ignacio Panach Navarrete

GenomIUm:
Un Método Basado en Patrones para el
Diseño de Interfaces de Usuario de
Acceso a Datos Genómicos

Carlos Efraín Iñiguez Jarrín



TESIS DOCTORAL DEPOSITADA EN CUMPLIMIENTO PARCIAL DE
LOS REQUISITOS PARA EL GRADO DE DOCTOR EN
INFORMÁTICA

Dirigida por:
Prof. Dr. Óscar Pastor López
Prof. Dr. Ignacio Panach

Diciembre 2019

GenomIUm: Un Método Basado en Patrones para el Diseño de Interfaces de Usuario de Acceso a Datos Genómicos

GenomIUm: Un Mètode Basat en Patrons per al Disseny de Interfícies d'Usuari d'Accés a Dades Genòmiques

GenomIUm: A Pattern Based Method for Designing User Interfaces for Genomic Data Access

Tesis defendida por Carlos Efraín Iñiguez Jarrín en diciembre del 2019 para la obtención del título de Doctor en Informática por la Universitat Politècnica de València.

Diseño de cubierta:

© Carlos Iñiguez

Autor

Carlos Efraín Iñiguez Jarrín, Universitat Politècnica de València. España

Supervisores

Prof. Dr. Óscar Pastor López, Universitat Politècnica de València, España

Prof. Dr. Ignacio Panach, Universitat de València, España

Comité de Evaluadores Externos:

Prof. Dra. Silvia Acuña Castillo, Universidad Autónoma de Madrid, España

Prof. Dra. Lourdes Moreno López, Universidad Carlos III, España

Prof. Dr. Víctor López Jaquero, Universidad de Castilla-La Mancha, España

Tribunal de Defensa:

Presidente: Prof. Dr. Vicente Pelechano, Universitat Politècnica de València, España

Secretario: Prof. Dra. Lourdes Moreno López, Universidad Carlos III, España

Vocal: Prof. Dra. María Dolores Lozano, Universidad de Castilla-La Mancha, España

Agradecimientos

A mis padres y hermanos que, pese a la distancia, se han valido de la tecnología para transmitirme el apoyo moral e incondicional para sortear los contratiempos que han aparecido en el camino.

A mis amigos Julio, Cecibel, Lenin, Nana, Ana León, Ana Marcen, José, Bea, Urko, Alberto, Ángel, Eduardo, quienes formaron parte fundamental de esta aventura doctoral y me enseñaron a transformar la adversidad en una oportunidad.

A mis mentores Oscar, Ignacio y Paco V., quienes han puesto luces, desde cada uno de sus mundos y en diferentes tiempos, para guiarme en la oscura, a veces desesperante, pero divertida y gratificante senda doctoral. Gracias Oscar por ser más que un tutor, un amigo, un confidente. Gracias Ignacio por compartir tus valiosos conocimientos, motivarme para conseguir mis objetivos y enseñarme a divisar las cosas desde una perspectiva práctica y objetiva. Gracias Paco Valverde por guiarme en mis primeros pasos en el programa doctoral.

A José Marín, por su preciado tiempo dedicado a revisiones y a sugerirme valiosos y oportunos comentarios.

A la SENESCYT y Escuela Politécnica Nacional por la oportunidad y apoyo económico para cumplir este objetivo.

A Dios por permitirme terminar esta etapa del camino, por enseñarme que somos frágiles y nada eternos, darme la fortaleza para seguir empujando y crear en mi mente la incertidumbre de que existe una mañana. Sin el regalo que me has dado, no habría estado aquí para escribir estas palabras.

A mi Música que, en las mejores y peores situaciones, nunca ha fallado como mi refugio, mi defensa y mi puerta de escape.

Resumen

La explosión de datos genómicos derivada de la secuenciación del ADN ha motivado el desarrollo de herramientas software que, a más de proveer la capacidad de almacenamiento y rendimiento para procesar los datos genómicos, incorporen interfaces de usuario que permitan a los investigadores (genetistas, analistas de bioinformática, médicos, biólogos) acceder a los datos genómicos almacenados y obtener conocimiento de ellos. La facilidad con la que los investigadores acceden a los datos genómicos depende en cierta medida de la facilidad con la que la interfaz de usuario (IU) puede ser usada. Es por eso por lo que, diseñar IUs intuitivas, eficaces y fáciles de usar es un requisito indispensable para desarrolladores y diseñadores de IUs inmersos en proyectos de software en este dominio.

El “diseño conceptual” de la IU es un artefacto principal de la etapa de diseño del proceso de desarrollo de la IU, en el cual desarrolladores y diseñadores plasman las decisiones de diseño para obtener IUs intuitivas, efectivas y fáciles de usar. En el dominio genómico, producir el diseño conceptual no es trivial. Las pocas o inexistentes guías y soluciones de diseño para abordar las necesidades de interacción del dominio genómico hacen del diseño conceptual una tarea desafiante para desarrolladores y diseñadores, novatos o expertos.

Con el fin de contribuir en el desarrollo de IUs que faciliten el acceso a los datos genómicos, diseñamos el método GenomIUm para cubrir el diseño conceptual de la IU. Para esto, GenomIUm, basado en el enfoque de diseño dirigido por patrones, define un proceso para diseñar el concepto de la IU y un catálogo de patrones de diseño que soporta a cada etapa del proceso. Así, GenomIUm guía al desarrollador de software en el diseño del concepto de la IU.

Esta investigación está guiada por la metodología Design Science promulgada por Roel Wieringa. Su enfoque en proyectos de investigación en Ingeniería de Software y Sistemas de Información hace de Design Science una metodología idónea para nuestra investigación. La metodología provee el marco de trabajo, los métodos de investigación y directrices para llevar a cabo la investigación y su aplicación asegura el rigor y validez científica de nuestros resultados. Para facilidad del lector, el presente documento de Tesis está organizado siguiendo la estructura de la metodología: inicia con la investigación del problema, sigue con el diseño de la solución propuesta para solucionar el problema y termina con la validación de la solución propuesta.

Para validar que GenomIUm cumple con los requisitos para los que fue diseñado, hemos ejecutado dos experimentos, uno por cada componente del método: los patrones y el proceso de diseño. El primer experimento evalúa el impacto producido por los patrones en la usabilidad de las IUs. El segundo experimento evalúa la percepción de desarrolladores de IUs respecto al uso del proceso de diseño. Adicionalmente, para ilustrar el uso de GenomIUm, describimos nuestra experiencia aplicando GenomIUm en un ambiente real.

Como proyectos futuros, planteamos i) enriquecer el catálogo de patrones con nuevos patrones identificados en IUs de aplicaciones emergentes en el dominio genómico y ii) aplicar GenomIUm en diversos casos de estudio con el fin de generalizar los efectos positivos reportados en esta Tesis.

Resum

L'Explosió de dades genòmiques derivada de la seqüenciació de l'ADN ha motivat el desenvolupament de ferramentes de programari que a més a més de proporcionar la capacitat d'emmagatzemament i rendiment per processar les dades genòmiques, incorporen interfícies d'usuari que permeten als investigadors (genetistes, analistes de bioinformàtica, metges, biòlegs) accedir a les dades genòmiques emmagatzemats i obtindre coneixement d'ells. La facilitat en la que els investigadors accedeixen a les dades genòmiques depèn de la facilitat en la que la IU puga ser utilitzada. Es per això pel que dissenyar IUs intuïtives, eficaces i senzilles d'utilitzar es converteix en un requisit indispensable per desenvolupadors i dissenyadors d'IUs immersos en projectes de programari en aquest domini.

El "disseny conceptual" de la IU és un artefacte principal de l'etapa de disseny del procés de desenvolupament de la IU, en el qual desenvolupadors i dissenyadors plasmen les decisions de disseny per a obtindre una UI intuïtiva, efectiva i fàcil d'utilitzar. En el domini genòmic, produir el disseny conceptual no és trivial. Les poques o inexistent guies i solucions de disseny per abordar les necessitats d'interacció del domini genòmic fan del disseny conceptual una tasca desafiadora per als desenvolupadors i dissenyadors, novells o experts.

Amb el fi de contribuir en el desenvolupament de IUs que faciliten l'accés a les dades genòmiques, nosaltres hem dissenyat el mètode GenomIUm per cobrir el disseny conceptual de la IU. Així, GenomIU, basat en l'enfocament de disseny dirigit per patrons, defineix un procés per a dissenyar el concepte de la IU i un catàleg de patrons de disseny que suporta a cadascuna de les etapes del procés. Així, GenomIUm guia al desenvolupador de programari en el disseny del concepte de la IU.

Aquesta investigació està guiada per la metodologia Design Science promulgada per Roel Wieringa. La seua aplicació en projectes d'investigació en Enginyeria de Programari i Sistemes d'informació fa de Design Science una metodologia idònia per la nostra investigació. La metodologia proveeix el marc de treball, els mètodes d'investigació i directrius per portar a terme la investigació i la seua aplicació assegura el rigor i validesa científica dels nostres resultats. Per facilitar la lectura, el present document de tesi està organitzat seguint l'estructura de la metodologia: inicia amb la investigació del problema, segueix en el disseny de la solució proposta per solucionar el problema i termina amb la validació de la solució proposta.

Per validar que GenomIUm compleix amb els requisits pels que va ser dissenyat, hem executat dos experiments, un per cada component del mètode: els patrons i el procés de disseny. El primer experiment avalua l'impacte produït pels patrons en la usabilitat de les IUs. El segon experiment avalua la percepció dels desenvolupadors de IUs respecte al ús del procés de disseny. Addicionalment, per il·lustrar l'ús de GenomIUm, descrivim nostra experiència aplicant GenomIUm en un ambient real.

Com projectes futurs, plantejem i) enriquir el catàleg de patrons amb nous patrons identificats en IUs d'aplicacions genòmic emergents i ii) aplicar GenomIUm en diversos casos d'estudi per generalitzar els efectes positius reportats en aquesta tesi.

Abstract

The explosion of genomic data derived from DNA sequencing has motivated the development of software tools that, in addition to providing the storage and performance to process genomic data, incorporate user interfaces that allow researchers (i.e., geneticists, analysts of bioinformatics, doctors, biologists) access to the stored genomic data and get knowledge of them. The ease with which researchers access genomic data depends to some extent on the ease with which the user interface (UI) can be used. That is why, designing intuitive, effective and easy-to-use UIs becomes an indispensable requirement for developers and designers which are involved in software projects in this domain.

The “conceptual design” of the UI is a main design artifact of the UI development process, in which developers and designers capture the design decisions to obtain intuitive, effective and easy-to-use UIs. In the genomic domain, producing the conceptual design is not trivial. The few or nonexistent guides and design solutions to address the interaction needs in the genomic domain make the conceptual design a challenging task for developers and designers, novice or experts.

In order to contribute to the development of UIs that facilitate access to genomic data, we design the GenomIUm method to cover the conceptual design of the UI. To do that, GenomIUm, based on the Pattern Oriented Design approach, defines a process to design the UI concept and a catalog of design patterns that supports each stage of the process. Thus, GenomIUm guides the software developer in the design of the UI concept.

This research is guided by the Design Science methodology enacted by Roel Wieringa. Its focus on Software Engineering and Information Systems research projects makes Design Science an ideal methodology for our research. The methodology provides the methodological framework, research methods, and guidelines for carrying out our research. The application of Design Science ensures the rigor and scientific validity of our results. For the convenience of the reader, this thesis document is organized according to the methodology structure. First, we describe the problem investigation. Then, we describe the design of the proposed solution to the problem. Finally, we describe the validation of the proposed solution.

To validate that GenomIUm meets the requirements for which it was designed, we executed two experiments, one for each GenomIUm component: the patterns and the design process. The first experiment evaluates the impact produced by the patterns on the usability of the UIs. The second experiment assesses the perception of UI developers regarding the use of the design process. Additionally, to illustrate the use of GenomIUm, we describe our experience applying GenomIUm in a real environment.

As future projects, we propose i) to enrich the catalog of patterns with new patterns identified in UIs of emerging applications in the genomic domain and ii) apply GenomIUm in various case studies in order to generalize the positive effects reported in this thesis.

Tabla de Contenidos

PARTE I. PREÁMBULO	1
Capítulo 1 Introducción	3
1.1 Motivación.....	3
1.2 Planteamiento del problema	4
1.3 Aproximación a la solución	6
1.4 Estructura de la Tesis.....	6
Capítulo 2 Metodología de Investigación	9
2.1 Objetivos y preguntas de investigación.....	10
2.2 El marco metodológico aplicado a la Tesis	12
2.3 Ciclo de diseño y ciclo empírico	13
2.3.1 Ciclo de Diseño	14
2.3.2 Ciclo Empírico	15
2.3.3 Ciclos aplicados a la Tesis.....	16
2.4 Conclusiones	18
PARTE II. INVESTIGACIÓN DEL PROBLEMA	19
Capítulo 3 Marco Conceptual	21
3.1 Un marco conceptual para los datos genómicos	21
3.1.1 Análisis de datos genómicos.....	24
3.1.2 Priorización de variaciones genéticas	25
3.2 Un marco conceptual para el diseño de interfaz de usuario.....	26
3.2.1 La interfaz de usuario	26
3.2.2 Diseño conceptual de la interfaz de usuario	27
3.2.3 Patrones	28
3.3 Conclusiones	33
Capítulo 4 Estado del Arte	35
4.1 Enfoques de diseño de interfaces basados en patrones	35
4.1.1 DCU	35
4.1.2 GDD.....	36
4.1.3 PSA.....	38
4.1.4 POD.....	40
4.2 Patrones en el dominio genómico	42
4.2.1 Patrones de diseño de interfaces existentes	42
4.2.2 Patrones de diseño de interfaces y el dominio genómico	46
4.3 Conclusiones	47

PARTE III. DISEÑO DE LA SOLUCIÓN.....	49
Capítulo 5 Requisitos de la Solución	51
5.1 Introducción.....	51
5.2 Requisitos de la solución prevista	51
5.2.1 El método SILE y el panorama de diseño de interfaces	52
5.2.2 Problemas de interacción con datos genómicos	53
5.2.3 Requisitos	54
5.3 Determinación respecto a alternativas de solución.....	54
5.4 Insumos necesarios para el método.....	57
5.5 Conclusiones	58
Capítulo 6 Método GenomIUm	59
6.1 Visión general.....	59
6.2 Catálogo de Patrones: GenomIU.....	60
6.2.1 Modelo conceptual del catálogo	62
6.2.2 Identificación de patrones genómicos	64
6.2.3 Evolución de los patrones genómicos	66
6.2.4 Inventario de Herramientas: Interfaces de Usuario estudiadas.....	69
6.2.5 Formato para describir los patrones.....	73
6.2.6 Patrones Genómicos.....	74
6.2.7 Interrelaciones entre patrones genómicos.....	104
6.3 Pasos del proceso de diseño	106
6.3.1 Diseño arquitectónico	106
6.3.2 Diseño estructural	107
6.3.3 Diseño de contenido.....	108
6.3.4 Refinamiento.....	109
6.4 Conclusiones	110
PARTE IV. VALIDACIÓN.....	111
Capítulo 7 Validación	113
7.1 Introducción.....	113
7.2 Validación de interfaces implementadas con patrones.....	113
7.2.1 Objetivo y preguntas de investigación	113
7.2.2 Planeamiento.....	115
7.2.3 Análisis de Datos e Interpretación de Resultados.....	127
7.2.4 Resumen	138
7.3 Validación del proceso de diseño.....	139
7.3.1 Objetivo y Preguntas de Investigación	139
7.3.2 Planeamiento.....	140
7.3.3 Análisis de datos e Interpretación de Resultados	144
7.3.4 Resumen	147

7.4 Caso de estudio: Genoma de los cítricos.....	147
7.4.1 Objetivo.....	147
7.4.2 Dominio y participantes.....	148
7.4.3 Resultados	157
7.5 Conclusiones	157
Capítulo 8 Discusión Final.....	159
8.1 Contribuciones.....	159
8.2 Resumen de publicaciones	161
8.2.1 Publicaciones	161
8.2.2 Participación en proyectos de investigación	163
8.2.3 Organización de congresos	163
8.3 Trabajos futuros	163
Referencias	165
Lista de Imágenes	175
Lista de Tablas	179

Acrónimos

ADN	Ácido desoxirribonucleico
CSCW	Trabajo Cooperativo Asistido por Computadora. Por sus siglas en inglés, <i>Computer Supported Cooperative Work</i>
CSUQ	Computer System Usability Questionnaire
DCU	Diseño Centrado en el Usuario
DE	Desviación Estándar
DP	Profundidad de lectura. Por sus siglas en inglés, <i>Read Depth</i>
DS	Design Science
EPN	Escuela Politécnica Nacional
FTP	Protocolo de Transferencia de Ficheros. Por sus siglas en inglés, <i>File Transfer Protocol</i>
GDD	Diseño Orientado por Objetivos. Por sus siglas en inglés, <i>Goal Directed Design</i>
GQ	Calidad del Genotipo. Por sus siglas en inglés, <i>Genotype Quality</i>
GUI	Interfaz Gráfica de Usuario. Por sus siglas en inglés, <i>Graphical User Interface</i>
HCI	Interacción Humano Computador. Por sus siglas en inglés, <i>Human Computer Interaction</i>
HGVS	Human Genome Variation Society
HTML	Lenguaje de marcas de HyperTexto. Por sus siglas en inglés, <i>Hypertext Markup Language</i>
ICGC	International Cancer Genome Consortium
IRS	Sistemas de recuperación de datos. Por sus siglas en inglés, <i>Information Recovery Systems</i>
IS	Ingeniería de Software
ISCN	International System for Chromosome Nomenclature
IU	Interfaz de Usuario
IVIA	Instituto Valenciano de Investigaciones Agrarias
KEGG	Kyoto Encyclopedia of Genes and Genomes
MCGH	Modelo Conceptual del Genoma Humano
MLM	Modelo Mixto Lineal
MPM	Meta/Pregunta/Medida
MVC	Modelo Vista Controlador
NCBI	National Center for Biotechnology Information
NGS	Secuenciación de Segunda Generación. Por sus siglas en inglés, <i>Next Generation Sequencing</i>
OMIM	Online Mendelian Inheritance in Man
PC	Pregunta de conocimiento

PD	Problema de Diseño
PI	Pregunta de Investigación
POD	Diseño Orientado por Patrones. Por sus siglas en inglés, <i>Pattern Oriented Design</i>
PROS	Centro de Investigación en Producción de Software
PSA	Enfoque Soportado por Patrones. Por sus siglas en inglés, <i>Pattern Supported Approach</i>
QUAL	Calidad de lectura
SENESCYT	Secretaría Nacional de Educación Superior Ciencia
SI	Sistemas de Información
SILE	Método de Búsqueda (S), Identificación (I), Carga (L) y Extracción (E)
SNV	Variación de un único nucleótido. Por sus siglas en inglés, <i>Single Nucleotide Variation</i>
TE	Tamaño del Efecto
UI	Unidad de Interacción
UPV	Universidad Politécnica de Valencia
URL	Localizador de recursos uniforme. Por sus siglas en inglés, <i>Uniform Resource Locator</i>
UX	Experiencia de usuario. Por sus siglas en inglés, <i>User experience</i>
VCF	Variant Call File

PARTE I.
PREÁMBULO

Capítulo 1

Introducción

1.1 MOTIVACIÓN

La interfaz de usuario es un componente importante de todo sistema interactivo; es a través de la interfaz que el usuario visualiza, explora, descubre y extrae conocimiento de los datos. El rol que desempeña la interfaz de usuario es crucial en varios dominios.

En el dominio genómico, en concreto, el conocimiento que los investigadores (genetistas, médicos, bioinformáticos, etc.) pueden obtener a través de la interfaz de usuario se traduce posiblemente en la creación de nuevas especies frutales, en el caso de programas de mejoramiento genético en la investigación agraria, o en nuevos tratamientos o maneras de prevenir enfermedades, en el caso de la medicina. En la “medicina personalizada”, por ejemplo, los investigadores interactúan con la interfaz de usuario de aplicaciones genómicas para analizar y obtener conocimiento del genoma de un paciente y aplicar dicho conocimiento para recomendar medicamentos, dosis específicas o tratamientos para la o las enfermedades del paciente.

El crecimiento exponencial del volumen de los datos derivado de la revolución digital ha creado la necesidad de evolucionar las interfaces de usuario a interfaces cada vez más sencillas y *usables* (fáciles de usar) para adaptarse a las necesidades de consumo masivo de datos. La necesidad de diseñar interfaces de usuario para las herramientas del dominio genómico, que mejoren la capacidad de visualizar u organizar la información genómica, es una demandante proposición (Javahery et al., 2004). Los investigadores requieren interfaces que permitan explorar y analizar inmensos volúmenes de datos genómicos, derivados de la secuenciación masiva de ADN y de diferentes bases de datos biológicas, para obtener conocimiento que apoye sus investigaciones.

En el dominio genómico, los mecanismos para diseñar interfaces son escasos haciendo del diseño una tarea desafiante para quienes la llevan a cabo. El artículo titulado *UX design, bioinformatics and cars* (Rowland, 2015), menciona algunos de los desafíos que los diseñadores de interfaces deben afrontar en el dominio genómico. El autor del artículo, un desarrollador web y diseñador de experiencia de usuario (UX) del Instituto Bioinformático Europeo, discute el cómo las herramientas bioinformáticas y sus interfaces están desarrolladas (referenciando al artículo *Si los carros fueran construidos por los bioinformáticos* (Filion, 2015) y ofrece, como diseñador, su opinión acerca de los desafíos en el diseño de interfaces en este dominio, los cuales se sintetizan en:

- *Investigación de usuario.* - Es necesario realizar investigación de usuario en etapas tempranas del proceso de diseño para entender qué es lo que los usuarios están tratando de hacer;
- *Usabilidad de Interfaces de usuario.* - La velocidad y la precisión son factores importantes en las aplicaciones bioinformáticas, sin embargo, la *usabilidad* de las interfaces y otras características no deben ser pasadas por alto;
- *Complejidad si, complicadas de usar no.* - Las herramientas bioinformáticas necesitan un buen diseño que exprese apropiadamente su nivel de complejidad pero que no haga complicado su uso;
- *Patrones de diseño.* - Es necesario pensar en la creación de un conjunto de patrones de diseño para las interfaces e interacciones de las herramientas bioinformáticas;
- *Participación multidisciplinar.* - Diseñar herramientas que proveen valor real a la comunidad, implica el trabajo conjunto de diseñadores, expertos del dominio, desarrolladores, y otros colaboradores.

El importante rol que juega la interfaz de usuario en la extracción de conocimiento en el dominio genómico junto con los desafíos que supone el diseño de tales interfaces, nos han motivado a contribuir con *GenomU* (un método para diseñar interfaces de usuario de acceso a datos genómicos) como soporte a desarrolladores de interfaces que se enfrentan a proyectos en este dominio.

1.2 PLANTEAMIENTO DEL PROBLEMA

El dominio genómico es un dominio donde los investigadores del área médica-genética (genetistas, biólogos, médicos, analistas en bioinformática) interactúan con cantidades masivas de datos provenientes de la secuenciación del material genético (Acido desoxirribonucleico, ADN), fuentes de bibliografía biomédica y fuentes de datos biológicos. En este dominio, el desarrollo de herramientas ha hecho un gran avance en aspectos de tratamiento de los datos como la extracción, almacenamiento y transformación de datos; sin embargo, los aspectos de acceso y explotación de los datos que están relacionados directamente con las características de la interfaz de usuario como son la priorización, filtrado, visualización, manipulación y navegación, han sido poco abordados (Pabinger et al., 2014).

En la práctica, el diseño de la interfaz en el dominio genómico está enfocado, en la mayoría de los casos, a los aspectos de funcionalidad de las herramientas más que en la manera en la que los usuarios, beneficiarios reales de la información, usan la interfaz para cumplir sus objetivos. No es difícil encontrar, en este dominio herramientas que realizan funcionalidades complejas pero que son complicadas de comprender, haciendo que su uso esté limitado a investigadores con habilidades informáticas capaces de ejecutar instrucciones de comandos de lenguajes de programación (p. ej. Python, R).

Diseñar interfaces de usuario para acceder a los datos genómicos es una tarea compleja que desafía a diseñadores y desarrolladores de interfaces a entender y comprender los complejos conceptos biológicos y sus relaciones, además de lidiar con problemas de interacción y usabilidad ajenos a su experticia. Además, las herramientas o directrices para guiar la construcción de la interfaz de usuario en el dominio genómico son muy escasas o

no han sido identificadas. Muchos de los desarrolladores pasan por alto el diseño de la interfaz de usuario y prefieren construir la interfaz a la par que codifican la funcionalidad de la aplicación genómica. El resultado, en el mejor de los casos, interfaces poco prácticas e intuitivas, es decir, interfaces que son difíciles de comprender, aprender y usar, y que afectan el trabajo de los investigadores.

En el contexto genómico, el panorama de diseño de interfaces de usuario, que tienen que ver con quienes diseñan las interfaces, cómo diseñan las interfaces y quienes usan las interfaces, puede ser visto desde tres perspectivas:

1. *Actores en el diseño de interfaces de usuario*: Las prácticas tradicionales de desarrollo de software en el dominio de trabajo estudiado confían el diseño de la interfaz de usuario a dos actores principales: i) analistas en bioinformática con conocimientos en análisis biológico, codificación de scripts de análisis de datos y escasa o ninguna experiencia en diseño y desarrollo de interfaces de usuario; y ii) desarrolladores de software con limitado conocimiento en biología y usabilidad y que dan mayor importancia a la facilidad de codificación.
2. *Concepción de las interfaces*. - Los desarrolladores construyen la interfaz de usuario a la par que codifican la funcionalidad de la aplicación. El proceso de diseño de la interfaz en este dominio prioriza la implementación de la interfaz en lugar de su diseño, una práctica que resulta con demasiada frecuencia en interfaces poco prácticas e intuitivas. Otro problema que afecta la concepción de las interfaces es la complejidad del dominio: los desarrolladores de software tienen dificultad para entender el dominio debido a la complejidad de los términos biológicos y sus relaciones. Esta barrera de comprensión del dominio impacta en la comunicación con los usuarios finales haciendo difícil determinar los problemas de interacción y en consecuencia las posibles soluciones para diseñar la interfaz.
3. *Usuarios finales*: Los usuarios que interactúan directamente con los datos genómicos son médicos, genetistas, bioinformáticos e investigadores quienes, en gran número, tienen profundos conocimientos en el dominio genómico pero su habilidad para usar el computador es habitualmente limitada, así como sus conocimientos en informática. Paradójicamente, muchas de las interfaces de usuario de las herramientas de gestión genómica están diseñadas para uso exclusivo de usuarios con conocimientos de informática, capaces de ejecutar, en la mayoría de los casos, complejos comandos que son difíciles de aprender y recordar que además afectan a la eficiencia y efectividad del trabajo de los investigadores y por ende a su satisfacción en el uso de estas interfaces.

Considerando lo mencionado, el problema abordado por esta Tesis es:

“En el dominio genómico, la interfaz de usuario es un componente clave para acceder a los datos y obtener conocimiento de ellos. La complejidad del conocimiento del dominio genómico y las escasas guías para diseñar interfaces de usuario en este dominio han llevado a los desarrolladores a implementar interfaces de usuario poco intuitivas y difíciles de usar para los investigadores (genetistas, bioinformáticos, médicos, biólogos)”.

Con el propósito de guiar en la producción de interfaces de usuario usables (fáciles de usar) y que cumplan con las necesidades de los usuarios finales, los enfoques de diseño de interfaz de usuario existentes promueven el uso de “patrones de diseño de interfaz de

usuario” como herramientas de apoyo para diseñar la interacción usuario-sistema. Uno de estos enfoques es el Diseño Orientado por Patrones (POD), el cual promueve la idea de componer la interfaz de usuario combinando patrones de interfaz basado en las interrelaciones entre los patrones. Los patrones de diseño de interfaz de usuario recopilan el conocimiento, experiencia y buenas prácticas de diseñadores y desarrolladores expertos en un dominio específico y describen tales experiencias en una forma fácil de entender por usuarios, desarrolladores y diseñadores, convirtiéndose en guías para solucionar problemas de interacción en la interfaz de usuario.

De esta manera, el uso de los patrones bajo el enfoque de POD se convierte en una interesante estrategia para abordar el problema del diseño de interfaces de acceso a datos en el contexto genómico. El problema que resuelve esta Tesis está enmarcado en la identificación, construcción y catalogación de soluciones de diseño repetitivas utilizadas en el diseño de interfaz de usuario y cómo aplicar dichas soluciones de diseño.

1.3 APROXIMACIÓN A LA SOLUCIÓN

En la literatura, se evidencia dos enfoques principales para diseñar la interfaz de usuario (Radeke et al., 2007): dirigido por modelos y basado en patrones de interfaces de usuario. El primero construye una interfaz concreta a partir de transformaciones de modelos que especifican la interfaz a un nivel abstracto. El segundo construye la interfaz reutilizando *patrones de interfaz de usuario*, soluciones de diseño de interfaz de usuario que han sido utilizadas recurrentemente en otras interfaces para abordar problemas comunes en el diseño de la interfaz de usuario, como bloques de construcción de la interfaz de usuario.

Esta Tesis sigue el segundo enfoque donde el diseño de la interfaz se realiza usando o reutilizando soluciones de diseño previamente definidas y probadas con el fin de “*no reinventar la rueda*”. Aunque la identificación, definición y uso de los patrones es evidente en varios dominios, en el dominio genómico no existen patrones que cubran los problemas específicos de interacción con los datos y es ahí donde se presenta el desafío.

La solución que proponemos en esta Tesis es un método de diseño de interfaces en el dominio genómico, al que hemos denominado GenomIUm. El método está basado en el POD lo cual hace que el método promueva el uso sistemático de patrones para diseñar las interfaces. Así, los patrones se convierten en bloques de construcción que el desarrollador utiliza para formar la interfaz de usuario.

La contribución principal de esta Tesis es el diseño del método GenomIUm que incluye la definición de: 1) un catálogo de patrones para desarrollar interfaces de usuario en el dominio genómico, 2) el proceso para usar el catálogo de patrones en el diseño de interfaces, y 3) la validación empírica del método.

1.4 ESTRUCTURA DE LA TESIS

Esta Tesis está organizada en cuatro partes y cada parte contiene capítulos. La parte I explica el contenido de la Tesis de manera general a través de dos capítulos. El Capítulo 1 es una introducción donde se describe la motivación y el problema que aborda la Tesis. El Capítulo 2, con un carácter más estructural y metodológico, describe la metodología de

investigación *Design Science* (DS) que guía nuestra investigación y que sirve como estructura de este documento de Tesis como tal.

Las tres partes restantes del documento describen el desarrollo de la Tesis siguiendo los lineamientos de la metodología DS para un ciclo de diseño. La metodología DS consiste en tres tareas de alto nivel (que se describen en detalle en el Capítulo 2):

- Investigación del problema
- Diseño de la solución
- Validación de la solución

Por lo tanto, las tres partes restantes del documento de Tesis se corresponden con cada una de las tareas de alto nivel de la metodología DS. Así, los capítulos dentro de cada parte detallan las actividades de la tarea de la metodología correspondiente. De esta manera:

La parte II trata la *investigación del problema* e incluye dos capítulos: 1) *marco de trabajo conceptual* (Capítulo 3), que define y describe los conceptos y términos empleados en esta Tesis y que facilitan la comprensión del lector, y 2) *estado del arte* (Capítulo 4), que describe el estado actual del diseño de interfaces de usuario en el dominio genómico y los métodos de diseño de interfaz de usuario existentes.

La parte III trata el *diseño de la solución* donde se describe por capítulos: el *método de diseño GenomIUm* (Capítulo 6) como solución propuesta de esta Tesis y que está basada en los *requisitos especificados* en el Capítulo 5.

La parte IV trata la *validación de la solución* que incluye un capítulo donde se presentan tres validaciones: la *validación empírica de los patrones* (Sección 7.2) que describe el experimento de usabilidad de interfaces de usuario que incorporan los patrones propuestos; la *validación empírica del método por parte de desarrolladores* (Sección 7.3); un *caso de estudio* donde se aplica el método a un problema del mundo real (Sección 7.4).

Finalmente, el Capítulo 8 presenta una discusión general sobre la investigación que incluye principalmente las respuestas a las preguntas de investigación planteadas en el Capítulo 2, el resumen de publicaciones derivadas de la investigación y una descripción de proyectos como futuros pasos a seguir en esta línea de investigación.

Capítulo 2

Metodología de Investigación

Esta investigación está basada en la metodología *Design Science* (DS) (Wieringa, 2014). Hemos escogido esta metodología por su enfoque en la investigación de proyectos de Sistemas de Información (SI) e Ingeniería de Software (IS) a través de la experimentación, observación del fenómeno de estudio y análisis de resultados, lo cual hace de esta metodología una guía adecuada para llevar a cabo esta investigación. Investigar en proyectos de SI e IS es investigar cómo el desarrollo, operación y mantenimiento son conducidos por los ingenieros de software y las partes interesadas, bajo diferentes condiciones (Wohlin et al., 2012).

Esta Tesis aplica la metodología DS para investigar cómo el diseño de interfaces de usuario en el dominio genómico es conducido por los diseñadores de interfaces de usuario y cómo estas interfaces de usuario satisfacen las necesidades de consumo de datos de los usuarios de este dominio.

El objeto de estudio de cualquier proyecto basado en DS es “estudiar un artefacto interactuando en su contexto del problema”, a lo cual la metodología lo denomina **tratamiento**. Cuando hablamos de “artefacto” nos referimos a un elemento de software (p. ej., método, proceso, aplicación de software, etc.) que es diseñado por los investigadores del proyecto DS y utilizado por las personas como una solución a un problema. El término “tratamiento” utilizado en la metodología DS puede ser entendido a través de una analogía con la medicina donde uno o varios tratamientos son aplicados a un enfermo como soluciones para curarlo. Desde este punto de vista, la “píldora” usada en el tratamiento de medicina para curar al enfermo es análoga al “artefacto” usado en el tratamiento de la metodología DS para solucionar un problema. Con esta idea, el objeto de estudio de esta Tesis es: estudiar el método GenomIUm (nuestro **artefacto**) interactuando con el **problema** del acceso a datos en el contexto genómico, como lo muestra la Imagen 2.1.

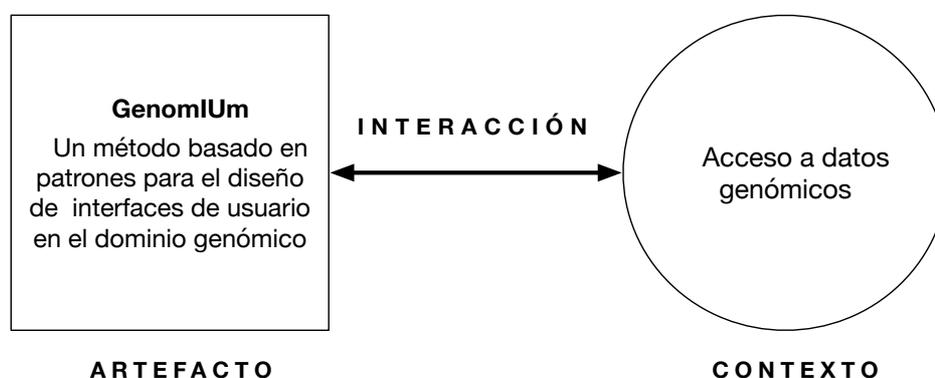


Imagen 2.1 El objeto de estudio del proyecto de investigación: el artefacto interactuando con el contexto para mejorar un problema en ese contexto.

Así como un tratamiento en la medicina usa la “píldora” como solución para curar al enfermo, el tratamiento en nuestro proyecto DS usa el método GenomIUm como solución para “curar” las dolencias o dificultades de diseñar interfaces para acceder (es decir, explorar, navegar, operar, visualizar) a los datos en el dominio genómico. La intención de este trabajo de investigación es que la aplicación del método GenomIUm contribuya a la mejora de la usabilidad de las interfaces de usuario, facilitando así la explotación y obtención de conclusiones a partir de los datos genómicos.

El tratamiento, incluyendo el artefacto y el conocimiento asociado a su creación y aplicación, constituye la principal contribución científica de la presente investigación.

2.1 OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN

El objetivo principal de la Tesis es:

Definir GenomIUm: un método para el diseño de interfaces de usuario de acceso a datos genómicos.

Como se mencionó en la Sección 1.1, la Tesis se fundamenta fuertemente en el Diseño Orientado por Patrones (POD) como estrategia para solucionar el problema de diseñar interfaces de acceso a datos genómicos. El objetivo del POD es producir el *diseño conceptual* de la interfaz de usuario. El diseño conceptual es el prototipo o diseño temprano de la interfaz que captura en esencia la estructura, comportamiento y presentación (Preece et al., 2015). Bajo el enfoque POD, el uso de los *patrones de interfaz de usuario* es fundamental para producir el diseño conceptual. Los patrones son utilizados como herramientas poderosas que describen el conocimiento de diseñadores expertos y se comportan como bloques de construcción que facilitan el diseño conceptual de la interfaz. En consecuencia, el método GenomIUm se centra en la producción del diseño conceptual a partir de la combinación de patrones de interfaz de usuario.

Por lo mencionado y para lograr el objetivo principal, es necesario contestar a las siguientes preguntas de investigación (PI):

- **PI1.** ¿Qué problemas existen en el acceso a datos genómicos?

- **PI2.** ¿Cuáles son los enfoques o herramientas existentes para el diseño de interfaces que pueden ser aplicables al contexto del acceso a datos genómicos?
- **PI3.** ¿Cómo definir el método GenomIUm para diseñar interfaces de usuario de acceso a datos genómicos?
 - **PI3.1** ¿Cómo definir un catálogo de patrones de diseño de interfaz de usuario específico para el acceso a datos genómicos?
 - **PI3.2** ¿Cómo guiar al desarrollador en el diseño conceptual de la interfaz de usuario?
- **PI4.** ¿Qué impacto produce el método GenomIUm en el diseño de interfaces de usuario de acceso a datos genómicos?
 - **PI4.1.** ¿Cuál es el impacto del uso de los patrones en la usabilidad de las interfaces de usuario?
 - **PI4.2.** ¿Cuál es el impacto del uso del proceso de diseño del método GenomIUm desde el punto de vista de los diseñadores de interfaces?
 - **PI4.3.** ¿Cuál es el impacto de la aplicación del método GenomIUm en un contexto del mundo real desde el punto de vista de las percepciones de las partes interesadas?

Para contestar estas preguntas, se plantean los siguientes objetivos específicos y se menciona, en cada uno de ellos, el método de investigación empleado para lograrlo:

Objetivo 1 (PI1). Para contestar la PI1, vamos a identificar los problemas de acceso a los datos genómicos a nivel de la interfaz de usuario mediante la revisión de literatura y la realización de entrevistas con investigadores del dominio como son genetistas, bioinformáticos y biólogos.

Objetivo 2 (PI2). Para contestar la PI2, vamos a revisar la literatura en busca de enfoques y herramientas existentes para diseñar interfaces de usuario y que puedan ser aplicadas al dominio genómico.

Objetivo 3 (PI3). Para contestar la PI3, vamos a definir un catálogo de patrones de interfaces de usuario específico del dominio genómico y el proceso a través del cual el catálogo puede ser aplicado para producir el diseño conceptual de la interfaz de usuario. Primero, para responder la PI3.1, construimos el catálogo de patrones estudiando sistemáticamente las interfaces de usuario existentes, identificando en ellas las soluciones de diseño que han sido empleadas repetitivamente por diseñadores y desarrolladores expertos. Segundo, para responder la PI3.2, definimos el proceso para aplicar el catálogo de patrones, tomando como base las estrategias empleadas por los métodos investigados en el Objetivo 2.

Objetivo 4 (PI4): Para contestar la PI4, vamos a realizar dos experimentos y un caso de estudio. El primer experimento, orientado a responder la PI4.1, es un experimento de comparación estadística que evalúa la usabilidad de las interfaces diseñadas con patrones en términos de efectividad, eficiencia y satisfacción. El segundo experimento, orientado a contestar la PI4.2, es un experimento dirigido a evaluar la percepción de los desarrolladores de software que usan el método para diseñar interfaces de usuario, en términos de utilidad, facilidad de uso e intención de uso. El caso de estudio, orientado a contestar la PI4.3, tiene como objetivo aplicar el método para resolver un problema en un contexto real (acceder a

los datos del genoma de las frutas cítricas) y evaluar la percepción de los usuarios de dicho contexto.

2.2 EL MARCO METODOLÓGICO APLICADO A LA TESIS

Para lograr los objetivos mencionados en la sección anterior y contestar a las PI's, la metodología provee un marco de trabajo que consiste en dos contextos interactuando con el proyecto DS. Estos contextos son el *contexto social* y el *contexto de conocimiento*, como lo muestra la Imagen 2.2.

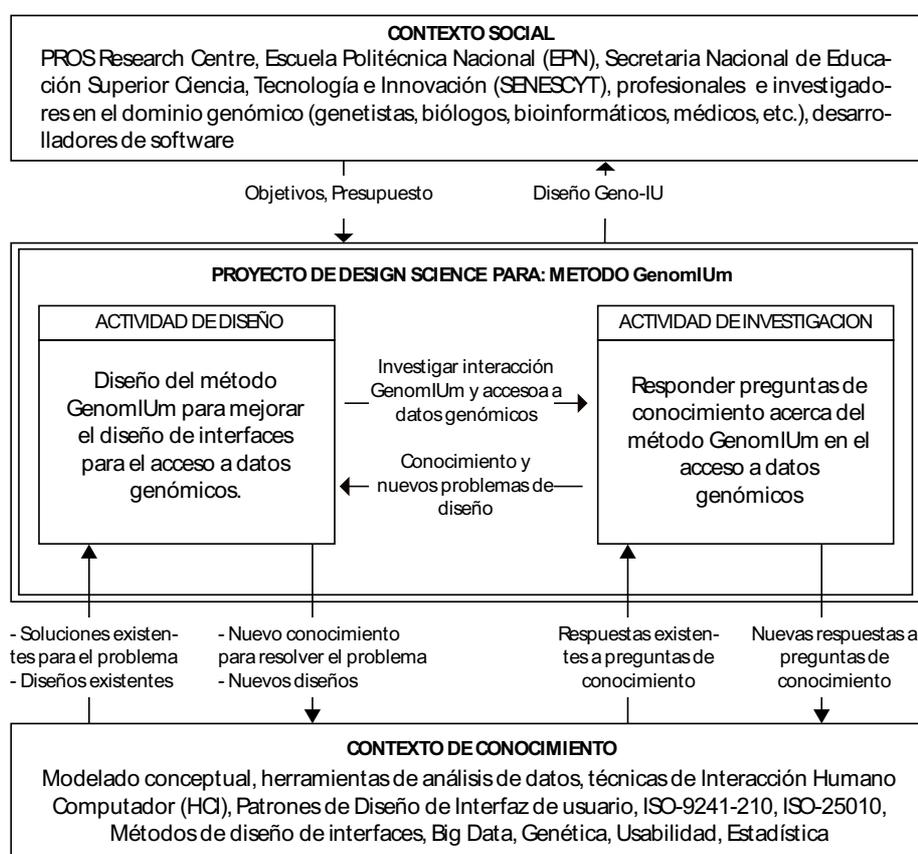


Imagen 2.2 Marco de trabajo de la metodología DS aplicado a la Tesis

El *contexto social* representa a las partes interesadas del proyecto incluyendo a las personas o instituciones que financian el proyecto y/o definen los objetivos o requisitos para el método GenomIUm. Una vez concluida la investigación, las partes interesadas reciben el método GenomIUm completamente diseñado. Las partes interesadas que

conforman el contexto social de esta Tesis se dividen en dos grupos. El primero lo conforman las partes interesadas que patrocinan el proyecto de investigación y son:

- Centro de Investigación de Métodos de Producción de Software (PROS) de la Universidad Politécnica de Valencia
- Escuela Politécnica Nacional del Ecuador (EPN)
- Secretaría Nacional de Educación Superior, Ciencia y Tecnología (SENESCYT)

El segundo grupo lo conforman las partes interesadas que son beneficiarios directos del producto resultante de la investigación, es decir el método GenomIUm, y son:

- Laboratorios biológicos e investigadores en el área de la medicina y genética.
- Diseñadores de interfaces de usuario

El *contexto de conocimiento* representa la literatura científica existente que se ha utilizado para llevar a cabo la investigación y también es donde residen la o las contribuciones producidas por esta investigación (especialmente artículos científicos que se detallan más adelante en la Sección 8.2.1). En esta Tesis, el contexto de conocimiento incluye las fuentes primarias de conocimiento como son literatura científica, profesional, técnica y comunicaciones orales existentes en las disciplinas HCI, IS, Big Data, y Genética e IS Empírica.

La *actividad de diseño* y la *actividad de investigación*, ubicadas dentro del cuadrado de doble línea de la Imagen 2.2, consisten en resolver *problemas de diseño (PD)* y responder a *preguntas de conocimiento (PC)*, respectivamente y durante el desarrollo del proyecto DS.

Los **PD** son problemas sobre cómo diseñar el método GenomIUm considerando las necesidades de las partes interesadas mencionados en el contexto social. Las soluciones a los PD son diseños nuevos (por ejemplo, modelos, prototipos de interfaz de usuario, patrones de diseño de interacción) que alteran el mundo real y están basadas en los objetivos y requisitos de las partes interesadas. Por lo tanto, evaluar estos diseños significa evaluar que estos diseños cumplan con los requisitos de las partes interesadas.

Las **PC** son preguntas enfocadas a conocer el contexto del problema, las maneras de diseñar el método, la manera en que el método interactúa en el contexto de problema, entre otras. Estas preguntas se responden investigando teorías, técnicas, modelos, etc. pertenecientes al contexto de conocimiento. Al contrario de las soluciones a los PD, las respuestas a las PC agregan conocimiento a nuestro entender del mundo real, más no lo alteran, y son independientes de los objetivos que persiguen las partes interesadas. Para responder a un PC, nosotros como investigadores recolectamos y analizamos datos a través de métodos de investigación (revisión de literatura, encuestas, entrevistas con expertos, experimentos, observación de casos, investigación de acción técnica, etc.). En efecto, en la Sección 7.2 y la Sección 7.3 presentamos dos experimentos que contestan PCs acerca de los efectos producidos por las partes que componen el método GenomIUm.

2.3 CICLO DE DISEÑO Y CICLO EMPÍRICO

En la práctica, para llevar a cabo las actividades de diseño e investigación en un proyecto DS, la metodología provee respectivamente dos ciclos iterativos y anidados: *Ciclo de Diseño* y *Ciclo Empírico*. Cada ciclo está compuesto de tareas y cada tarea involucra resolver PDs y contestar PCs.

2.3.1 Ciclo de Diseño

El ciclo de diseño es un proceso orientado al diseño del artefacto de la investigación y puede ser visto como un sub-ciclo de un típico ciclo de ingeniería enfocado a la resolución de problemas. El ciclo de ingeniería está compuesto de cinco tareas:

- T1. Investigación del problema. – Identificar el fenómeno que debe ser mejorado. Esto incluye identificar las causas que generan el fenómeno, las partes interesadas en el contexto del problema y los objetivos que se desea alcanzar. Además, se definen los conceptos que giran alrededor del problema.
- T2. Diseño del tratamiento. – Diseñar artefactos para tratar el problema. Para esto, se especifican sus requisitos, se estudian tratamientos existentes que pueden resolver el problema y se decide si adaptar un tratamiento existente o diseñar uno nuevo.
- T3. Validación del tratamiento. – Verificar que el o los diseños de los artefactos traten el problema.
- T4. Implementación del tratamiento. - Tratar el problema con uno de los artefactos diseñados.
- T5. Evaluación de la implementación. - Evaluar el éxito del tratamiento en condiciones reales.

De estas cinco tareas del ciclo de ingeniería, el ciclo de diseño abarca las tareas T1, T2 y T3, como lo muestra la Imagen 2.3. Las tareas T4 y T5 que están en color gris claro corresponden al ciclo de ingeniería y ha sido incluidas en el gráfico para expresar la diferencia entre los dos ciclos. Dentro de cada tarea se describe los PD (anotadas con el signo “!”) y PC (anotadas con el signo “?”) que deben ser abordados.



Imagen 2.3 Ciclo de diseño de la metodología DS. Adaptado de (Wieringa, 2014)

En consecuencia, el ciclo de diseño está compuesto de las tareas T1, T2, y T3 y no aborda la implementación del tratamiento ni su evaluación. Para abordar la T3, el ciclo de diseño se enlaza con el ciclo empírico, el cual se describe en la siguiente sección.

2.3.2 Ciclo Empírico

El ciclo empírico es un proceso orientado a contestar preguntas de conocimiento científico de manera racional, donde el investigador diseña la configuración de la investigación (o estudio empírico, como por ejemplo un experimento) y analiza los datos producidos de esta configuración. El ciclo empírico se muestra en la Imagen 2.4 y se compone de cinco tareas que han sido denotadas con la letra T' para diferenciarse de las tareas del ciclo de diseño.

T'5. Análisis de datos

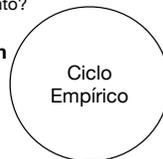
- Descripciones?
- Conclusiones estadísticas?
- Explicaciones?
- Generalizaciones?
- Respuestas a preguntas de conocimiento?

T'4. Ejecución de la investigación

- ¿Que sucedió?

T'3. Validación

- Validez del objeto de estudio?
- Validez de la especificación del tratamiento?
- Validez de especificación de las medidas?
- Validez del diseño de la inferencia



T'1. Análisis del problema de investigación

- Marco de trabajo conceptual?
- Preguntas de conocimiento?
- Población?

T'2. Diseño de la investigación y la inferencia

- Objetos de estudio?
- Especificación de tratamientos?
- Especificación de medidas y métricas?
- Inferencias?

Imagen 2.4 Ciclos Empírico de la metodología DS. Adaptado de (Wieringa, 2014)

Las tareas del ciclo empírico son:

- T'1. Análisis del problema de investigación. - Definir el contexto del problema sobre el cual vamos a realizar el estudio, las preguntas de conocimiento y la población objetivo de la cual obtendremos los datos (es decir, los participantes del experimento o estudio).
- T'2. Diseño de la investigación e inferencia. - Diseñar el estudio empírico definiendo las variables dependientes que queremos medir y cómo medirlas (las métricas), las variables independientes (conocidos como tratamientos), y los objetos experimentales (los problemas que proponemos resolver a los participantes). Además, diseñamos la configuración del estudio (por ejemplo, en un experimento de comparación estadística, diseñamos como los tratamientos y problemas serán asignados a los participantes), definimos el procedimiento sistemático que se seguirá para ejecutar el estudio y definimos, según la configuración del estudio empírico, los métodos estadísticos para realizar inferencias y obtener conclusiones.
- T'3. Validación del diseño de la investigación. - Validar las amenazas que pueden afectar al estudio empírico y a los resultados. Específicamente realizamos cuatro tipos de validaciones (Wohlin, 2012): *validez del constructo*, *validez interna*, *validez externa* y *validez de conclusión*. Por cada tipo de validación, evaluamos cada amenaza asociada y verificamos su presencia, y si aplica, describimos cómo la hemos evitado o minimizado.

- T⁴. Ejecución de la investigación. – Ejecutar el estudio empírico según el procedimiento sistemático definido en la T².
- T⁵. Análisis de los datos. – Analizar los datos producidos del estudio empírico aplicando el diseño de la inferencia especificado en la T².

2.3.3 Ciclos aplicados a la Tesis

En el desarrollo de esta Tesis, aplicamos un ciclo de diseño y tres ciclos empíricos. Para mejor explicación de la aplicación de estos ciclos, estructuramos los ciclos en forma de lista como lo muestra la Imagen 2.5, indicando las tareas de cada ciclo (con la letra T o Tⁿ), las actividades dentro de cada tarea y el capítulo o sección de la Tesis (columna “Capítulo/sección”) donde se desarrollan dichas actividades.

Como habrá notado el lector, la estructura de este documento de Tesis está basada en la secuencia de tareas de los ciclos de la metodología DS. Todos los capítulos de la Tesis, exceptuando los capítulos 1 y 2, están relacionados con las tareas de los ciclos: la T1 del ciclo de diseño se trata en los capítulos 3 y 4; la T2, en el 5 y 6; la T3 que involucra los tres ciclos empíricos, en el 7. Los capítulos 1 y 2 del documento de Tesis no se corresponden con ninguna parte de los ciclos ya que estos se enfocan en brindar una introducción al documento de Tesis y a explicar a fondo la metodología de investigación, que es el capítulo en el cual nos encontramos al momento.

Siguiendo con la Imagen 2.5, iniciamos con la tarea T1 del ciclo de diseño investigando el problema, acto seguido realizamos la tarea T2, el diseño del tratamiento, y en la tarea T3, que corresponde a la validación del tratamiento, realizamos tres ciclos empíricos: el primero valida los patrones, el segundo el proceso de diseño y el tercero es un caso de estudio donde se aplica GenomIUm en un caso real. Las tareas que componen los ciclos son descritas a continuación.

En la *investigación del problema* (T1), el objetivo es investigar, definir y explicar el problema utilizando varios métodos de investigación (p. ej., entrevistas, estudios de literatura, encuestas, observación de casos de estudio, experimentos de casos simples o experimentos de diferencia estadística, casos de estudio observacionales). En particular, hemos realizado estudios de la literatura, observaciones de casos de estudio, así como también, encuestas y entrevistas a usuarios del dominio. Como resultado de la investigación del problema se obtienen productos como: el *estado de arte* (es decir, las diferentes soluciones existentes y actuales que han sido propuestas para resolver el problema), el marco de trabajo conceptual (conjunto de definiciones que son la base para definir el problema y su solución), y un primer esbozo de los requisitos que dan forma al método de diseño GenomIUm, nuestro artefacto.

El *diseño del tratamiento* (T2) consiste principalmente en tres acciones: 1) definir los requisitos de la solución; 2) decidir la estrategia de solución al problema bajo estudio considerando los requisitos definidos; 3) diseñar la solución. Así como en la IS una especificación de requisitos describe el comportamiento deseable de un software, los requisitos definidos en esta tarea describen el comportamiento deseable tanto del tratamiento como del artefacto que se intenta proveer. Con estos requisitos, evaluamos la factibilidad de adaptar una solución existente o construir una nueva. En nuestro caso hemos

decidido crear una nueva solución: el método GenomIUm. Tras la decisión, procedemos a diseñar el tratamiento y el artefacto para abordar el problema investigado.

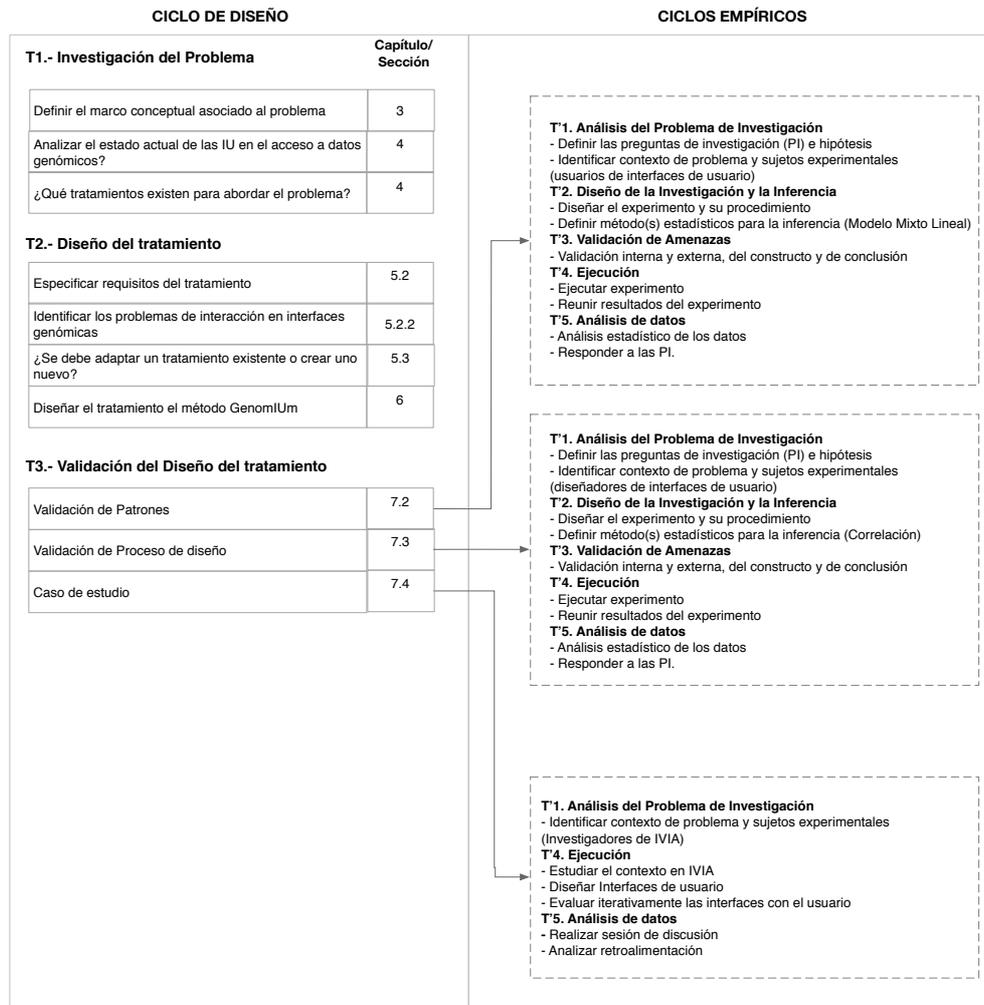


Imagen 2.5 Ciclos aplicados a la Tesis

La *validación del tratamiento* (T3) consiste en justificar que GenomIUm cumple con los requisitos de las partes interesadas antes de ser implementado. Ya que el método GenomIUm está compuesto de un conjunto de patrones y un proceso de diseño, validamos cada una de sus partes y posteriormente validamos los efectos producidos por el método en un ambiente real. Para esto, realizamos tres ciclos empíricos correspondientes a los siguientes estudios: dos experimentos controlados que validan los patrones y el proceso de

diseño, respectivamente, y un caso de estudio donde el método GenomIUm es aplicado en un ambiente real.

Los resultados empíricos de la tarea de validación son evidencias de valor para considerar a GenomIUm como una relevante contribución a la comunidad científica.

2.4 CONCLUSIONES

Este capítulo ha descrito la metodología de investigación *Design Science (DS)* que guía la presente Tesis. Como parte de la metodología aplicada a esta Tesis se exponen:

- Los objetivos de la investigación.
- Las preguntas de investigación derivadas de los objetivos.
- El marco de trabajo propuesto por la metodología para lograr los objetivos de la Tesis.
- Las tareas del Ciclo de Diseño y el Ciclo Empírico para llevar a cabo la investigación.

Como conclusiones importantes de este capítulo se destacan:

- El marco de trabajo aplicado a la Tesis consta de dos contextos que interactúan con el proyecto de investigación: a) el *contexto social* está representado por las partes interesadas en el proyecto de investigación (PROS, SENESCYT, laboratorios genéticos, investigadores, diseñadores de interfaces de usuario) y b) el *contexto de conocimiento* está representado por las teorías, técnicas, principios alrededor del campo de interacción humano-computador, modelado conceptual, genética, usabilidad, etc.
- El proyecto de investigación se desarrolla a lo largo de un ciclo de diseño y tres ciclos empíricos, aplicando métodos de investigación como entrevistas, observaciones de casos, experimentos, etc.

**PARTE II.
INVESTIGACIÓN DEL
PROBLEMA**

Capítulo 3

Marco Conceptual

Este capítulo define los conceptos, conocidos también como “constructos” (Wieringa, 2014), inmersos en el ámbito del método GenomIUm y su contexto. El marco de trabajo conceptual crea un acuerdo sobre los conceptos y definiciones utilizados en esta Tesis, evitando de esta manera ambigüedades y facilitando la comprensión del problema y la estructura de la solución propuesta: el método GenomIUm.

3.1 UN MARCO CONCEPTUAL PARA LOS DATOS GENÓMICOS

La genómica¹ es el estudio del genoma, es decir, el estudio de la secuencia de ADN que contiene el conjunto completo de genes de un organismo o una especie. La secuencia de ADN es una secuencia de cuatro bases bioquímicas (*Adenina*, *Citosina*, *Guanina* y *Timina*) representadas por cuatro letras (A, C, G, T) respectivamente, conocidas como el *código genético*, que se encuentra en los cromosomas contenidos en el núcleo de las células. En el caso del ser humano, por ejemplo, el genoma humano se refiere a la secuencia de ADN de los cromosomas contenidos en el núcleo de cada célula humana.

Estudiar el genoma de cualquier especie involucra estudiar cantidades masivas de datos relacionados al ADN provenientes de fuentes heterogéneas en origen y formato. Nos referimos a este tipo de datos como *datos genómicos* y las fuentes de donde provienen incluyen:

- *Secuenciación masiva del ADN*.- Masiva cantidad de datos generados por las tecnologías de Secuenciación de Segunda Generación (Metzker, 2010) (NGS, por sus siglas en inglés) que secuencian el ADN de especies a gran escala y producen ficheros de texto plano de formato específico (p.ej., ficheros VCF (Samtools, 2016))
- *Literatura biomédica*. - Bibliografía (p.ej. artículos, revistas, reportes clínicos) almacenada en repositorios digitales (p.ej., PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>)).
- *Bases de datos biológicas*. - Bases de datos de secuencias de elementos genéticos.

La Tabla 1 muestra algunas de las bases de datos existentes indicando el elemento genético que abordan.

¹ <https://www.genome.gov/genetics-glossary/genomics>

Tabla 1 Ejemplos de bases de datos biológicas

Bases de datos	URL	Elemento genético abordado
Nucleotide (NCBI)	https://www.ncbi.nlm.nih.gov/nucleotide/	nucleótidos
EMBL-EBI	https://www.ebi.ac.uk/ena	
UniProt	https://www.uniprot.org	proteínas
Ensembl	http://www.ensembl.org	genomas
KEGG	https://www.genome.jp/kegg/	rutas metabólicas
OMIM	https://www.omim.org/	enfermedades humanas
Gene Ontology	http://geneontology.org/	funciones de los genes

Para entender y contextualizar los datos genómicos, es necesario entender los conceptos biológicos complejos e inmersos en el dominio genómico y cómo estos conceptos están relacionados. Para esto, el modelo conceptual se convierte en una herramienta útil ya que a través del modelo es posible describir y hacer explícito el conocimiento de cualquier dominio. En este sentido, el Modelo Conceptual del Genoma Humano (MCGH) (Reyes R., 2018) fue concebido para abordar la complejidad del dominio genómico, especificando visualmente sus conceptos y las relaciones entre ellos. Aunque el MCGH está dirigido al ser humano como especie, el modelo sirve como base para generalizar los conceptos generales del genoma. El MCGH es un modelo propuesto por el grupo de investigación PROS y, aunque no es la única propuesta de modelado del genoma (Ram et al., 2004), (Bornberg-Bauer et al., 2002), (Bernasconi et al., 2017), el MCGH destaca su valor por erigirse desde una perspectiva holística (global) y de “compromiso ontológico” (Jacquette, 2002); un acuerdo compartido entre los implicados en el dominio en relación con los términos y su significado utilizados en el discurso.

El MCGH es un modelo en continua evolución marcada por un proceso sostenido de investigación de nuevos conceptos en el dominio que incluye la colaboración de expertos en el dominio en extensas jornadas de modelamiento conceptual que conducen a la agregación consensuada de los conceptos a la estructura conceptual y, consecuentemente, a la consecución de un modelo de conocimiento refinado, sólido y holístico. La evolución del MCGH ha pasado por las versiones 1 y 1.1 hasta consolidarse actualmente en la segunda versión (MCGH v2) (Reyes R., 2018). Esta versión define los elementos implicados en el contexto genómico a través de seis vistas:

- Vista Estructural. - Muestra la estructura del genoma incluyendo sus componentes (por ejemplo, cromosomas), considerando el tipo de segmentos de ADN del cual está conformado.
- Vista de Transcripción. - Muestra los componentes y conceptos involucrados en el proceso de transcripción que conduce, a través de la correspondiente traducción, a la “síntesis de proteínas”, es decir, la manera en la que los genes expresan su funcionalidad a través de la producción de proteínas.
- Vista de Variaciones. - Muestra las diferencias encontradas en la comparación de secuencias de ADN.
- Vista de Rutas metabólicas. - Muestra los componentes involucrados en las reacciones químicas que ocurren dentro de una célula.

- Vista de Fuentes de Datos y Bibliografía. - Muestra las fuentes de datos desde donde se extrae información para almacenar en el modelo y enlaces a documentos clínicos que respaldan la información almacenada.

Cada vista presenta una parte del modelo conceptual total y detalla los componentes involucrados en la vista. Por ejemplo, la Imagen 3.1 presenta la vista estructural del MCGH donde la clase “CHROMOSOME” hace referencia a los cromosomas de una especie (clase “SPECIES”), es decir, a un grupo de organismos con ADN semejante. Cada cromosoma puede estar caracterizado por su patrón de bandas (clase “CYTOBAND”) que representa las regiones dentro de la estructura del cromosoma que son visibles gracias a las técnicas de tinción (técnicas que mejoran el contraste de la imagen vista por microscopio y que son aplicadas en una fase específica del ciclo celular). La clase “HOTSPOT” representa los lugares en la secuencia de ADN donde puede producirse con mayor probabilidad el entrecruzamiento entre los cromosomas de los progenitores (padre y madre) para formar un par de cromosomas durante el proceso de división celular (o fase de meiosis). La clase “CHROMOSOME_ELEMENT” representa los fragmentos significantes dentro de la secuencia de ADN del cromosoma (p. ej., gen, exón, etc.). La clase “VARIATION” representa a una *variación genética* o alteración en la secuencia de ADN de un cromosoma y puede ser la causante de enfermedades o de rasgos característicos como, por ejemplo, el color de ojos.

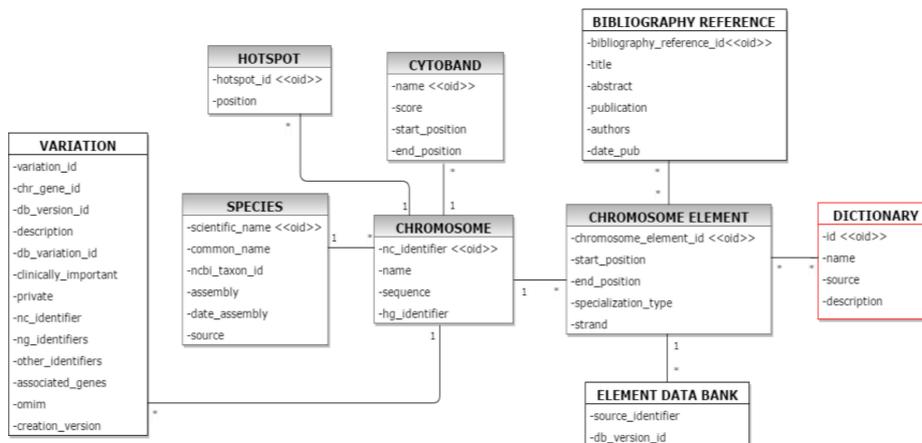


Imagen 3.1 Vista estructural del MCGH versión 2 (MCHG v2). Fuente: (Reyes R., 2018)

Dado que la información de los elementos del cromosoma es extraída de diversas fuentes de datos, la clase “ELEMENT DATA BANK” representa la relación de cada elemento cromosómico con las fuentes de datos de donde se ha extraído su información. La literatura o publicaciones (bibliografía) cuyo contenido refiere al elemento cromosómico está representada por la clase “BIBLIOGRAPHY_REFERENCE”. La clase “DICTIONARY” representa un glosario de términos que ha sido incluido en el modelo para solucionar el problema de que un término puede servir para describir muchos

conceptos. De esta manera, esta clase proporciona una breve descripción del término y la fuente bibliográfica de donde se ha obtenido la información.

Las vistas del MCGH están conectadas entre sí. Una vista puede contener clases que son particulares de la vista y otras clases que provienen de otras vistas, conectadas de una manera lógica. Estas conexiones reflejan las conexiones entre las vistas. En la vista estructural que mencionamos anteriormente (Imagen 3.1), por ejemplo, las clases con cabecera de color oscuro representan las clases particulares de la vista estructural (es decir, HOTSOPT, CYTOBAND, CHROMOSOME, SPECIES, CHROMOSOME ELEMENT), mientras que las clases con cabecera de color blanco (es decir, VARIATION, BIBLIOGRAPHY_REFERENCE, ELEMENT DATA BANK y DICTIONARY) representan a clases provenientes de otras vistas (p. ej., la clase “VARIATION” es una clase particular de la vista “variaciones”). De esta manera, la relación entre la clase CHROMOSOME y la clase VARIATION representa la conexión entre la vista “estructural” y la vista de “variaciones”. Tanto la vista estructural como las otras vistas, incluyendo la descripción de las clases y relaciones en cada una, están detalladas en (Reyes R., 2018).

Además de clarificar los elementos inmersos en el dominio, el MCGH se convierte en una sólida y coherente estructura para el diseño de interfaces de usuario orientadas a la explotación de los datos en este dominio. En efecto, el MCGH es un componente angular de esta Tesis ya que los diseños de interfaz que proponemos como soluciones a problemas de diseño de interfaz de usuario son explicados en términos de los conceptos y relaciones incluidos en el MCHG. Aunque el MCGH está dirigido al genoma humano, el modelo define los términos biológicos básicos inmersos en la genómica, por lo tanto, el MCGH puede utilizarse como marco de referencia para la estructura biológica de otras especies.

3.1.1 Análisis de datos genómicos

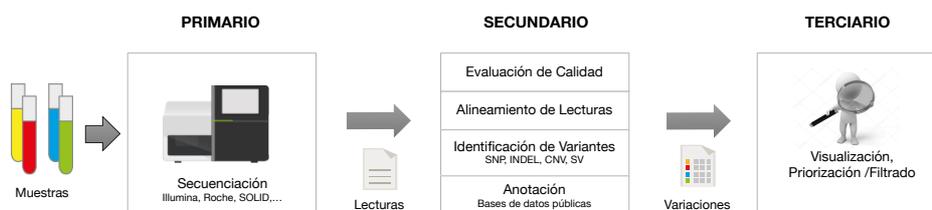


Imagen 3.2 Flujo de trabajo básico para proyectos de secuenciación del genoma, adaptado de Pabinger et. al. (Pabinger, 2014) y GATK (Broad Institute, 2019)

Cuando hablamos de analizar los datos genómicos, nos referimos desde una perspectiva general a un flujo de trabajo que incluye tres tipos de análisis (primario, secundario y terciario) y que tiene como punto de entrada la o las muestras que se desea analizar. Este flujo de trabajo se muestra en la Imagen 3.2.

En el *análisis primario* se obtienen masivas cantidades de datos genómicos derivadas de la secuenciación masiva a gran escala del ADN realizada con tecnologías NGS (p. ej., Illumina) dando como resultado el conjunto de “lecturas” de la secuenciación almacenado en ficheros

de texto. En el *análisis secundario*, se procesan las lecturas a través de una serie de transformaciones que permiten evaluar la calidad de las lecturas, alinear las lecturas, identificar variaciones genéticas, anotar variaciones genéticas (usando generalmente fuentes externas de información de acceso público). El resultado de este análisis es un fichero de texto con las variaciones genéticas identificadas.

Los análisis primario y secundario comparten desafíos de extracción y transformación de datos; desafíos que han sido bastante explotados al punto de disponer actualmente de herramientas potentes operadas en su mayoría por lenguaje de comandos que automatizan parcial o completamente los procesos de tratamiento de datos.

A diferencia del análisis primario y secundario, el *análisis terciario* tiene que ver con el desafío de extraer conocimiento de los datos; un desafío que involucra directamente a la interfaz de usuario ya que es en la interfaz donde el usuario interactúa (buscar, visualizar, filtrar) con los datos del análisis secundario para obtener conocimiento de ellos. En el análisis terciario, la interfaz de usuario debe proveer mecanismos apropiados e intuitivos para facilitar a los investigadores el acceso al conocimiento encerrado en los datos genómicos. Son cada vez más evidentes en la literatura las propuestas de herramientas para “priorizar variaciones genéticas”, muchas de ellas asegurando que implementan interfaces de usuario fáciles de usar y aprender y que facilitan la exploración, filtrado y curación de los datos.

La investigación y los resultados de esta Tesis se ubican en el contexto del análisis terciario.

3.1.2 Priorización de variaciones genéticas

La priorización de variaciones genéticas es un paso esencial en el análisis genómico y consiste en filtrar grandes cantidades de variaciones genéticas para identificar un subconjunto relevante de variaciones candidatas para un estudio específico (Dashti et al., 2017).

La priorización de variaciones genéticas se ubica en el análisis terciario del análisis genético y es transcendental en diferentes campos de investigación. En la investigación médica, por ejemplo, la priorización de variaciones genéticas consiste en analizar un extenso número de variaciones genéticas e identificar, a partir de la información asociada a las variaciones, aquellas que son causantes de una enfermedad bajo estudio (Roca et al., 2018).

La información asociada a la variación genética proviene de diferentes fuentes (literatura biomédica, bases de datos de biológicas, etc.) e incluye información de metadatos de la variación (nombres e identificadores), referencia a reportes clínicos relacionados con la variación, información de la relación de la variación con diferentes fenotipos, entre otras. Es por eso por lo que, filtrar y determinar las variaciones relevantes para un estudio se convierte en un trabajo complejo y demandante que generalmente está dirigido a especialistas bioinformáticos.

En este contexto, interfaces de usuario diseñadas para acceder a la información de las variaciones, visualizar el conjunto de variaciones y aplicar condiciones de filtro para segregar y extraer las variaciones más relevantes para el estudio, son de mucha ayuda para los especialistas bioinformáticos.

En esta Tesis, proponemos un método para diseñar dichas interfaces de usuario para la priorización de variaciones genéticas. Para ilustrar el método propuesto en esta Tesis, presentamos un caso de uso (Sección 7.4) donde aplicamos el método para diseñar interfaces de usuario para la priorización de variaciones genéticas en el contexto del genoma de los cítricos.

3.2 UN MARCO CONCEPTUAL PARA EL DISEÑO DE INTERFAZ DE USUARIO

3.2.1 La interfaz de usuario

La ISO: 9241-2010:2010 define a la interfaz de usuario como *todos los componentes de un sistema interactivo (software o hardware) que proveen información y controles al usuario para cumplir tareas específicas con el sistema interactivo*. La interfaz de usuario es un elemento crucial de todo sistema interactivo ya que la interfaz de usuario se convierte en la superficie de contacto en la cual el usuario interactúa con los datos del sistema. Esta superficie de contacto refleja las propiedades físicas de los participantes en la interacción, hace intuitivas las funciones que se pueden realizar y da al usuario la sensación de control sobre tales funciones (Laurel et al., 1990).

En el dominio genómico, hemos definido a la **interfaz de usuario de acceso a datos genómicos** como *la superficie de contacto donde el usuario interactúa con el sistema para obtener conocimiento de los datos genómicos*.

En la definición de la ISO existe varios términos (p. ej., “usuario”, “tarea”, “sistema interactivo”) que es necesario definir y contextualizar para entender el concepto de interfaz de usuario en el dominio genómico. A continuación, definimos estos y otros términos (derivados de los primeros) y contextualizamos aquellos que tienen mayor relevancia en nuestro dominio de estudio.

Usuario. – El término se refiere a *la persona quien usa la interfaz de usuario* que en el dominio genómicos puede ser un investigador del área de la genética, medicina, biología, bioinformática.

Tarea. - La ISO: 9241-2010:2010 define a una tarea como *las actividades requeridas para lograr un objetivo*. En el dominio genómico, las tareas son todas aquellas actividades que son realizadas por el usuario y que son requeridas para llevar a cabo un objetivo de nivel superior y relacionado al análisis y extracción de conocimiento de los datos genómicos.

Sistema interactivo. – Un sistema interactivo, o simplemente “sistema”, está definido por la ISO: 9241-2010:2010 como *la combinación de hardware, software y/o servicios que recibe información de los usuarios y la comunica a los usuarios*. En el contexto genómico, un sistema interactivo está representado por un sistema de gestión de datos genómicos que implementa interfaces de usuario que son utilizadas por el usuario para extraer conocimiento de los datos.

Interacción. – Este término hace referencia al *diálogo usuario-sistema entendido como la manera en la que el usuario navega, explora, filtra o manipula los datos de un sistema genómico para obtener conocimiento*. Entender el concepto de interacción es básico para el diseño de la interfaz. Diseñar la interacción significa diseñar productos interactivos para apoyar a las personas en

su vida cotidiana y laboral (Preece, 2015). Considerando que el interfaz de usuario es un producto interactivo, diseñar la interfaz de usuario significa diseñar la interacción. Así, la preocupación central del diseño de la interacción es desarrollar interfaces de usuario usables, es decir, interfaces que sean “intuitivas” y efectivas y que provean una *experiencia de usuario* satisfactoria.

Interfaz Intuitiva. - Con el término “intuición” nos referimos al “aprendizaje”. Una interfaz intuitiva *es aquella que es fácil de aprender*. La facilidad de aprender una interfaz de usuario sucede en gran medida por dos razones: 1) porque los mecanismos en la interfaz son tan evidentes que guían al usuario en lo que tiene que hacer o 2) porque el usuario está ya familiarizado con la manera con la que tiene que usar la interfaz debido a experiencias similares previas. Estas dos razones están cubiertas bajo el término *affordance* utilizado por Norman (Norman, 2013) para definir la “relación” entre las propiedades de un objeto y las capacidades de la persona que interactúa con dicho objeto. En efecto, la primera razón tiene que ver con las propiedades de la interfaz de usuario y la segunda razón con las capacidades del usuario. Esta relación (o *affordance*) es la que determina la manera en la que el objeto puede ser utilizado y contribuye a que el uso del objeto sea fácil de aprender. En consecuencia, si para un usuario determinado son desconocidos los mecanismos que incorpora una interfaz, intuir las acciones que debe hacer en la interfaz será una tarea difícil para dicho usuario.

Experiencia de usuario. – La ISO: 9241-2010:2010 define a la experiencia de usuario como *las percepciones y respuestas de las personas resultantes del uso y/o uso anticipado de un producto, sistema o servicio*.

3.2.2 Diseño conceptual de la interfaz de usuario

El diseño conceptual de la interfaz de usuario es un *diseño inicial (prototipo) de alto nivel de la interfaz de usuario que captura la estructura y el flujo de trabajo de la misma* (Seffah et al., 2005) (Wallach et al., 2012). Por ejemplo, la Imagen 3.3, presenta el diseño conceptual de la interfaz de usuario de la intranet de la Universidad Politécnica de Valencia (<https://intranet.upv.es>).

El diseño conceptual sienta las bases de cómo va a ser la interfaz de usuario y es importante en el proceso de diseño de la interfaz de usuario porque es utilizado por el diseñador/desarrollador para explorar, comunicar y evaluar sus ideas con los usuarios finales y se convierte en un instrumento para la evaluación de usabilidad de la interfaz de usuario.

Mayhew (Mayhew, 1999), ubica al diseño conceptual en la primera fase del ciclo de vida de la ingeniería de usabilidad donde es utilizado para obtener retroalimentación del usuario antes de iniciar la implementación (codificación) de la interfaz de usuario. En otras palabras, el diseño conceptual de la interfaz de usuario es útil como instrumento de evaluación de la usabilidad de la interfaz de usuario y guía el refinamiento de la interfaz de usuario con el propósito de obtener, a partir de un diseño conceptual básico, un diseño de interfaz de usuario completamente detallado. Por ejemplo, la Imagen 3.4. muestra el diseño conceptual refinado tras varias iteraciones del diseño conceptual presentado en la Imagen 3.3.

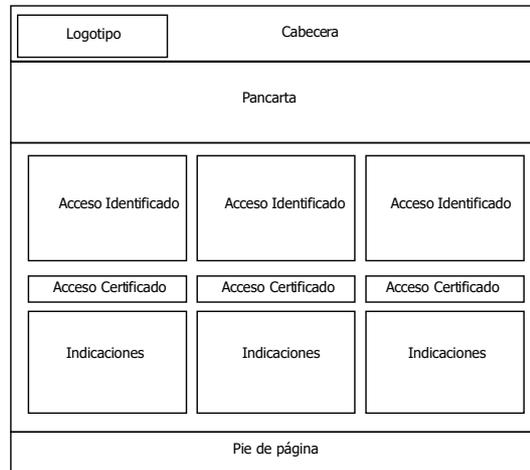


Imagen 3.3 Ejemplo de diseño conceptual básico



Imagen 3.4 Diseño conceptual de la Interfaz refinado

3.2.3 Patrones

Origen

El término *patrón* tuvo sus orígenes en el área de la Arquitectura a través de los libros escritos por el arquitecto Christopher Alexander al final de los años 70. El libro *A Pattern*

Language (Alexander et al., 1977) de Alexander recoge las experiencias de ocho años en entender la naturaleza del proceso de construcción de edificios, pueblos y casas a través de un total de 253 patrones. A través del término *patrón*, Alexander intenta proporcionar recomendaciones y buenas prácticas para el diseño y construcción de edificaciones que puedan ser entendidas por personas no expertas en diseño ni en arquitectura, con el objetivo de generar una mejor experiencia de quienes las habitan.

La intención detrás de la idea de Alexander inspiró el uso del término en otros ámbitos. Es así como, desde su origen en el dominio de la construcción, el término patrón ha sido aplicado para solucionar problemas de diseño comunes en varios dominios de aplicación y el diseño de la interfaz de usuario es uno de estos dominios.

A finales de los 80's, Howard Cummingham (creador del WikiWikiWeb, el primer sitio wiki) y Kent Beck (uno de los creadores de la programación extrema o XP) adaptan el término *patrón* al diseño de interfaces y crean cinco patrones de interfaz para Smalltalk (uno de ellos es *Window per task*). Mas tarde, Jeniffer Tidwell (Tidwell, 1999) y Martin Van Welie (Van Welie et al., 2003) desarrollan colecciones de **patrones de interacción** para interfaces de usuario en la web.

Actualmente, la lista de patrones de diseño de interacción publicados es tan larga como los problemas que estos resuelven, incluyendo el cómo organizar el contenido en la interfaz de usuario, cómo capturar la entrada de datos del usuario utilizando formularios y controles o cómo presentar y navegar por los datos.

Definición

Alexander definió el término patrón en función de lo que hace: “... describe un **problema** que ocurre una y otra vez en nuestro **entorno**, y entonces describe el núcleo de la **solución** a ese problema, de tal manera que se pueda usar esta solución un millón de veces, sin hacer lo mismo dos veces”, página x del libro *A Pattern Language* (Alexander, 1977).

Para ilustrar la definición del término acuñado por Alexander, hemos escogido al azar el patrón “Transición de entrada” (“**Entrance transition**”, patrón 112, pág. 552 del libro *A Pattern Language*). El problema que trata el patrón Transición de entrada es *diseñar un cambio o transición de ambientes entre la parte pública y privada de una edificación*. El humano no percibe una buena sensación cuando la entrada de la casa es directa desde la calle, es decir, no existe un cambio de ambiente entre la calle y la casa.

La solución que describe Alexander a este problema es “*Crear un espacio de transición entre la calle y la puerta de entrada. Une el camino que conecta la calle y la entrada a través de este espacio de transición, y márkelo con un cambio de luz, un cambio de sonido, un cambio de dirección, un cambio de superficie, un cambio de nivel, quizás por pasarelas que hagan un cambio de recinto, y sobre todo con un cambio de vista.*”. Tal solución está acompañada de una descripción gráfica, como lo muestra la Imagen 3.5.

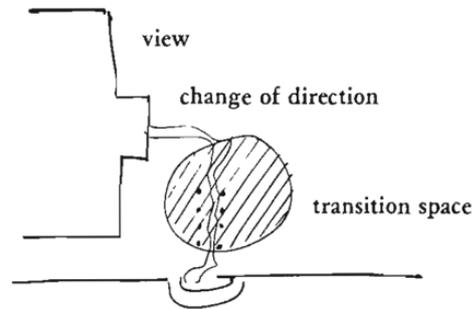


Imagen 3.5 Descripción gráfica del patrón “Entrance transition”. (Alexander, 1977)

En la Imagen 3.6 mostramos ejemplos reales del patrón “Transición de Entrada”. La Imagen 3.6a es un ejemplo donde el patrón no está aplicado; no existe transición entre la calle y el interior de las casas. La figura Imagen 3.6b muestra el patrón aplicado a través de las gradas que crean un cambio de nivel entre la calle y la puerta principal de la edificación.

El objetivo de Alexander no solo fue construir edificaciones que sean atractivas sino también, y más importante, edificaciones que aporten una buena experiencia a las personas. Es importante notar que la solución propuesta es de tipo conceptual, lo que implica que su implementación debe adaptarse a la situación particular, lo cual convierte a los patrones en herramientas potentes para el diseño de soluciones.



a)



b)

Imagen 3.6 Ejemplo donde a) no se cumple el patrón “Entrance Transition” y b) donde se cumple el patrón.

A diferencia de los patrones de diseño en el dominio de la Arquitectura, los patrones de diseño en el dominio del Diseño de Interfaces abordan problemas de interacción entre el usuario y la interfaz, y proveen una descripción estructurada de un concepto que expresa una solución no trivial al problema (Granlund et al., 2001). En la literatura, este tipo de patrones es conocido bajo varios nombres: *patrones de diseño de interfaz de usuario* (Toxboe, 2007), *patrones de diseño de interacción* (Borchers, 2000), *patrones de Interfaz Humano-Computador* (Granlund, 2001), o simplemente como *patrones de diseño* (Yahoo, 2006). Para simplificar la diversidad de nombres, nos referimos a este tipo de patrones simplemente como *patrones*. Por lo tanto, en el dominio del diseño de la interfaz de usuario:

Patrón es la descripción estructurada de una solución recurrente a un problema común de diseño de la interfaz de usuario, dentro de un contexto específico.

Formato

El formato original utilizado por Alexander para describir patrones en el dominio de la Arquitectura consta de las siguientes partes (Alexander, 1977):

- *Nombre.* - Nombre corto para identificar al patrón.
- *Imagen.* - Muestra un ejemplo arquetípico del patrón.
- *Introducción.* - Párrafo que establece el contexto para el patrón, explicando cómo el patrón ayuda a completar otros patrones.
- *Titular del problema.* - Una o dos oraciones, en letra negrita, que describen la esencia del problema.
- *Cuerpo del problema:* Sección más larga que describe el trasfondo empírico del patrón, la evidencia para su validez, el rango de maneras diferentes en que el patrón puede ser manifestado en una construcción.
- *Solución.* - Es el núcleo del patrón y describe el campo de las relaciones físicas y sociales que son requeridas para resolver el problema indicado, en el contexto indicado. Esta solución está redactada en la forma de una instrucción, así el diseñador sabe exactamente qué necesita hacer para construir el patrón. Estas instrucciones no son “imposiciones” sino más bien “indicaciones” de los aspectos esenciales que se deben considerar para solucionar el problema. Estas indicaciones son expresadas de una manera muy general y abstracta, dejando al lector solucionar el problema por sí mismo y a su propia manera, adaptando la solución a sus preferencias y a las condiciones locales. La descripción de la solución incluye las relaciones esenciales y necesarias para solucionar el problema.
- *Diagrama.* - Muestra la solución en forma de un diagrama con etiquetas para indicar sus componentes principales.
- *Relaciones.* - Párrafo que describe los patrones que están inmersos y son necesarios para la conformación del patrón actual.

Alexander también señala dos propósitos fundamentales directos del formato: 1) presentar cada patrón conectado a otros patrones de tal manera que sea fácil para el lector navegar por todos los 253 patrones del conjunto y crear todas las posibles combinaciones, y 2) presentar el problema y la solución de cada patrón de tal manera que el lector pueda por sí mismo modificarlo sin perder la esencia central del patrón.

Por otro lado, cada patrón tiene varios niveles de significancia. Alexander define un sistema de puntuación basado en marcas:

- Dos asteriscos (**). – Las propiedades constantes y comunes para todas las maneras posibles de resolver el problema han sido completamente identificadas y resumidas.
- Un asterisco (*). – Las propiedades constantes y comunes para todas las maneras posibles de resolver el problema han sido parcialmente identificadas, pero la solución propuesta puede ser mejorada. Con “parcialmente” se refiere a que a pesar de que se conoce que existen soluciones al problema dado, estas soluciones no están cubiertas en la redacción del patrón.
- Sin asterisco. – Las propiedades constantes y comunes para todas las maneras posibles de resolver el problema no han sido identificadas. No se ha conseguido identificar las propiedades constantes y comunes. Existen ciertamente maneras de resolver el problema de diferente manera a la que se ha escrito. En estos casos se indica la solución para proveer al lector con al menos una manera de resolver el problema.

Ventajas del uso de patrones

Los patrones proveen mayores beneficios que las directrices o guías de estilo para diseñar la interfaz de usuario (Seffah, 2015). A diferencia de los directrices o guías de estilo, los patrones explican a menudo los pasos o reglas que deben ser aplicadas como solución al problema. Entre los principales beneficios de usar patrones se menciona:

- *Transferencia de conocimiento.* – Los patrones son herramientas para capturar y disseminar el conocimiento y experiencias de diseño.
- *Guías de Usabilidad.* – Ayudar a los desarrolladores en el desarrollo de sistemas interactivos más usables.
- *Diseño de la Interacción.* – Los patrones son herramientas útiles para diseñar la interacción. Los patrones se convierten en herramientas poderosas para diseñar la interacción (el diálogo usuario-sistema) que debe satisfacer la interfaz de usuario. Cada patrón provee una solución de diseño de interfaz de usuario descrita de forma textual o gráfica que resuelve un problema de interacción en la interfaz de usuario.
- *Comunicación.* – Los patrones se convierten en una *Lingua Franca* la cual es fácil de entender por diseñadores, desarrolladores y usuarios, independientemente del conocimiento y experiencia.
- *Reutilización.* – Los patrones evitan que el diseñador pierda tiempo *reinventando la rueda* para resolver un problema conocido ya que las soluciones provistas por los patrones son el resultado del intercambio de experiencias entre diseñadores que se han enfrentado a un mismo problema. Por lo tanto, los patrones reúnen un conjunto de buenas prácticas que pueden ser reutilizadas para enfrentar un problema de diseño en la interfaz de usuario.

3.3 CONCLUSIONES

En este capítulo hemos definido el conjunto de conceptos, o *constructos*, sobre los cuales se fundamenta el método GenomIUm propuesto en esta Tesis. El marco conceptual está dividido en dos grupos conceptuales: 1) los datos genómicos y, 2) el diseño de interfaces de usuario.

El primero define los conceptos relacionados con los datos genómicos y su contexto. Estos conceptos son claves para entender el dominio genómico incluyendo los conceptos biológicos de la estructura genética y los conceptos relacionados con las fuentes de datos inmersas en el dominio.

El segundo grupo aborda los conceptos inmersos en el diseño de interfaces tales como interfaz de usuario, usuario, tarea, sistema interactivo, interacción, interfaz intuitiva, experiencia de usuario, diseño conceptual de la interfaz, patrones de interacción, que son claves para entender la estructura del método GenomIUm.

Los conceptos descritos en el marco conceptual son utilizados en el Capítulo 6, donde se describe en detalle el método GenomIUm y el catálogo de soluciones que lo soporta, y el Capítulo 7 donde se describe la validación del método.

Capítulo 4

Estado del Arte

Este capítulo presenta el estudio acumulado en el área del diseño de interfaces de usuario en el dominio genómico, basado en la búsqueda y revisión de literatura. El contenido de este capítulo contextualiza la investigación y ayuda a comprender las razones que motivan la decisión de proponer el método GenomIUm.

Dado que la solución que se propone en esta Tesis está basada en la definición y uso de patrones de diseño de interfaz de usuario, en la Sección 4.1 analizamos los enfoques de diseño soportados por patrones y en la Sección 4.2 analizamos los patrones existentes para el diseño de interfaces en el dominio genómico.

4.1 ENFOQUES DE DISEÑO DE INTERFACES BASADOS EN PATRONES

4.1.1 DCU

El Diseño Centrado en el Usuario (DCU) es una filosofía de diseño que aplicada al ámbito del diseño de interfaces de usuario tiene como objetivo diseñar interfaces de usuario de calidad, que sean fáciles de usar y aprender. El DCU se enfoca en primer lugar en las necesidades, capacidades y comportamiento del usuario para luego enfocarse en diseñar soluciones ajustadas a estas necesidades, capacidades y formas de comportarse (Norman, 2013). La Imagen 4.1 muestra el proceso de diseño del DCU descrito en la ISO 9241-210 como un ciclo iterativo de 4 actividades:

- *Actividad 1: Entender y especificar el contexto de uso:* Identificar a los usuarios para quienes estará dirigido el producto, para qué lo usarán y en qué condiciones.
- *Actividad 2: Especificar requisitos:* Identificar los objetivos del usuario que deberán satisfacerse.
- *Actividad 3: Producir soluciones de diseño:* Esta fase se puede subdividir en diferentes etapas secuenciales, desde el diseño de las primeras soluciones conceptuales de la interfaz de usuario hasta su diseño refinado y final.
- *Actividad 4: Evaluación:* Validar las soluciones de diseño, es decir, verificar si el producto satisface los requisitos o, por el contrario, se detectan problemas de usabilidad.

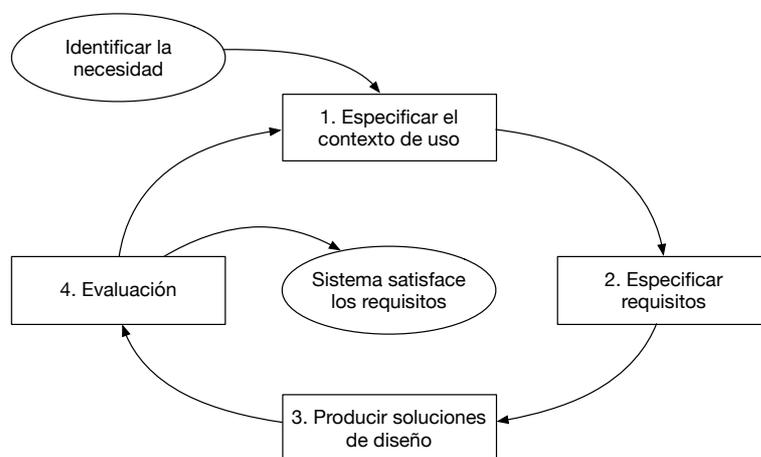


Imagen 4.1 Proceso de Diseño Centrado en el Usuario (ISO 9241-210)

El DCU es una metodología integral de desarrollo de sistemas interactivos que establece la ruta para ir desde la “investigación de usuario” (Actividad 1 y 2) hacia la “producción de soluciones” (Actividad 3 y 4) garantizando que estas soluciones se ajusten a las necesidades de los usuarios que las usan. Nosotros creemos que la aplicación del DCU puede mejorar sustancialmente la usabilidad de las interfaces de usuario de las aplicaciones en el dominio genómico. Los métodos y técnicas involucradas en el DCU son útiles para abordar las actividades del proceso, sin embargo, para la actividad de producción de soluciones (Actividad 3) los métodos recomendados (ISO TR 16982:2002) (p. ej., esbozo de interfaces, guías de estilo) y técnicas utilizadas (p. ej., diseño inicial de la interfaz de usuario, arquitectura de información y prototipos de interfaces) en el diseño y evaluación no están conectadas de una manera que guíe y facilite al desarrollador de software a producir soluciones.

El método GenomIUm que proponemos en esta Tesis está enmarcado dentro de la producción de soluciones del DCU y soporta específicamente el desarrollo de un artefacto clave creado en la Actividad 3: el *diseño conceptual de la interfaz de usuario* (ver Sección 3.2.2). Para esto, GenomIUm adopta el principio del DCU que motiva el uso de patrones de interfaz de usuario para la producción de soluciones. Es por eso por lo que GenomIUm propone un proceso de diseño y un catálogo de patrones para abordar el desarrollo del concepto de la interfaz de usuario.

4.1.2 GDD

Diseño Orientado por Objetivos (GDD por su nombre en inglés *Goal-Directed Design*) (Cooper et al., 2004) es una metodología centrada en el usuario que aborda el proceso completo de diseño de interfaz de usuario, desde la fase de investigación de usuario hasta la fase de diseño de soluciones; el proceso previo a la etapa de la construcción o implementación de la interfaz.

Como su nombre lo indica, la característica principal de GDD es dirigir el diseño de la interfaz de usuario a partir de los “objetivos” de las personas que usarán la interfaz. La metodología rompe la idea tradicional de diseñar la interfaz de usuario para cubrir tareas puntuales y aisladas. GDD se enfoca en los objetivos; un nivel más alto y sostenible en el tiempo ya que las tareas cambian como cambia la tecnología mientras que los objetivos permanecen sin cambios debido a que son independientes de la tecnología. Para GDD, los objetivos son el fin mientras que las tareas son el medio para llegar a este fin.

Desde una perspectiva general, el diseño de la interfaz de usuario puede ser visto como un proceso de dos actividades: la investigación de usuario y el diseño de soluciones. GDD resuelve el problema de convertir los requisitos (obtenidos en la actividad de investigación de usuario) en estructuras del diseño conceptual de la interfaz de usuario (en la actividad de diseño de soluciones), proponiendo un proceso sistemático de seis fases como se muestra en la Imagen 4.2.

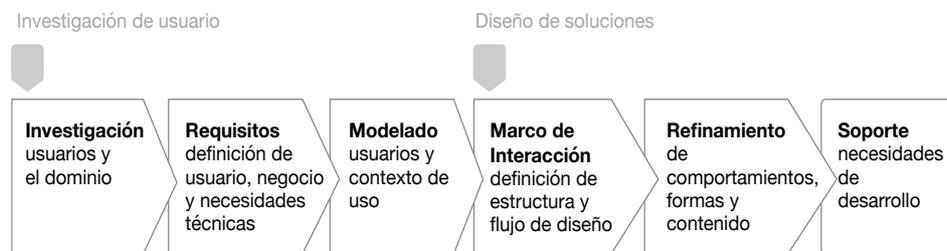


Imagen 4.2 Proceso de diseño de GDD, tomado y adaptado de (Cooper et al., 2007)

Además del proceso sistemático, GDD provee herramientas para soportar el proceso. Dos de las herramientas principales son las *Personas* y los *Patrones de Diseño de Interacción* y las dos están estrechamente relacionadas.

Las *Personas* son herramientas utilizadas por los analistas, desarrolladores y diseñadores para modelar a los usuarios que usarán la interfaz con el fin de tener una idea clara de las características de los usuarios y los objetivos que estos desean lograr. El concepto de *Personas* fue introducido en el libro *The inmates are running the asylum* (Cooper, 2004), y desde su uso tiene gran aceptación por la comunidad de Experiencia de Usuario (UX). Las *Personas* son arquetipos de usuarios reales, es decir, no son personas reales, pero representan a estas personas reales durante todo el proceso de diseño. Las *Personas* se crean en la fase de “Modelado”, reuniendo la información recabada en la fase de “Investigación”, pero su uso se extiende a lo largo de todo el proceso de diseño.

Los *Patrones de Diseño de Interacción* entran en escena en la fase “Marco de interacción”. En esta fase se crea el “marco de interacción”, que no es más que el diseño conceptual de la interfaz de usuario, definiendo la estructura de la interfaz y su flujo de interacción en función de las necesidades y características de los usuarios. Los desarrolladores crean el diseño conceptual de la interfaz de usuario definiendo aspectos de diseño visual (forma) y comportamiento utilizando patrones de diseño de interacción que son seleccionados en función de información de las *Personas*, es decir, en función de las necesidades y características de los usuarios y su contexto de uso.

Respecto al objetivo de esta Tesis, las conclusiones que podemos obtener del GDD son las siguientes:

- GDD cubre todo el proceso de diseño e introduce herramientas poderosas como las Personas y los Patrones de Diseño de Interacción. Mientras las Personas están pensadas para organizar la información obtenida en la investigación de usuario y representar al usuario durante todo el proceso de diseño, los Patrones de Diseño de Interacción están pensados para trasladar la información de las Personas a estructuras de la interfaz de Usuario. En este sentido, la influencia que tienen las Personas en la selección de los patrones es un punto para destacar del GDD y que puede ser de mucha utilidad tanto en la definición como en la selección de patrones para solucionar los problemas en el dominio genómico.
- A pesar de la evidente relación entre Persona y Patrones, la metodología no ofrece detalles de cómo llevar esta relación a la práctica. La metodología ofrece información suficientemente detallada sobre cómo construir la Persona, pero los detalles de cómo aplicar los patrones para construir la interfaz de usuario a partir de la información de la Persona no son abordados en profundidad.

4.1.3 PSA

El PSA (*Pattern Supported Approach*) (Granlund, 2001) es un enfoque para el proceso de diseño de la interfaz de usuario que amplía el uso de los patrones a todas las actividades del proceso del DCU (Imagen 4.1). El PSA revoluciona la idea de usar patrones exclusivamente en la actividad de diseño de la interfaz extendiendo el uso de los patrones a las actividades anteriores al diseño (es decir, definición del sistema, análisis de usuario y tareas).

El PSA propone utilizar un conjunto de patrones particular como soporte a cada actividad del proceso de diseño. Así, el conjunto de patrones en el PSA está categorizado en patrones de *dominio, procesos de negocio, tareas, diseño de navegación y estructura y diseño de interfaz gráfica de usuario*, como se muestra en la Imagen 4.3.

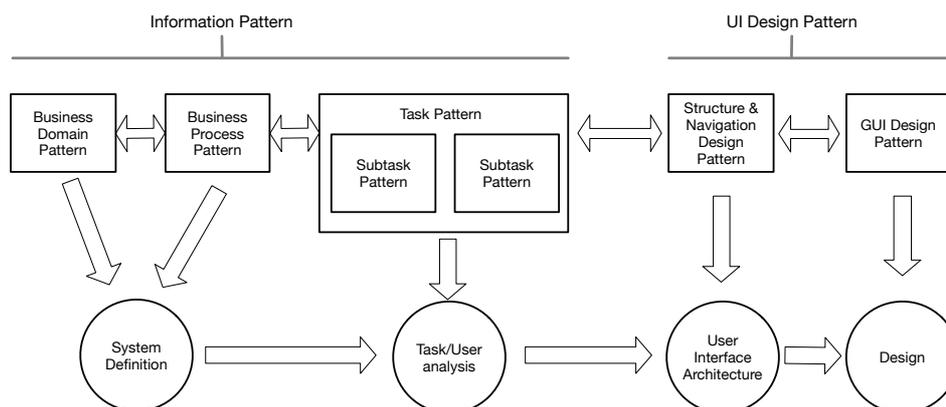


Imagen 4.3 Marco de trabajo del PSA. Fuente: (Seffah, 2015)

Las flechas de doble sentido entre las diferentes categorías de patrones indican que las categorías de patrones están relacionadas de manera bidireccional formando una red conectada de patrones. Por ejemplo, los patrones de tareas (Task Pattern) están relacionados bidireccionalmente con los patrones de estructura y navegación (Structure & Navigation Design Pattern) y estos a su vez relacionados bidireccionalmente con los patrones de interfaz gráfica (GUI Design Pattern). Es importante destacar que el PSA establece un punto de entrada a la red de patrones y sugiere al desarrollador un conjunto de patrones relacionados en las diferentes actividades del proceso de diseño.

De lo mencionado, delineamos tres conceptos valiosos que el PSA introduce en el proceso de diseño de la interfaz de usuario:

- Concepto 1: *Cobertura*. - El PSA no solo se centra en el diseño como actividad dentro del proceso de diseño de la interfaz de usuario sino también en las actividades anteriores al diseño como son la definición del sistema, la identificación de usuarios y tareas. Es decir, el PSA cubre todo el proceso de diseño de la interfaz de usuario.
- Concepto 2: *Patrones estructurados para cubrir el proceso de diseño*. – Para cubrir todo el proceso de diseño, el PSA utiliza un conjunto de patrones estructurado en categorías donde cada categoría soportan una actividad del proceso de diseño.
- Concepto 3: *Relaciones entre categorías de patrones*. - Las categorías de patrones que soportan cada actividad del proceso están relacionadas bidireccionalmente formando una red de patrones que el PSA utiliza para sugerir patrones adecuados en cada actividad del proceso.

Considerando la influencia de los conceptos del PSA en el propósito de esta Tesis (diseñar el concepto de la interfaz de usuario), mencionamos que:

- La cobertura del PSA, mencionada en el concepto 1, es muy amplia para nuestro propósito. A diferencia del interés del PSA en abordar todas las actividades del proceso de diseño, nuestro interés se centra en la actividad de diseño, donde se lleva a cabo el diseño conceptual de la interfaz y su refinamiento.
- El uso de los patrones como soporte al proceso de diseño, mencionado en el concepto 2, es una idea útil para nuestro propósito. Esto nos motiva a crear un conjunto de patrones estructurado en categorías y enfocado a resolver problemas de acceso a datos genómicos que, en lugar de soportar el proceso de diseño completo, soporte concretamente la actividad de diseño del proceso.
- La característica relacional del conjunto de patrones, mencionada en el concepto 3, es útil para nuestro propósito. Sin embargo, es importante mencionar que a pesar de que el PSA sugiere un conjunto de patrones para cada fase del proceso de diseño, el PSA no aborda el cómo usar las interrelaciones entre los patrones para enlazar los patrones al diseño conceptual de la interfaz (Seffah, 2010); un concepto importante para nuestro propósito.

4.1.4 POD

El Diseño Orientado por Patrones (POD) (Javahery et al., 2002) es un enfoque de diseño que está enmarcado en la actividad de producción de soluciones del proceso de diseño de interfaces de usuario y que provee un marco de trabajo para guiar al desarrollador paso a paso por el proceso de diseño conceptual de la interfaz de usuario a través de la combinación lógica de patrones. A diferencia del PSA, el POD aborda las interrelaciones entre los patrones y el cómo usarlas para enlazar los patrones al diseño conceptual de la interfaz de usuario. El POD emerge para soportar la idea de que para que los patrones sean útiles, los desarrolladores deben saber cómo combinarlos para crear diseños completos o parciales (Seffah, 2015). De esta manera, el POD aborda aspectos claves de la aplicación de patrones como *cuándo* son aplicables (en el proceso de diseño conceptual de la interfaz de usuario), *como* pueden ser usados y *por qué* pueden ser o no combinados con otros patrones. El POD se basa en dos consideraciones principales:

Primero, el desarrollador debe seguir un proceso que lo guía paso a paso en las decisiones de diseño. Por ejemplo, para el diseño web, Javahery define un proceso en 4 pasos: 1) definir la arquitectura, 2) establecer la estructura interna de la página, 3) identificar elementos de contenido, y 4) organizar la interacción entre el contenido.

Segundo, el desarrollador debe explotar las relaciones semánticas existentes entre los patrones. Javahery establece cinco categorías de relaciones entre patrones de diseño web (Javahery et al., 2007) que pueden ser aplicables en otras librerías de patrones. Este conjunto de relaciones está basada en la clasificación de relaciones entre patrones establecidas originalmente por (Zimmer et al., 1995) y se convierte en un mecanismo que guía a los desarrolladores en las decisiones de diseño, permitiendo hacer suposiciones como: “Para algún problema P, si aplicamos el Patrón X, entonces el patrón Y y Z se aplican como subordinados, pero el Patrón S no puede ser aplicado ya que este es un competidor”. Aquí resumimos el conjunto de relaciones considerando a X e Y como dos patrones en el extremo de una relación:

- **Similar.** X es similar a Y si los dos patrones pertenecen a la misma categoría y proveen diferentes soluciones a un mismo problema dentro de un contexto similar. En consecuencia, X e Y pueden ser reemplazados el uno por el otro y aún resolver el problema de diseño o pueden ser utilizados al mismo tiempo en un diseño como soluciones al problema de diseño.
- **Competidor.** X es competidor de Y si los dos patrones pertenecen a una misma categoría y no pueden ser utilizados al mismo tiempo en un diseño.
- **Superordinado.** X es superordinado de Y si Y sirve como bloque de construcción para crear X. Esta es una relación de composición del tipo “*todo-parte*” donde X es el *todo* y Y la *parte*.
- **Subordinado.** X es subordinado de Y si X puede ser incluido en Y. Y es también superordinado de X. Es decir, X es la *parte* y Y es el *todo*.
- **Vecino.** X es vecino de Y si los dos patrones pertenecen a la misma categoría.

Para ilustrar el POD, Javahery presenta un caso de estudio en el ámbito del diseño de interfaces web aplicando UPADE, un lenguaje de patrones de diseño web que está organizado en categorías de patrones que soportan las etapas del proceso de diseño web.

La Imagen 4.4 muestra un resumen del lenguaje UPADE organizado en tres categorías de patrones: 1) patrones de arquitectura 2) patrones estructurales 3) patrones de soporte a la navegación. Los patrones de arquitectura abordan la organización de todo el sitio.

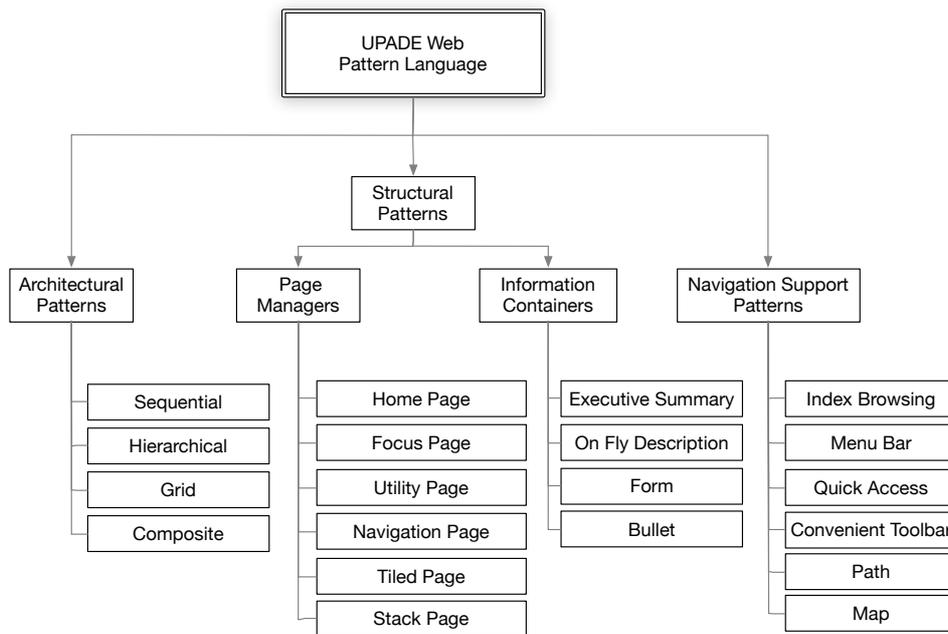


Imagen 4.4 Resumen del lenguaje de patrones UPADE. Fuente: (Seffah, 2015)

Los patrones estructurales se enfocan en soluciones para la estructura interna de la página web. Los patrones estructurales se dividen a su vez en patrones de administración de página, que se enfocan en la estructura de los elementos de la página, y los patrones contenedores de información que se enfocan en los elementos de contenido dentro de la página. El último grupo de patrones, los patrones de soporte a la navegación, sugiere diferentes modelos de navegación entre el contenido. En consecuencia, los desarrolladores inician el diseño de un sitio web estableciendo la arquitectura del sitio, respondiendo a preguntas como ¿Cuáles interfaces de usuario contendrá el sitio? ¿Cómo navegar entre ellas? Para esto, el desarrollador se apoya de los patrones de arquitectura.

En la siguiente etapa de diseño, los desarrolladores definen la estructura interna de cada interfaz de usuario apoyándose de los patrones de administración de página y de contenedores de información. Finalmente, los desarrolladores, apoyados de los patrones de soporte a la navegación, se enfocan en organizar la interacción entre los contenidos de la página y entre las páginas.

Por lo mencionado, consideramos que el POD es un enfoque adecuado sobre el cual fundamentar el método GenomIUM, por las siguientes razones:

- **Covertura.**- El POD se enfoca en el diseño conceptual de la interfaz de usuario, en lugar de todo el proceso de diseño como lo abordan el PSA o el DCU.
- **Uso de patrones.**- El POD usa un conjunto de patrones estructurados en categorías que soportan el proceso de diseño. Javahery ha utilizado el lenguaje de patrones Web UPADE para ilustrar el POD. La estructura de UPADE y sus patrones abordan el diseño web general, sin embargo, si se trata de abordar un proyecto de diseño web en un dominio específico, la estructura de UPADE debe ser adaptada para incluir categorías de patrones específicos del dominio. Bajo esta concepción, para proponer un catálogo de patrones que soporte el diseño de interfaces en el dominio genómico, pretendemos adaptar la estructura de UPADE agregando una categoría de patrones que traten los problemas específicos del acceso a datos genómicos.
- **Relaciones semánticas.**- Un principio fundamental del POD es permitir al desarrollador explotar las relaciones semánticas existentes entre los patrones. De esta manera, el desarrollador puede saber *cuándo* los patrones son aplicables, *como* pueden ser usados y *por qué* pueden ser o no combinados con otros patrones.

4.2 PATRONES EN EL DOMINIO GENÓMICO

El objetivo de esta sección es trazar una línea base sobre la relación entre los patrones de diseño de interfaz de usuario y el dominio genómico. Para este propósito, realizamos un estudio de literatura con una perspectiva de lo general a lo particular, que incluyó dos pasos. Primero, identificamos las librerías de patrones de diseño de interfaz existentes y aplicables en diferentes dominios. Segundo, analizamos si las librerías encontradas abordan los problemas de interacción en el dominio genómico o están de alguna manera vinculadas con el dominio genómico.

Utilizamos el término “librerías” para hacer referencia a un “conjunto de patrones” y evitar ambigüedades con términos como *catálogos de patrones* o *lenguajes de patrones*, que se mencionarán más adelante.

Las siguientes dos secciones tratan los dos pasos mencionados.

4.2.1 Patrones de diseño de interfaces existentes

La literatura muestra varios trabajos relacionados con patrones de diseño de interfaces. Utilizando los motores de búsqueda Scopus, IEEEExplorer y Springer, encontramos 185 documentos relacionados con frases que expresan el concepto de patrones de diseño de interfaz de usuario (p. ej., pattern, user interface pattern, user interface design pattern, HCI pattern, user interface pattern language, etc.) y publicados entre el 2015 y el 2019. En el contenido de estos documentos, especialmente en la sección “trabajos relacionados”, los autores hacen referencia a varias “librerías de patrones” existentes en la comunidad de diseño de interfaces de usuario. A partir del contenido de estos documentos, recopilamos un conjunto de 25 librerías que codificamos (para efectos de explicación) y listamos en la Tabla 2 y que analizamos desde varias perspectivas:

Tabla 2 Librerías de patrones encontradas en la literatura

Código	Librería	Referencia
1	Experiences - A Pattern Language for User Interface Design	(Coram et al., 1996)
2	Design reuse in hypermedia applications development	(Rossi et al., 1997)
3	Gama: A pattern language for computer supported dynamic collaboration	(Schümmer, 2003)
4	User interfaces Patterns for Hypermedia Applications	(Rossi et al., 2000)
5	User interaction design patterns for information retrieval	(Schmettow, 2006)
6	Just-ui: A user interface specification model	(Molina Moreno et al., 2002)
7	MyUI	(Peissner et al., 2011)
8	Interaction Design Patterns in Recommender Systems	(Cremonesi et al., 2015)
9	Usability patterns in games	(Folmer, 2006)
10	Designing Social Interfaces	(Crumlish et al., 2009)
11	A pattern approach for end user centered groupware development.	(Schümmer, 2005)
12	Patterns for computer-mediated interaction	(Schummer et al., 2013)
13	The design of sites: patterns, principles, and processes for crafting a customer-centered Web experience	(Van Duyne et al., 2003)
14	Designing Web Interfaces: Principles and Patterns for Rich Interactions	(Scott et al., 2009)
15	A pattern language for Web usability	(Graham et al., 2002)
16	UI-Patterns	(Foxboe, 2007)
17	Designing interfaces: Patterns for effective interaction design	(Tidwell, 2010)
18	Welie.com Patterns in Interaction Design	(Van Welie, 2008)
19	Pattern Tap	(ZURB, 2017)
20	Design patterns for user interface for mobile applications	(Nilsson, 2009)
21	User Interface Design Patterns	(Laakso, 2003)
22	The Data Visualisation Catalogue	(Ribeca, 2015)
23	UI Design Patterns for Navigation Complex Data Sets	(espero, 2015)
24	Software-Assisted Knowledge Generation in the Culture Heritage Domain: A Conceptual framework	(Martín-Rodilla, 2016)
25	User interface patterns in recommendation-empowered content intensive multimedia applications	(Cremonesi et al., 2017)

Catálogos o lenguajes. – Los autores de las librerías de patrones diferencian si la librería presentada es un “catálogo de patrones” o “lenguaje de patrones”. Como lo menciona (Kruschitz et al., 2010), una librería calificada como catálogo de patrones, también denominado *colección* de patrones, se caracteriza por i) soportar navegación rápida dentro de

un repositorio, ii) presentar un sistema de patrones no totalmente conectado (algunos patrones pueden estar aislados sin tener una relación con sus antecesores o predecesores), iii) no cubrir, usualmente, todos los problemas de diseño de un dominio específico. Por otro lado, una librería calificada como “lenguaje de patrones” se caracteriza por i) presentar un sistema de patrones totalmente conectado ii) cubrir todos los problemas de un dominio en concreto.

La Tabla 3 muestra las librerías de patrones encontradas clasificadas por catálogo o lenguaje de patrones.

Tabla 3 Tipos de librerías de patrones

Tipo de librería	Código de Librería
Catálogo de patrones	17, 19, 22, 24, 4, 7, 8, 9, 10, 14, 15, 20, 21, 23, 25
Lenguaje de patrones	1, 2, 3, 5, 6, 11, 12, 13, 16, 18

Dominio de aplicación. – Las librerías de patrones encontradas están dirigidas a dominios como las aplicaciones Web, móviles, escritorio, Hipermedia e ingeniería Web, Trabajo Cooperativo Asistido por Computadora (CSCW, por sus siglas en inglés), Sistemas de Recomendación. La Tabla 4 muestra los dominios mencionados y las librerías de patrones correspondientes a cada dominio.

Tabla 4 Dominios de aplicación de las librerías de patrones

Dominio	Código de Librería
Sitios Web	13, 16, 19, 26
CSCW	3, 11, 12
Diseño de Interacción	1, 17,18
Hipermedia e Ingeniería Web	2, 4
Herencia Cultural	24
Sitios Web Sociales	10
Big Data	23
Aplicaciones para usuarios con necesidades especiales	7
Usabilidad Web	15
Visualización de datos	22
Aplicaciones Móviles	20
Aplicaciones RIA	14
Especificación de interfaces	6
Juegos	9
Aplicaciones de escritorio	20, 21
Sistemas de recomendación	8
Aplicaciones multimedia	25
Sistemas de recuperación de datos (IRS)	5

Tipo de recurso. – El recurso utilizado con más frecuencia para publicar las librerías son los artículos de conferencias o revistas científicas (librerías 1, 2, 3, 4, 5, 6, 7, 8, 9, 25), sitios web (librerías 16, 17, 18, 19, 20, 21, 22) y libros (librerías 10, 11, 12, 13, 14, 15, 26). Otros recursos

menos frecuentes son las tesis (librería 24) y los seminarios en línea (librería 23). La literatura menciona que las librerías publicadas en línea (sitios web) son las más utilizadas por los diseñadores o desarrolladores debido a su facilidad de acceso. De las librerías publicadas en sitios web, encontramos que solo tres de ellas, listadas en la Tabla 5, documentan los patrones cumpliendo con los elementos básicos de un “patrón”: *problema, contexto, solución*. Las tres librerías reúnen un total de total de 404 patrones que principalmente están dirigidos al dominio de aplicaciones Web.

Tabla 5 Librerías de patrones en línea

Código librería	Librería	Autor	Número de patrones
17	Designing interfaces: Patterns for effective interaction design URL: https://designinginterfaces.com/patterns/	Tidwell	125
16	UI-Patterns URL: http://ui-patterns.com/patterns	Toxboe	148
18	Welie.com Patterns in Interaction Design URL: http://www.welie.com/patterns/index.php	Welie	131

Organización.- La organización de los patrones dentro de las librerías es también un tema relevante en la comunidad de patrones (Van Welie, 2003) (Hübscher et al., 2011) ya que impacta en el uso de las librerías: la organización de los patrones ayuda a los diseñadores o desarrolladores a buscar los patrones. Varios principios de organización han sido propuestos como por ejemplo agrupar los patrones por función, por similitud de problemas, por el defecto de usabilidad que abordan (p. ej., un grupo de patrones que incrementen la velocidad de ejecución de una tarea), por tarea y tipo de usuario al que están dirigidos (p. ej., grupos de patrones que tratan con seleccionar, encontrar o crear cosas considerando que los usuarios pueden ser principiantes, intermedios o expertos).

Tabla 6 Grupos de patrones definidos por Duyne. Traducido de (Van Duyne, 2003)

Letra	Nombre
A	Género del sitio
B	Crear un marco de navegación
C	Crear una poderosa Pagina de Inicio
D	Escribir y manejar contenido
E	Construir confianza y credibilidad
F	E-Commerce básico
G	E-Commerce avanzado
H	Ayudar a los clientes a completar las tareas
I	Diseñar una distribución efectiva de la página
J	Hacer que la búsqueda sea rápida y relevante
K	Hacer que la navegación sea fácil
L	Acelerar tu sitio

En la librería 13, Duyne (Van Duyne, 2003) estructura los patrones formando 12 grupos de patrones ordenados por letra y nombre como se muestra en la Tabla 6, donde cada grupo contiene patrones temáticamente relacionados. Así, si el desarrollador necesita establecer la distribución del contenido de la página, entonces se apoya de los patrones del “Grupo I. Diseñar una distribución efectiva de la página”. Además, los grupos están organizados en correspondencia al proceso de diseño haciendo que un grupo de patrones sea utilizado después de utilizar algún grupo anterior en la lista. Por ejemplo, en el proceso de diseño de un sitio, los patrones del grupo A y B tratarían el género del sitio y la estructura de navegación, luego, el grupo D trataría problemas específicos del contenido del sitio y, finalmente, los grupos J, K, y L tratarían con cómo mejorar el mecanismo de búsqueda, cómo facilitar la navegación y cómo hacer que el sitio sea más rápido, respectivamente.

Welie et. al. (Van Welie, 2003) adoptan un enfoque de “arriba hacia abajo” para organizar jerárquicamente los patrones de diseño de interacción, desde los problemas de diseño de más alto nivel hasta los de más bajo nivel. Esta organización es ilustrada en su colección de patrones publicada en www.welie.com/patterns (librería 18).

Descriptivos o generativos. – Vanderdonck et. al. (Vanderdonck et al., 2010) clasifica los patrones en descriptivos y generativos. Los *patrones descriptivos* describen el problema, el contexto y la solución potencial para resolver el problema y su descripción, y generalmente ofrecen una descripción bastante detallada, pero lo suficiente general para poder ser aplicados a un amplio espacio de problemas, interpretando la descripción del patrón a un contexto de uso particular. Por la forma de su descripción, el uso de estos patrones está dirigido a personas, principalmente líderes de proyectos, diseñadores, analistas y desarrolladores. Por otro lado y contrario a los patrones descriptivos, los *patrones generativos* consisten en una representación orientada a objetos que ofrece la expresividad suficiente para, a partir de su representación, facilitar su implementación, ya sea por “interpretación” de código (p. ej., patrones expresados en HTML que son interpretados en tiempo de ejecución por un motor de representación) o por “generación automática de código” (p. ej., patrones expresados en un lenguaje de programación o transformaciones modelo-código en el Desarrollo Dirigido por Modelos). Por su forma de representación, el uso de estos patrones está dirigido a autómatas (p. ej., algoritmos).

Aplicando la perspectiva de clasificación de Vanderdonck et. al. a nuestro conjunto de librerías encontramos que con excepción de la librería 6, todas las librerías caen dentro de la categoría “patrones descriptivos”. La mayoría de los patrones ofrecen un grado profundo de detalle en su descripción, pero no lo suficiente para permitir al patrón ser aplicado para solucionar problemas en otros contextos. La librería 6 cae dentro de los patrones generativos ya que los patrones en esta librería están representados en modelos a partir de los cuales, aplicando transformaciones de modelo a código, se puede obtener el código final del patrón de manera automática.

4.2.2 Patrones de diseño de interfaces y el dominio genómico

A partir del estudio presentado en la sección anterior, llegamos a dos resoluciones respecto al estado de los patrones de diseño de interfaces en el dominio genómico.

- *No existen patrones específicos para el dominio genómico.* – Aunque son varios los dominios cubiertos por las librerías estudiadas (ver Tabla 4), no existe una librería con patrones específicos para diseñar interfaces de usuario en el dominio genómico.
- *Algunos patrones pueden ser adaptados al dominio genómico.* – Aunque las librerías de patrones estudiadas no son específicas del dominio genómico, algunos patrones contenidos en ellas han sido aplicados para resolver problemas relacionados al dominio genómico. Esta situación nos ha llevado a pensar que la descripción de estos patrones puede ser adaptada para cubrir un problema de interacción específico dentro del dominio genómico, es decir, el problema, contexto y solución del patrón deben ser descritos en términos específicos del dominio genómico. Por ejemplo, uno de los ejemplos del patrón “Radial Table”, documentado en la página 327 del libro de Tidwell (Tidwell, 2010) (librería 17 de la Tabla 2), muestra el patrón aplicado para visualizar la secuencia genética de una bacteria y la relación entre sus elementos a través de una representación circular. Tidwell es clara al describir que el patrón Radial Table es útil cuando se tiene una larga lista o tabla de ítems y se necesita mostrar arbitrariamente las relaciones entre ellos. Sin embargo, desde el punto de vista genómico, esta descripción no menciona el problema específico de interacción con datos genómicos que se desea solucionar. Describiendo el problema, contexto y solución del patrón Radial Table en términos genómicos, este patrón podría ser de mucha ayuda para desarrolladores de interfaces en el dominio genómico.

Considerando estas resoluciones, en esta Tesis nos planteamos contribuir con un conjunto de patrones que aborden los problemas de diseño de interfaz específicos para interactuar con datos genómicos. A estos patrones los hemos denominado *patrones genómicos* y se detallan más adelante, en la sección 6.2.5. Los patrones genómicos junto con patrones que tratan problemas de diseño generales de la interfaz de usuario serán parte de un catálogo de patrones que servirá como instrumento para apoyar al método GenomIUm. Así como los patrones de Duyne (Van Duyne, 2003) o Welie (Van Welie, 2003), el catálogo de patrones propuesto estructura los patrones siguiendo una organización particular que hemos definido para abordar el diseño de la interfaz en el dominio genómico. Los patrones genómicos, el catálogo de patrones y la organización interna de los patrones son detalladas más adelante, en el Capítulo 6.

4.3 CONCLUSIONES

En este capítulo se ha presentado el estado del arte que soporta esta Tesis y que ha sido tratado desde dos líneas principales:

- Línea 1: Enfoques de diseño de interfaces basados en patrones
- Línea 2: Patrones en el dominio genómico.

El conocimiento adquirido en el estudio de estas líneas nos proporciona una sólida base para decidir sobre las mejores prácticas y enfoques que podemos aplicar para diseñar GenomIUm, el método para diseñar interfaces de usuario en el dominio genómico y que se

propone en esta Tesis. Este conocimiento puede resumirse en función de las líneas mencionadas:

- En relación con la línea 2, el estudio de la literatura reportó tres enfoques relevantes en el diseño de interfaces basado en patrones, estos son: el diseño centrado en el usuario (DCU), enfoque soportado por patrones (PSA) y el diseño orientado por patrones (POD). El DCU y el PSA abordan todas las fases del proceso de diseño de la interfaz de usuario (es decir, análisis, diseño, implementación y evaluación) con la diferencia de que el DCU motiva el uso de patrones en la fase de diseño del proceso, mientras que el PSA motiva el uso de patrones en todas las fases del proceso de diseño y hace uso de las relaciones entre los patrones para sugerir al desarrollador patrones adecuados en cada fase del proceso. El tercer enfoque es el POD que se enmarca en la fase de diseño del proceso y se centra en el desarrollo del diseño conceptual de la interfaz de usuario; el artefacto clave para la posterior implementación de la interfaz. Para esto, el POD se basa en dos principios fundamentales: la definición de un proceso de diseño conceptual de la interfaz y la explotación de las relaciones semánticas entre patrones a lo largo de cada fase del proceso de diseño conceptual. Para nuestro propósito, consideramos que el alcance de los enfoques DCU y PSA respecto a la cobertura del proceso de diseño de la interfaz de usuario es demasiado amplio. Nosotros nos enfocamos en la fase de diseño del proceso, donde se produce el concepto de la interfaz de usuario. Es por eso por lo que, de los tres enfoques estudiados, el POD es el que más se alinea con nuestro propósito de diseñar interfaces en el dominio genómico.
- En relación con la línea 3, hemos determinado que no existen patrones específicos para el diseño de interfaces en el dominio genómico, sin embargo, consideramos que algunos patrones en las librerías de patrones existentes pueden ser adaptados para abordar algunos problemas específicos en el dominio genómico.

El conocimiento obtenido en este capítulo es de utilidad para decidir sobre los detalles de la solución que intentamos proveer y que servirá para resolver el problema del diseño de interfaces en el dominio genómico.

PARTE III.
DISEÑO DE LA SOLUCIÓN

Capítulo 5

Requisitos de la Solución

5.1 INTRODUCCIÓN

En la Sección 2.3, correspondiente a los ciclos de la metodología DS, se mencionó que la “tarea de diseño” (T2) consiste en tres acciones principales:

- Especificar los requisitos que debe cumplir la solución para resolver el problema de diseñar interfaces de usuario de acceso a datos genómicos
- Tomar decisiones frente a las posibles alternativas de solución en función de los requisitos.
- Diseñar la solución.

Este capítulo se enfoca en las dos primeras acciones y pone los cimientos para la tercera acción. Para esto:

- En la Sección 5.2 se definen los requisitos que son las *propiedades deseadas de la solución* (Wieringa, 2014). Por lo tanto, estos requisitos son útiles para buscar potenciales soluciones y decidir sobre posibles alternativas de solución (si utilizar o adaptar una solución existente o diseñar una nueva solución para el problema bajo estudio).
- En la Sección 5.3, basados en estos requisitos y el conocimiento obtenido del estado del arte, decidimos la mejor estrategia de solución. La decisión resultante será la base para la tercera acción, es decir, diseñar la solución propuesta por esta Tesis, la cual se tratará en el Capítulo 6.

5.2 REQUISITOS DE LA SOLUCIÓN PREVISTA

Dentro de los intereses de nuestro Centro de Investigación (PROS) está la gestión de datos genómicos a través de un método sistemático de cuatro niveles, denominado SILE, que se inicia con el nivel de búsqueda de datos genómicos y termina con el nivel de

explotación de datos. Un requisito concreto del PROS es apoyar al nivel de “explotación de datos” de SILE con una solución que guíe el diseño de interfaces de usuario para interactuar con los datos genómicos.

Esta sección describe detalladamente dicho requisito del Centro PROS especificando los problemas de interacción con los datos genómicos que debe cubrir la solución requerida. De esta manera, en las subsecciones 5.2.1 y 5.2.2 se provee información acerca del método SILE y los problemas de interacción con los datos genómicos, respectivamente. Con esta información, en la subsección 5.2.3 se delinean los requisitos que debe cumplir la solución para abordar el problema del diseño de interfaces en el contexto genómico.

5.2.1 El método SILE y el panorama de diseño de interfaces

SILE es un método instituido y propuesto por el PROS para la gestión sistemática de datos genómicos (Reyes R., 2014) y ha sido aplicado en el descubrimiento de información genética relacionada a distintas enfermedades del ser humano tales como catarata congénita (Navarrete-Hidalgo et al., 2018), sensibilidad al alcohol (Román et al., 2016) y neuroblastoma (Burriel et al., 2017). Actualmente, SILE es aplicado en la academia para identificar variaciones genéticas de las enfermedades de Crohn, migraña, epilepsia y cáncer de mama (Palacio et al., 2018). Su versión más elaborada y completa ha sido recientemente presentada en la Tesis Doctoral de Ana León (León Palacio, 2019). SILE consta de cuatro niveles:

- **Buscar (Search).** - Determinar los límites del análisis genético e identificar las fuentes de datos de donde extraer la información.
- **Identificar (Identification).** – A partir de la información buscada, determinar los datos relevantes y confiables para cargar en una base de datos estructurada según un modelo holístico del genoma (Pastor Lopez et al., 2016).
- **Cargar (Load).** – Extraer, transformar y cargar datos identificados en el nivel anterior.
- **Explotación (Exploitation).** – Visualizar, descubrir, explorar los datos genómicos almacenados con el fin de analizar e interpretar los datos.

En especial, la descripción del nivel de “Explotación” hace referencia a la priorización de variaciones genéticas con el objetivo de analizar, visualizar, descubrir los datos y derivar conclusiones.

El método SILE constituye una columna vertebral a la cual se han adherido enfoques, métodos y estrategias para apoyar diferentes aspectos de la gestión de datos genómicos. Un ejemplo de esto es el *método sistemático de calidad de datos* (Palacio, 2018) que ha sido adaptado a los tres primeros niveles (es decir, buscar, identificar y cargar) para apoyar la carga selectiva de datos y reducir la gran cantidad de datos a un conjunto de datos relevante, útil y confiable a través de la aplicación de ciertos parámetros de calidad.

Precisamente, esta Tesis tiene el objetivo de contribuir con una solución para diseñar interfaces de usuario para la explotación de datos.

5.2.2 Problemas de interacción con datos genómicos

Para identificar los problemas de interacción con los datos genómicos, no solo nos centramos en los problemas de interacción experimentados por miembros del PROS desde una perspectiva académica, sino también en los problemas de interacción experimentados en la industria. Los problemas identificados son el resultado de i) reuniones de discusión con el grupo de investigación del Genoma Humano del PROS de la UPV, y ii) entrevistas a investigadores y colaboradores del dominio genético como son: laboratorio genético Tellmegen², Centro de Genómica del Instituto Valenciano de Investigaciones Agraria (IVIA³) y estudiantes en ingeniería biomédica de la Universidad Politécnica de Valencia (UPV).

Algunos de los problemas encontrados fueron repetitivos en las entrevistas y reuniones realizadas. Después de consolidar la información, encontramos un total de 16 problemas de interacción:

- *Problema 1.-* Buscar elementos genéticos (p. j., genes o variaciones) por su nombre o identificador
- *Problema 2.-* Segregar variaciones genéticas según su genotipo
- *Problema 3.-* Filtrar las variaciones respecto al fenotipo relacionado
- *Problema 4.-* Representar gráficamente la distribución de variaciones genéticas, genes, proteínas, etc.
- *Problema 5.-* Filtrar el conjunto de variaciones genéticas estableciendo condiciones inherentes a la variación
- *Problema 6.-* Comparar muestras de variaciones genéticas para saber diferencias o similitudes entre ellas.
- *Problema 7.-* Explorar y navegar la información genética sin experimentar interrupciones en la recuperación de información.
- *Problema 8.-* Reducir la vista tabular de las variaciones genéticas a columnas específicas de interés.
- *Problema 9.-* Comparar dos o más secuencias genómicas o anotaciones respecto a una localización exacta de la secuencia genética.
- *Problema 10.-* Filtrar el conjunto de variaciones estableciendo condiciones relacionadas a los efectos que ésta produce.
- *Problema 11.-* Reproducir los pasos de un análisis genético previamente realizado.
- *Problema 12.-* Limitar el análisis de variaciones genéticas a un sector específico de la secuencia genética.
- *Problema 13.-* Representar gráficamente las relaciones entre el material genético de diferentes cromosomas.
- *Problema 14.-* Representar gráficamente el complejo cromosómico de una especie en concreto.
- *Problema 15.-* Analizar un elemento genético de interés en su contexto genómico, es decir, relacionar al elemento genético con anotaciones de diferentes fuentes.

² <https://www.tellmegen.com/>

³ <http://www.ivia.gva.es>

- *Problema 16.-* Cargar muestras contenidas en ficheros con diferentes formatos para posteriormente analizarlas.

La lista de problemas se convierte en el espacio de problemas de diseño que abordaremos en el diseño de interfaces de usuario de acceso a datos genómicos.

5.2.3 Requisitos

Después de estudiar la necesidad de soportar el nivel de “explotación” de SILE y los aspectos del panorama de diseño de interfaces en el ámbito genómico descritos en la Sección 5.2.1 así como los problemas de interacción en el consumo de datos genómicos descritos en la Sección 5.2.2, consideramos que una solución apropiada a esta necesidad y estos problemas es proponer un **método de diseño de interfaces de usuario para la explotación de datos genómicos**. Por lo tanto, sintetizando lo mencionado en las dos subsecciones anteriores, los requisitos que debe cumplir el método son:

Requisito 1 – Producir interfaces usables. – Las interfaces diseñadas con el método propuesto deben evidenciar un impacto positivo en su usabilidad en términos de eficiencia, efectividad y satisfacción.

Requisito 2 – Abordar los problemas de interacción de acceso a datos genómicos. – El método debe proveer soluciones a los problemas de interacción con los datos genómicos identificados en la Sección 5.2.2.

Requisito 3 – Proveer instrumentos que faciliten la comunicación entre los implicados en el diseño de la interfaz y la comprensión del problema de interacción. – El método debe contar con instrumentos que sirvan como canal de comunicación entre los desarrolladores, diseñadores y usuarios finales, integrando los aspectos del dominio (conceptos biológicos y sus relaciones) y técnicos (aspectos de implementación). Los instrumentos deben permitir entender el problema de interacción y comprender la solución para poder implementarla.

Requisito 4 – Guiar el diseño conceptual de la interfaz de usuario. - El método debe conducir a desarrolladores/diseñadores, novatos y expertos en el dominio genómico, a través de los diferentes pasos de la producción del diseño conceptual de la interfaz de usuario.

Requisito 5 – Reutilizar conocimiento. – El método propuesto debe facilitar a desarrolladores novatos el diseño de interfaces de usuario para el acceso a datos genómicos aprovechando y reutilizando los conocimientos, experiencias, estrategias, directrices y buenas prácticas aplicadas por diseños expertos.

Requisito 6 – Soportar el diseño conceptual. – El método debe apoyar la creación temprana de prototipos con el fin de que el usuario final pueda evaluar el concepto de diseño de la interfaz de usuario.

A partir de estos requisitos, decidimos en la siguiente sección si la solución más adecuada es utilizar o adaptar un método de diseño de interfaces existente, o crear un nuevo método.

5.3 DETERMINACIÓN RESPECTO A ALTERNATIVAS DE SOLUCIÓN

Para tomar una determinación frente a las alternativas de solución al problema que planteamos en la Tesis, analizamos en qué medida los enfoques existentes y librerías de

patrones estudiadas en el Capítulo 4 (Estado del Arte) cumplen con los requisitos especificados en la sección anterior. Esto permitirá decidir si utilizar o adaptar una solución existente, o crear una nueva solución.

La Tabla 7 muestra los enfoques evaluados en función de los requisitos. Los requisitos 1, 5 y 6 son cumplidos por todos los enfoques ya que estos enfoques están “basados en patrones” y que el uso de los patrones promueve la producción de interfaces usables, la reutilización de conocimiento y facilita el diseño de prototipos de interfaz de usuario. Hasta aquí, todos los enfoques se convierten en alternativas adecuadas para nuestro propósito. Sin embargo, el cumplimiento de los requisitos 2, 3 y 4 evidencia diferencias entre los enfoques. Para efectos de claridad en la explicación, primero analizamos el requisito 4 y luego los requisitos 2 y 3:

Respecto al requisito 4, el único enfoque que cumple es el POD. Esto es porque a diferencia de los otros enfoques que cubren de manera general todas las etapas del desarrollo de la interfaz de usuario, el POD guía exclusivamente, y de manera sistemática, el proceso para diseñar el concepto de la interfaz de usuario, que es lo que perseguimos en esta Tesis. Por lo tanto, concluimos, hasta aquí, que el POD se convierte en una alternativa de solución potencial al problema que queremos resolver.

Tabla 7 Evaluación de enfoques respecto a los requisitos de la solución

Requisito	DCU	PSA	GDD	POD
1.- Producir interfaces usables	SI	SI	SI	SI
2.- Abordar los problemas de interacción de acceso a datos genómicos	NO	NO	NO	NO
3.- Proveer instrumentos que faciliten la comunicación entre los implicados en el diseño de la interfaz y la comprensión del problema de interacción	NO	NO	NO	NO
4.- Guiar el diseño conceptual de la interfaz de usuario	NO	NO	NO	SI
5.- Reutilizar conocimiento	SI	SI	SI	SI
6.- Soportar el diseño conceptual	SI	SI	SI	SI

Respecto a los requisitos 2 y 3, ninguna de los enfoques cumple con estos requisitos. Esto tiene sentido ya que los enfoques de diseño de interfaces que estamos analizando no condicionan su uso a un dominio de aplicación específico mientras que los requisitos 2 y 3 condicionan a la solución a proveer instrumentos que traten con las complejidades del dominio genómico. Es importante considerar que los enfoques analizados promueven el uso de patrones como instrumentos para abordar los problemas de interacción y facilitar la comunicación entre los miembros. En consecuencia, estos enfoques pueden adaptarse a un dominio específico incluyendo como instrumentos a patrones que traten los problemas de dicho dominio. Esto nos lleva a la búsqueda de librerías de patrones que traten los problemas de interacción en el dominio genómico y que puedan ser integradas y utilizadas en los enfoques. En este sentido, el análisis realizado en la Tabla 4 del Estado de Arte es de

utilidad y evidencia que las librerías estudiadas están dirigidas a dominios de aplicación como Sitios Web, Sistemas Colaborativos (CSCW), Diseño de Interacción, Hipermedia e Ingeniería Web, Herencia Cultural, Sitios Web Sociales, Juegos, Aplicaciones de Escritorio, Sistemas de Recomendación, entre otros. No existen librerías de patrones específicas del dominio genómico. Por lo tanto, concluimos que para cumplir los requisitos 2 y 3 es necesario considerar la definición de un conjunto de patrones dirigido al dominio genómico.

Apoyados en las conclusiones obtenidas en el análisis realizado, decidimos que la solución más apropiada para el problema de diseñar interfaces de usuario en el dominio genómico es:

- Crear, considerando el fundamento del enfoque POD, un método de diseño de interfaces que, soportado por patrones específicos del dominio genómico que deben ser previamente definidos, guíe el diseño de las interfaces de usuario de acceso a datos genómicos, desde su diseño conceptual inicial hasta su diseño conceptual refinado y listo para ser implementado. A este método lo hemos denominado **GenomIUm**.

Las razones que han motivado el uso de patrones en GenomIUm y cómo se alinean estas razones con los requisitos planteados en la Sección 5.2 se muestran en la Tabla 8.

Tabla 8 Razones de usar patrones mapeadas a los requisitos de GenomIUm

Razón	Descripción	Requisito
Experiencias previas satisfactorias	Las experiencias anteriores en el diseño de interfaces (Iñiguez-Jarrin et al., 2017) nos lleva a confiar en que el uso de los patrones puede ser un mecanismo viable y útil para hacer frente a la complejidad de diseñar interfaces de usuario en el contexto genómico.	Todos
Mejora de la usabilidad	Diseñar una interfaz usable implica utilizar soluciones de diseño eficaces y eficientes que sean familiares para los usuarios. Esto sintetiza la ley de Jakob (Nielsen, 2000) sobre la experiencia de usuario de internet “ <i>Los usuarios pasan la mayoría de su tiempo en otros sitios que en tu propio sitio. Esto significa que ellos prefieren que tu sitio funcione igual que los sitios que ya conocen</i> ”. Por lo tanto, si el diseñador quiere mejorar la usabilidad de su interfaz debe utilizar patrones, estándares, es decir, soluciones utilizadas en otras aplicaciones que sean conocidas para el usuario que en consecuencia reduzcan la curva de aprendizaje.	1
La falta de librerías de patrones para el dominio genómico	Aunque existen varias colecciones de patrones de interfaz de usuario en diferentes dominios, estos no son suficientes para abordar los problemas específicos de acceso a datos genómicos. Es necesario contar con un catálogo de patrones particular para el contexto genómico.	2
Facilidad de comunicación	Los patrones facilitan la comunicación entre los implicados en el proceso de diseño ya que su descripción está pensada para ser leída y entendida fácilmente por desarrolladores, diseñadores y usuarios finales.	3
Toma de decisiones	Los patrones pueden aplicarse en diferentes etapas del proceso que guía el diseño conceptual de la interfaz (p. ej., estructura del contenido, navegación) ayudando al desarrollador, novato o experto, a tomar decisiones de diseño en cada paso del proceso.	4

Reutilización de conocimiento	El uso de patrones en el diseño permite trasladar el conocimiento de desarrolladores e ingenieros en usabilidad expertos a diseñadores y desarrolladores más novatos de tal manera que el conocimiento de los expertos pueda ser reutilizado.	4, 5
Uso como bloques de construcción reusables	Los patrones pueden ser vistos como bloques de construcción reutilizables que combinados adecuadamente pueden soportar la creación del diseño conceptual de la interfaz.	6

5.4 INSUMOS NECESARIOS PARA EL MÉTODO

Desde una perspectiva general, un método puede ser definido como un enfoque para desarrollar un producto de software y ha de contener todo lo necesario para lograrlo incluyendo un “proceso” (la manera de cómo hacer algo) y los “recursos o productos” empleados, entre otras cosas (adaptado de (Henderson-Sellers et al., 2014)). Esto quiere decir que para crear el método GenomIUM, necesitamos definir un proceso sistemático y recursos (el catálogo de patrones, en nuestro caso) que nos permita alcanzar el objetivo: diseñar el concepto de la interfaz de usuario. Clasificando estas necesidades en problemas de diseño (PD) y preguntas de conocimiento (PC) como lo expone la metodología DS, se plantean los siguientes:

PD 1: “Construir un catálogo de patrones de interfaz de usuario específico para el dominio genómico”. Para solucionar este PD, las preguntas que necesitamos responder son:

- PC 1. ¿Cómo identificar los patrones?
- PC 2. ¿Cómo documentar los patrones?

Una vez construido el catálogo de patrones, el siguiente problema de diseño es:

PD 2: “Construir un proceso sistemático para diseñar interfaces utilizando el catálogo de patrones”. Para solucionar este PD, las preguntas que necesitamos responder son:

- PC 3. ¿Qué principios de combinación de patrones deben ser considerados para formar el diseño conceptual de la interfaz de usuario?

La solución a los PD y las respuestas a las PC supone el diseño de varios recursos e instrumentos:

Recurso 1: Un catálogo de patrones de interfaz de usuario para acceso a datos genómicos. - Un conjunto de soluciones de diseño de interfaz de usuario reutilizables que cubra los problemas de interacción específicos del acceso a datos genómicos. Este recurso se convierte en la solución al PD 1. Para construir este recurso, necesitamos los siguientes instrumentos:

- *Instrumento 1:* Un proceso para identificar las soluciones de diseño. – Un proceso sistemático para identificar y definir dichas soluciones a partir de los problemas identificados en el instrumento 1. Este recurso responde a la PC 1.
- *Instrumento 2:* Un formato para documentar patrones. – Una plantilla que sea entendible tanto para diseñadores, desarrolladores y usuarios. Este recurso responde a la PC 2.

Recurso 2: Un proceso sistemático de diseño de interfaz de usuario. – Un conjunto de fases sucesivas, basadas en el enfoque PDD, que facilite el diseño conceptual de la interfaz de usuario utilizando el catálogo de patrones. Este recurso responde a la PD 2.

- *Instrumento 3: Directrices para combinar los patrones.* – Una instrucción a seguir para la combinación de patrones en el diseño conceptual de la interfaz de usuario. Este recurso responde a la PC 3.

El desarrollo de estos recursos e instrumentos, son tratados a profundidad en el Capítulo 6.

5.5 CONCLUSIONES

En este capítulo se ha definido los requisitos que debe cumplir una solución que resuelva el problema de diseñar interfaces en el contexto genómico. Estos requisitos se sintetizan en que la solución debe:

- Apoyar la “explotación” de datos genómicos (visualización, exploración, operación y navegación).
- Producir interfaces que permitan a los usuarios lograr sus objetivos de manera eficiente, eficaz y produzcan un buen nivel de satisfacción en su uso.
- Abordar los problemas específicos de interacción con los datos genómicos.
- Proveer instrumentos que faciliten la comprensión del problema y la comunicación.
- Reutilizar el conocimiento de diseñadores expertos.
- Facilitar el diseño conceptual de la interfaz de usuario.

Evaluando las posibles alternativas de solución descritas en estado de arte respecto a la cobertura de los requisitos expuestos, hemos resuelto que la solución para diseñar interfaces de usuario en el dominio genómico es crear un método de diseño de interfaces adaptando el enfoque de Diseño Orientado por Patrones (POD). A este método lo denominamos *GenomIUm*.

Finalmente, con la perspectiva de implementar el método *GenomIUm*, se han delineado los recursos necesarios para su efecto, tales como:

- El proceso sistemático que guía el diseño conceptual de la interfaz de usuario
- Un catálogo de patrones de interfaz de usuario específico para el dominio genómico que soporte al proceso sistemático.

Capítulo 6

Método GenomIUm

En este capítulo describimos el método GenomIUm que apoya al diseñador de interfaces en la definición del diseño conceptual de la interfaz de acceso a datos genómicos. En la Sección 6.1, presentamos una visión general del método mencionando sus componentes: i) el proceso del diseño conceptual de la interfaz de usuario y ii) el catálogo de patrones que soporta al proceso. En lo que sigue del capítulo, profundizamos en cada componente: en la Sección 6.2 desarrollamos el catálogo de patrones y en la Sección 6.3 detallamos cada uno de los pasos del proceso.

6.1 VISIÓN GENERAL

GenomIUm consiste en un proceso de diseño y un catálogo de patrones que soporta al proceso, como lo muestra la Imagen 6.1. Como se puede observar, el método GenomIUm está enmarcado en la actividad de diseño del proceso general de desarrollo de la interfaz de usuario y se enfoca en producir el artefacto clave de esta actividad: el *diseño conceptual de la interfaz de usuario*. GenomIUm está basado en el enfoque POD. Es por eso por lo que define un proceso de diseño y un catálogo de patrones para soportar al proceso.

En GenomIUm, el desarrollador diseña el concepto de la interfaz de usuario (un prototipo) combinando patrones, basado en las relaciones semánticas entre los patrones. El proceso de diseño cubre desde la arquitectura de información del sistema interactivo hasta llegar al diseño detallado de cada una de las interfaces de usuario que componen el sistema. El proceso consiste en cuatro pasos: *diseño arquitectónico*, *diseño estructural*, *diseño de contenido* y *refinamiento*.

El catálogo de patrones que soporta al proceso se denomina GenomIU. El catálogo estructura los patrones en categorías aplicables a cada paso del proceso de diseño: la categoría *patrones de información* apoya al paso de diseño arquitectónico, la categoría *patrones de página* apoya al diseño estructural y las categorías *patrones de contenido* y *patrones de navegación* apoyan al diseño de contenido. De esta manera, el diseñador aborda cada paso del proceso de diseño con la categoría de patrones correspondiente. Además, los patrones del catálogo están relacionados semánticamente formando una red de patrones conectados. Así, el diseñador no solo se apoya en la categoría de patrones correspondiente a cada paso del proceso, sino que puede explotar las relaciones semánticas entre los patrones y decidir la mejor combinación de patrones para el diseño del concepto de la interfaz de usuario.

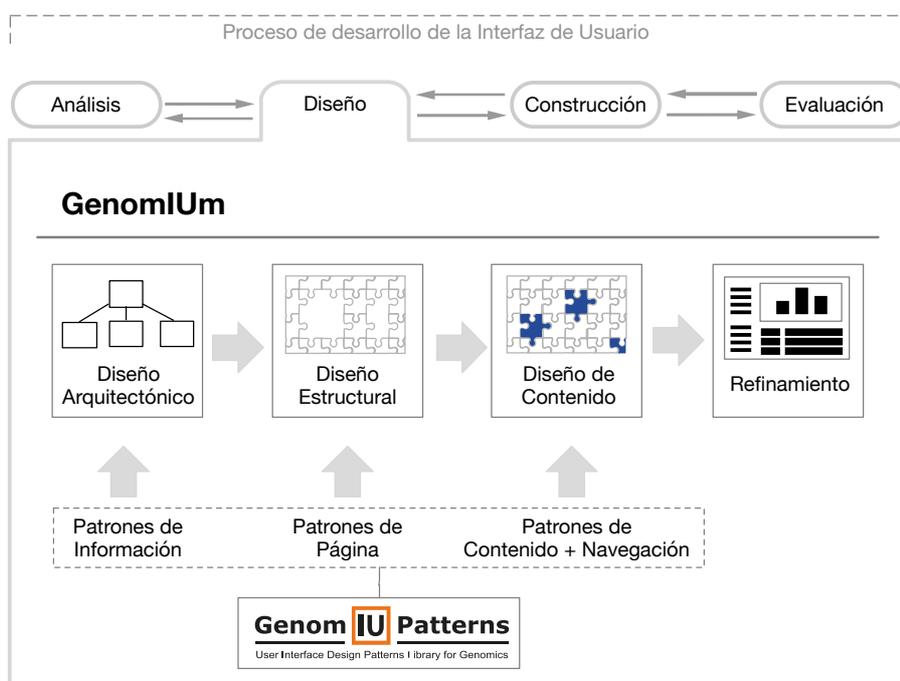


Imagen 6.1 Proceso de diseño de interfaz de usuario del método GenomIU

Las siguientes subsecciones describen el catálogo de patrones GenomIU y los pasos del proceso diseño donde el catálogo es aplicado.

6.2 CATÁLOGO DE PATRONES: GENOMIU

El catálogo GenomIU, que apoya al proceso de diseño de GenomIU, está estructurado en categorías de patrones que apoyan a cada paso del proceso de diseño conceptual de la interfaz de acceso a datos genómicos. Dado que la Web es la plataforma de mayor uso para las herramientas de acceso a datos genómicos, el catálogo GenomIU reúne patrones de diseño de interfaz Web.

Para apoyar todo el proceso de diseño, las categorías de patrones se componen de patrones *genéricos* y *específicos*. Mientras los patrones genéricos tratan problemas generales de diseño Web (p. ej., arquitectura de información del sitio, organización interna de la interfaz de usuario, menús de navegación), los patrones específicos tratan problemas concretos del diseño de interfaces de acceso a los datos genómicos (p. ej., cargar ficheros de información genómica, visualizar y explorar la secuencia genética).

Los patrones genéricos que conforman el catálogo GenomIU son una colección de patrones existentes en las librerías publicadas por Duyne (Van Duyne, 2003), Toxboe (Toxboe, 2007), Welie (Van Welie, 2008) y Tidwell (Tidwell, 2010), correspondientes a las librerías 13, 16, 17, 18 de la Tabla 2 de librerías estudiadas en el estado de arte. A pesar de

que los patrones genéricos son elementos importantes del catálogo, la descripción de estos patrones está fuera del interés de esta Tesis. No obstante, el lector puede revisar su descripción en la página o recurso original del autor.

Los patrones específicos que conforman el catálogo GenomIU, al contrario de los patrones genéricos, no están documentados en la literatura. Nuestro estudio sobre patrones en el estado de arte (Sección 4.2) reportó que no existen patrones documentados para el acceso a datos en el dominio genómico. Los patrones específicos han sido documentados por nosotros después de analizar las interfaces de usuario de varias aplicaciones en el dominio. El conjunto de patrones específicos se convierte en una de las contribuciones directas de esta Tesis por lo que dedicamos una sección completa para describirlos (Sección 6.2.5).

La Imagen 6.2 muestra un resumen del catálogo GenomIU estructurado en las cuatro categorías de patrones que apoyan el proceso de diseño del método GenomIU. Estas categorías son: *patrones de información*, *patrones de página*, *patrones de contenido* y *patrones de navegación*. La categoría patrones de contenido se compone a su vez de dos categorías de menor nivel: *patrones genómicos* y *patrones contenedores de Información*. En esta estructura, los patrones específicos están resaltados con línea discontinua y corresponden a la categoría de patrones genómicos, las demás categorías son patrones genéricos. En la práctica, GenomIU contiene un gran número de patrones genéricos, sin embargo, en la imagen resumen del catálogo han sido incluidos algunos de ellos únicamente para ilustrar el contenido de cada categoría.

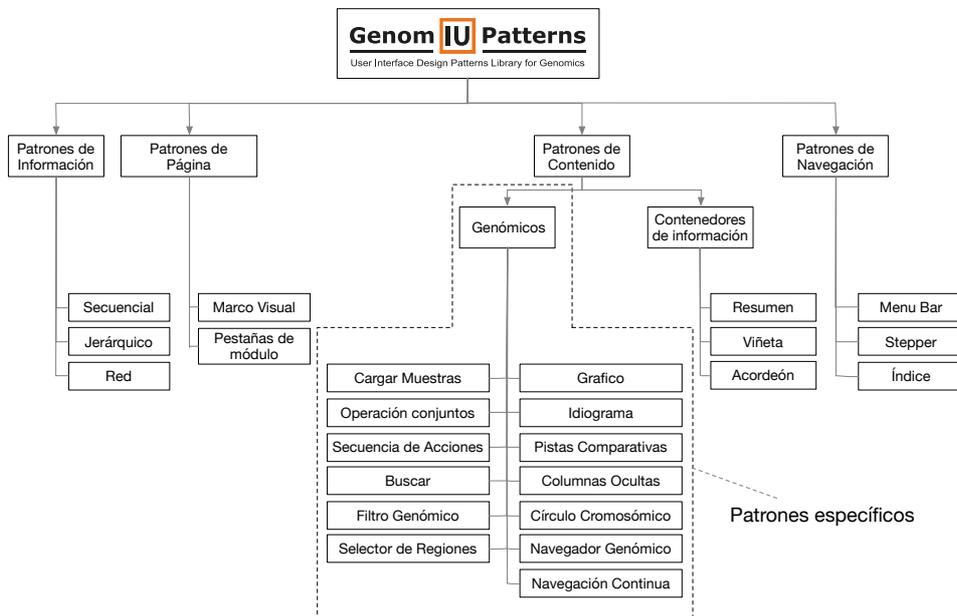


Imagen 6.2 Resumen del catálogo de patrones GenomIU

A continuación, se describe las categorías de patrones del catálogo GenomIU:

Patrones de Información. - Describen soluciones para la organización de la información de todo el sistema, definiendo unidades de presentación (que representan a las interfaces de usuario del sistema) y sus enlaces. Los patrones de esta categoría son explicados en función de nodos y enlaces, donde los nodos representan las unidades de presentación y los enlaces la navegación entre ellas. Por ejemplo, el patrón “Secuencial” establece una secuencia de nodos lineal.

Patrones de Página. – Describen soluciones para la distribución interna de la unidad de presentación, definiendo sectores para el contenido de la página. Por ejemplo, el patrón “Marco Visual” describe una solución para estandarizar la estructura visual de todas las interfaces del sitio interactivo estableciendo sectores de contenido comunes para todas las páginas como por ejemplo la cabecera de la página, pie de página, contenido, etc.

Patrones de Contenido. - Describen soluciones para lograr las tareas genómicas. Estos patrones se subdividen en:

- *Patrones Genómicos.* – Describen soluciones a los problemas de interacción (operación, visualización, navegación) con los datos genómicos que fueron descritos en la Sección 5.2.2. El patrón “Idiograma”, por ejemplo, describe una solución para representar gráficamente la distribución ordenada de los cromosomas de una especie.
- *Patrones Contenedores de información.* – Describen mecanismos de interacción relacionados a la presentación general del contenido y que son complementarios a los patrones genómicos. El patrón “Resumen” propone una solución para presentar de manera resumida un contenido extremadamente largo e importante para el usuario y que requiere suficiente espacio para ser presentado.

Patrones de Soporte a la Navegación. – Describen soluciones para organizar la interacción de la navegación por las páginas. El patrón “Índice”, por ejemplo, presenta una lista de índices navegables a las diferentes páginas que componen el sistema interactivo y que generalmente es ubicado en páginas principales para ayudar al usuario a explorar el sistema interactivo.

6.2.1 Modelo conceptual del catálogo

El catálogo GenomIU está construido bajo el modelo conceptual que se muestra en la Imagen 6.3, donde la clase “Catálogo GenomIU” representa el catálogo y la clase “Solución Diseño” las soluciones de diseño que lo conforman. Una solución de diseño es un diseño de interfaz de usuario, o una porción de esta, que representa una solución a un problema de diseño en la interfaz de usuario.

Todo patrón es una solución de diseño, sin embargo, no toda solución de diseño puede ser considerada como patrón. Para distinguir si una solución de diseño es un patrón, hemos adoptado la regla de oro ampliamente utilizada por la comunidad de patrones y denominada *Regla de tres* la cual establece que un patrón válido es una solución de diseño que evidencia al menos tres ejemplos de implementación. Si una solución de diseño no cumple con la regla de tres, es decir, el número de ejemplos de implementación de la solución es menor a tres, entonces consideramos a esta solución como *proto-patrón*. Un proto-patrón es una solución

candidata que aún no se sabe que pasa las pruebas de patrón de la regla de tres (Appleton, 2000) (Seffah, 2010) (Borchers, 2000). En nuestro modelo, patrón y proto-patrón están representados por las clases especializadas “Patrón” y “Protopatrón”. La condición de la regla de tres, exclusiva para considerar a una solución como patrón, está indicada con la condición “{3 ejemplos}” precisamente sobre la relación de especialización de la clase “Patrón”.

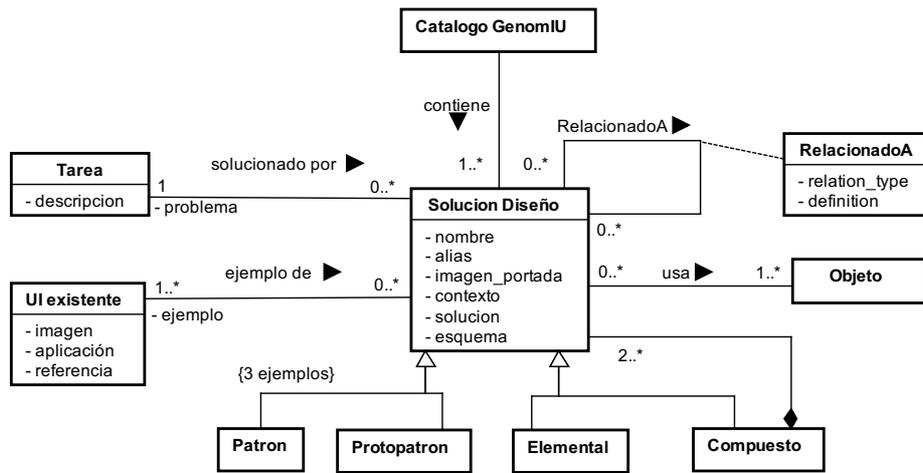


Imagen 6.3 Modelo del catálogo de patrones

La clase “Solución Diseño”, considerada ya sea como patrón o como proto-patrón, está definida por varios atributos los cuales están descritos como partes de un formato explicado más adelante en la Sección 6.2.5. La clase “Tarea” representa el *problema* de interacción que resuelve la solución de diseño. La clase “UI existente” representa a los *ejemplos* de interfaces de usuario donde se evidencia la implementación de la solución de diseño. Las interfaces de usuario “ejemplo” son el resultado de un estudio de interfaces de usuario de herramientas existentes en el dominio genómico y en el dominio de manejo de datos masivos (Big Data). Estas herramientas están detalladas en la Sección 6.2.4.

La clase “Objeto” hace referencia a los objetos del MCGH que intervienen en la solución del patrón. Por ejemplo, una solución de diseño que trata el problema de visualizar el complejo *cromosómico* de una *especie* bajo estudio debe incluir dos participantes claves del MCGH: los objetos “Cromosoma” y “Especie”.

En el catálogo, muchas de las soluciones de diseño están conectadas estableciendo relaciones semánticas entre ellas. Estas *relaciones* están representadas por la relación recursiva “RelacionadaA” de la cual deriva la clase asociativa “RelacionadoA” (con el mismo nombre de la relación) que almacena la información sobre la definición y tipo de la relación. Existen cuatro tipos de relaciones que hemos definido basados en las relaciones descritas por Welie (Van Welie, 2003) y semejantes a las relaciones en el modelado de Orientación a Objetos.

Usando el término “patrón” (como equivalente a “solución de diseño” o “proto-patrón”) para simplificar la redacción, las relaciones se definen como:

- **Asociación.** – Los patrones se asocian cuando ocurren en un contexto similar relacionado a un problema de diseño, cuando ellos son alternativas para resolver la misma clase de problemas o cuando contribuyen al logro de una tarea. Esta relación existe entre patrones de una misma categoría. Esta es una relación bidireccional del tipo “**relacionado a**”.
- **Especialización.** – Los patrones pueden ser especializaciones de otros patrones, es decir, un patrón puede ser una versión específica de otro patrón. Esta relación existe entre patrones de una misma categoría y es una relación del tipo “**es un**”.

A parte de la asociación y la especialización, existe la relación del tipo “Agregación” que está representada por las clases “Elemental” y “Compuesta” y que hace referencia a la relación entre un *todo* y sus *partes*: la clase Compuesta (el todo) es un patrón resultante de la agregación de otros patrones más elementales (las partes). Por lo tanto, esta relación se define como:

- **Agregación.** – Una solución de diseño puede usar en su solución a otro patrón, es decir, un patrón puede ser el resultado de “agregar” otros patrones. Es una relación de la forma “**tiene un**”. Esta relación existe entre patrones de diferente o misma categoría.

Los tipos de relaciones serán ilustradas con ejemplos existentes en el dominio genómico más adelante en la 6.2.7.

6.2.2 Identificación de patrones genómicos

Como se mencionó al inicio de esta sección (Sección 6.2), el catálogo de patrones GenomIU está conformado por patrones genéricos y específicos. A pesar de que los patrones genéricos son piezas importantes en el catálogo, su identificación no es de interés en esta Tesis puesto que estos patrones han sido identificados y publicados previamente por otros autores. En efecto, los patrones genéricos del catálogo GenomIU provienen de librerías de patrones publicadas, específicamente de aquellas mencionadas en la sección 4.2.1. Nuestra atención se dirige exclusivamente a la identificación de los patrones genómicos, es decir, los patrones específicos del dominio genómico los cuales son la contribución real de la Tesis.

Los patrones genómicos han sido identificados considerando un enfoque de resolución de problemas y basado en el estudio de ejemplos. Con *resolución de problemas* nos referimos a que para identificar los patrones primero identificamos los problemas y después buscamos solución a dichos problemas. Con *estudio de ejemplos* nos referimos a que los patrones son obtenidos de la observación y análisis de interfaces de usuario implementadas en herramientas existentes. Seffah (Seffah, 2010) describe al “estudio de ejemplos” como una estrategia adecuada para la identificación y captura de patrones. Escogimos esta estrategia debido a su carácter observacional de factible realización frente a la alternativa de contar con un experto o mentor en diseño o usabilidad o contar con el conocimiento profundo en directrices y estilos de diseño.

Con base en este enfoque, el proceso sistemático para identificar los patrones incluye en términos generales: i) estudiar ejemplos de interfaces, ii) identificar en estos ejemplos los diseños que han sido utilizados repetitivamente como soluciones a los problemas de interacción y iii) documentar las soluciones. La Imagen 6.4 muestra en detalle el proceso mencionado y a continuación se detallan sus pasos:

Paso 1: Seleccionar Problema. – El punto de entrada del proceso es el problema de interacción que queremos resolver.

Paso 2: Buscar en librerías existentes. – Muchas de las soluciones de diseño que estamos buscando pueden haber sido documentadas previamente por otros autores. Por lo tanto, buscamos patrones que aborden el problema bajo estudio en librerías de patrones publicadas por otros autores, específicamente en las librerías publicadas en línea que hemos identificado en la Tabla 5 ya que estas librerías son usadas con mayor frecuencia por desarrolladores y diseñadores de interfaces web. Si existen patrones que aborden el problema, los seleccionamos y seguimos al paso 7, caso contrario seguimos al paso 3.

Paso 3: Seleccionar herramientas. – Buscamos herramientas de software de acceso a datos y seleccionamos aquellas que abordan total o parcialmente la tarea/problema seleccionado en el paso 1. Para esto, utilizamos un *inventario de herramientas* que se describe más adelante en la Sección 6.2.4.

Paso 4: Inspeccionar Interfaces de Usuario. – Coleccionar capturas de pantalla de las interfaces de usuario de cada herramienta seleccionada e inspeccionarlas visualmente. La inspección consiste en estudiar los diseños que han sido implementadas por los diseñadores para solucionar el problema de interacción bajo estudio.

Paso 5: Identificar soluciones recurrentes. – A partir de la inspección visual en el paso anterior, identificamos las soluciones de diseño implementadas repetitivamente para solucionar el problema de interacción. Si una solución de diseño ha sido utilizada en al menos tres interfaces de usuario diferentes para solucionar el problema bajo estudio, entonces consideramos a esta solución de diseño como “patrón”, de acuerdo a la “Regla de Tres” (Appleton, 2000) (Seffah, 2010). Caso contrario, si la solución de diseño ha sido utilizada para tratar el problema pero no ha sido implementada en al menos tres interfaces de usuario diferentes, entonces consideramos a esta solución de diseño como solución candidata o “proto-patrón” (Appleton, 2000).

Paso 6: Documentar solución. – Documentar las soluciones que fueron consideradas como patrones o proto-patrones en el paso 5.

Paso 7: Documentar solución existente. – El patrón encontrado y publicado por otros autores debe ser adaptado (aumentando o reduciendo la información del patrón original) a la realidad del dominio genómico. Las soluciones adaptadas en este paso presentan un alias (nombre asignado por el autor original del patrón y la dirección de página web o referencia bibliográfica del patrón original).

Paso 8: Agregar al catálogo. – Los patrones y proto-patrones documentados son agregados al catálogo GenomIU dentro de la categoría “patrones genómicos”.

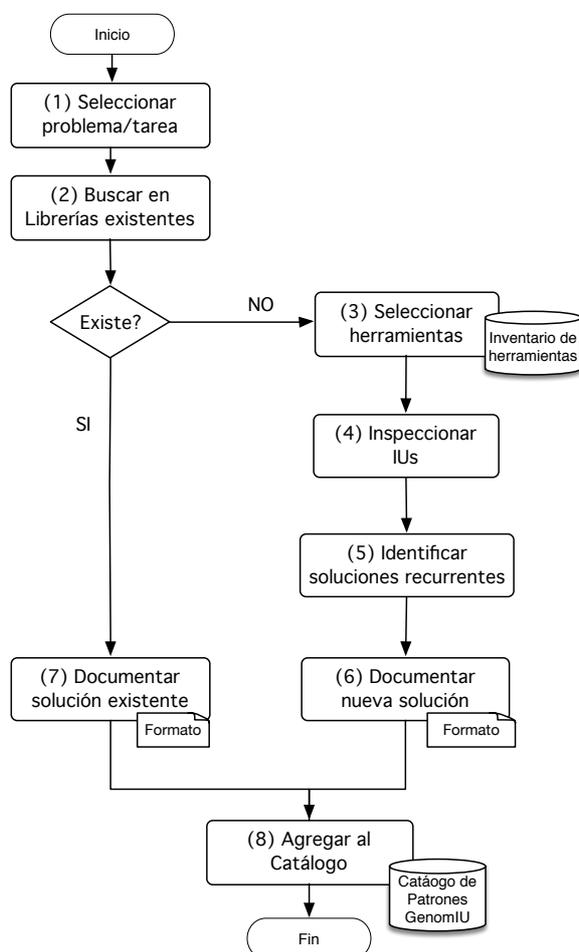


Imagen 6.4 Proceso para identificar patrones genómicos

6.2.3 Evolución de los patrones genómicos

Durante la investigación de esta tesis, el catálogo de patrones ha evolucionado continuamente y en base a una perspectiva “deductiva”. Con *deductivo* nos referimos a que, fundamentados en la idea de que manejar datos genómicos significa manejar datos masivos (Stephens et al., 2015), en las primeras iteraciones de la identificación de patrones nos concentramos en buscar soluciones de diseño a problemas generales de interacción con datos masivos para luego concentrarnos en las soluciones a problemas de interacción específicos del dominio genómico. De esta manera, el primer conjunto de patrones consistió en patrones dirigidos a solucionar problemas de interacción en el dominio Big Data, y

paulatinamente, el conjunto de patrones fue creciendo, incluyendo patrones más enfocados en el dominio genómico.

El primer conjunto de patrones fue publicado en (C. E. Iñiguez-Jarrín et al., 2018) y estuvo conformado por cuatro patrones que fueron el resultado de estudiar las interfaces de usuario de las herramientas utilizadas en el dominio de Big Data. Estos patrones se resumen a continuación:

Nombre: visualizar, conectar, filtrar

Problema: ¿Cómo puede el usuario visualizar los múltiples cambios en el comportamiento de los datos causados por un simple filtro de datos?

Solución: Proporcionar un lienzo en el que el usuario pueda crear gráficos de datos interactivos y colocar y organizar los gráficos de acuerdo con sus necesidades. Cuando el usuario aplica una condición de filtro al interactuar con un gráfico de datos, los gráficos restantes responden automáticamente actualizando su estado de acuerdo con la condición del filtro que afecta a todo el conjunto de datos.

Nombre: Caja de Operaciones

Problema: ¿Cómo aplicar operaciones de conjuntos sobre los datos?

Solución: Proporcione al usuario un conjunto de operaciones (p. ej., unión, intersección, complemento, diferencia) para computar conjuntos de datos que compartan al menos un atributo común del esquema de datos. El usuario debe seleccionar como entrada, al menos dos conjuntos de datos y luego seleccionar la operación a realizar. El nuevo conjunto de datos resultante de la comparación puede ser reutilizado en posteriores operaciones.

Nombre: Entrega implícita

Problema: ¿Cómo se puede reducir el efecto de retardo producido al mostrar y navegar a través de grandes volúmenes de datos?

Solución: Recuperar y entregar los datos a la interfaz de usuario a medida que el usuario navega a través de un conjunto de datos de interés. Los datos se cargan solo cuando el usuario expresa implícitamente su intención. Al contrario de la intención explícita en la que el usuario usa, por ejemplo, el control de paginación para moverse por una tabla, la intención implícita permite al usuario moverse libremente a través de la tabla sin indicar la página que quiere visitar. Esta solución es aplicable para recuperar cualquier objeto de datos del esquema de datos subyacente.

Nombre: Filtro móvil

Problema: ¿Cómo puede el usuario definir condiciones de filtro y aplicarlas de manera fluida dentro del espacio de análisis?

Solución: permita al usuario crear condiciones de filtro a partir de cualquier atributo existente en el esquema de datos subyacente. Esto minimiza los movimientos innecesarios del usuario en el espacio de análisis al permitirle colocar los filtros en cualquier lugar de este espacio.

Para validar la aplicabilidad de estos cuatro patrones en el diseño de interfaces, realizamos un experimento controlado que evaluó el impacto causado por estos patrones en la usabilidad de interfaces de usuario. Esta validación se presenta en la Sección 7.2. como parte de la validación de esta Tesis.

A partir de la retroalimentación obtenida del experimento y las sugerencias recibidas en el congreso científico donde los patrones fueron presentados (CAiSE 2018), realizamos

ajustes a los patrones, incluyendo cambios en la descripción o en el nombre del patrón e incluso la simplificación de un patrón en patrones más concretos. Por ejemplo, el patrón “entrega implícita” fue renombrado “navegación continua” con el fin de expresar el propósito concreto de la solución del patrón; el patrón “filtro móvil” sufrió modificaciones en su nombre y descripción para adaptarlo al dominio genómico (actualmente se denomina “filtro genómico”); el patrón “caja de operaciones” fue cambiado el nombre a “operación de conjuntos”; el patrón “visualizar, conectar, filtrar”, con una descripción compleja, fue separado en patrones más elementales de visualización y filtrado de datos (estos patrones actualmente se denominan patrón “gráfico” y “filtro genómico”).

Aunque este primer esfuerzo de identificación produjo únicamente cuatro patrones, este esfuerzo marcó un punto de partida para la progresiva identificación de nuevas soluciones de diseño cada vez más orientadas al dominio genómico. Es así como, después iteraciones del proceso de identificación de patrones, reunimos un total de 13 patrones que se convierten en soluciones a los problemas de interacción identificados y enumerados en la Sección 5.2.2. La Tabla 9 muestra los patrones identificados y su correspondencia con el problema de consumo de datos que aborda.

Tabla 9 Problemas cubiertos por los patrones

Nombre Patrón	Problemas																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Cargar Datos																	X
Operación de conjuntos						X											
Secuencia de Acciones											X						
Buscar	X																
Filtro Genómico		X	X		X					X							
Selector de regiones												X					
Gráfico				X													
Idiograma														X			
Pista Comparativa									X								
Columnas Ocultas								X									
Círculo cromosómico													X				
Navegador genómico															X		
Navegación Continua							X										

Este conjunto de patrones es el resultado de la ejecución iterativa del proceso de identificación de patrones (Sección 6.2.2) durante el tiempo de este proyecto de investigación.

6.2.4 Inventario de Herramientas: Interfaces de Usuario estudiadas

En esta sección describimos el “inventario de herramientas”, es decir, el conjunto de herramientas cuyas interfaces de usuario hemos estudiado para capturar los patrones genómicos. El inventario contiene herramientas existentes en la industria y en la academia y se convierte en un recurso importante en nuestra investigación. Las interfaces de usuario de dichas herramientas representan la fuente de búsqueda de soluciones de diseño y la evidencia que sustenta a los patrones genómicos.

El inventario está en constante actualización ya que es el resultado de la búsqueda continua de herramientas utilizadas en el análisis terciario del análisis genómico y en el manejo masivo de datos.

Par mejor explicación del inventario, clasificamos las herramientas en función del tipo de interfaces que implementan. Es decir, describimos las características de un tipo de interfaz y luego listamos las herramientas que implementan dicho tipo de interfaz. Así, tenemos tres categorías de interfaces de usuario: visualización de la secuencia genómica, búsqueda de bibliografía y análisis de secuencia.

Visualización de genoma. – Estas interfaces de usuario contienen visualizaciones potentes que permiten analizar, en diferentes escalas, los elementos del genoma. Navegadores genómicos y herramientas de visualización independientes implementan este tipo de interfaces. Las visualizaciones incorporadas en las interfaces están generalmente generadas con librerías gráficas independientes. La Tabla 10 lista las herramientas y librerías involucradas en este tipo de interfaces de usuario.

Tabla 10 Herramientas y librerías de visualización genómica

Navegadores		
Nombre	Referencia	Soporte
G-compass	http://www.h-invitational.jp/g-compass/	Web
Ensembl Genome Browser	https://www.ensembl.org/index.html	Web
Decipher	https://decipher.sanger.ac.uk/	Web
TumorMap	https://tumormap.ucsc.edu/?#	Web
Genome Data Viewer	https://www.ncbi.nlm.nih.gov/genome/gdv/browser	Web
VEGA	http://vega.archive.ensembl.org/Homo_sapiens/Info/Index	Web
GBrowse	http://www.gmod.org/wiki/GBrowse	Web
JBrowse	https://jbrowse.org/	Web
Gramene	http://www.gramene.org/	Web
UCSC Xena	https://xena.ucsc.edu/	Web
LookSeq	https://www.sanger.ac.uk/science/tools/lookseq	Web
Sequece Viewer	https://www.ncbi.nlm.nih.gov/projects/sviewer/	Web
Genome Maps	http://genomemaps.org/	Web
UTGB	http://utgenome.org/	Web
GEMINI	https://gemini.readthedocs.io/en/latest/content/browser.html	Web
Integrative Genomics Viewer	http://software.broadinstitute.org/software/igv/home	Escritorio
Herramientas de visualización independientes		
Nombre	Referencia	Soporte

SnapGene Viewer	https://www.snapgene.com/snapgene-viewer/	Escritorio
J-Circos	https://sourceforge.net/projects/jcircos/	Escritorio
HEFaMp	https://hb.flatironinstitute.org/	Web
Genoma Decorator Page	https://www.ncbi.nlm.nih.gov/genome/tools/gdp	Web
Librerías		
Nombre	Referencia	Lenguaje*
ChromoWheel	http://chromowheel.cgb.ki.se	Html, SVG
CGView	http://wishart.biology.ualberta.ca/cgview/	Java
Ideogram.js	https://eweitz.github.io/ideogram/	JavaScript
Circos	http://circos.ca/	Perl
BioCircos.js	http://bioinfo.ibp.ac.cn/biocircos/	JavaScript
RCircos	https://bitbucket.org/henryhzhang/rcircos/src/master/	R
BioCircos (para R)	https://github.com/lvulliard/BioCircos.R	R

* Lenguaje de programación base necesario para su ejecución.

Los navegadores genómicos son entornos interactivos que permiten visualizar la posición de los elementos genómicos en la secuencia genómica e integrar información de diferentes fuentes de datos genómicos para analizar los elementos genómicos en contexto. Las interfaces de los navegadores genómicos se caracterizan por ofrecer visualización del tipo “resumen más detalle” (una vista de mayor nivel que resume la secuencia genómica completa y otra de menor nivel que presenta los detalles de una región de la secuencia seleccionada en la vista de mayor nivel) y controles para navegar por la visualización (p. ej., acercamiento, alejamiento, desplazamiento).

Las herramientas de visualización independientes implementan interfaces de usuario que permiten generar visualizaciones del genoma las cuales pueden ser configuradas previamente por el usuario a través de paneles de configuración incluidos en la interfaz y que contienen parámetros de visualización como son alineación (horizontal, vertical), etiquetas de leyenda y tamaño de letra, color (paleta de colores para diferentes figuras). La mayoría de las interfaces de este tipo incorporan controles para explorar la visualización y la opción para descargar la imagen producida. Por ejemplo, J-Circos (An et al., 2015) es una herramienta desarrollada en lenguaje Java que lleva el concepto visual y estático de Circos (Krzywinski et al., 2009) (un paquete de software para visualizar datos en diseño circular y que es aplicado a los datos genómicos) a una visualización interactiva permitiendo a los investigadores estudiar de mejor manera interacciones cromosómicas complejas que son difíciles de visualizar a partir de los datos secuenciados del genoma. Genoma Decorator Page es un servicio web del NCBI que permite representar el complejo cromosómico de una especie seleccionable y ofrece la opción de establecer los parámetros de visualización como alineación del gráfico, alineación de las etiquetas del gráfico y permite descargar la visualización en varios formatos (p. ej., pdf, png, jpg, eps).

Por otro lado, existe una gran variedad de librerías gráficas que son utilizadas por los desarrolladores para generar visualizaciones del genoma e integrarlas a interfaces de usuario. Algunas de las librerías generan visualizaciones estáticas o interactivas que pueden ser integradas en la web y otras que se comportan como un ensamblado que debe ser integrado

a una plataforma para funcionar. Por ejemplo, Ideogram.js y BioCircos.js son librerías de JavaScript que, incluidas en páginas web, generan visualizaciones de los cromosomas, en el caso de Ideogram.js, y de las relaciones entre los genes o posiciones cromosómicas, en el caso de BioCircos.js. RCircos (Zhang et al., 2013) y BCircos (Vulliard et al., 2019) (una adaptación de BioCircos.js para R) son librerías gráficas que generan imágenes estáticas para visualizar datos genómicos que funciona como paquetes de R.

Búsqueda de bibliografía. – Las interfaces de usuario de las herramientas de búsqueda de bibliografía biomédica, en su mayoría implementadas en plataforma web, soportan la recuperación o el análisis de literatura biomédica. La Tabla 11 lista las herramientas que implementan este tipo de interfaces. Por ejemplo, PubMed es un sitio web utilizado para recuperar información de la literatura biomédica que contiene hasta la fecha consultada⁴ (25 de julio 25 de 2019), más de 30 millones de citaciones de artículos médicos-científicos completos almacenados en la base de datos MEDLINE, una de las bases más importantes en la medicina. Las interfaces de usuario de búsqueda de bibliografía permiten analizar los enlaces en la literatura, autores, citaciones, etc. Este es el caso de la herramienta VOSviewer, un sistema que se ejecuta como programa de escritorio que acepta la lista de bibliografía exportada de PubMed y que analiza y muestra gráficamente las relaciones bibliográficas. Otra herramienta que se enfoca en analizar bibliografía de PubMed es PubNet (Douglas et al., 2005), sin embargo, esta herramienta no consta en la lista debido a que no existe referencia de acceso a la herramienta y no ha sido posible analizarla.

Tabla 11 Herramientas de bibliografía biomédica

Nombre	URL	Soporte
Dimensions	https://www.dimensions.ai/	Web
VOSviewer	http://www.vosviewer.com/	Escritorio
PubMed	https://www.ncbi.nlm.nih.gov/pubmed/	Web
BioGPS	http://biogps.org	Web
MasterMind	https://mastermind.genomenon.com/	Web

Análisis de secuencia. – Las interfaces de usuario de herramientas de análisis de secuencia permiten acceder y trabajar con datos de variaciones genéticas obtenidos a partir de ficheros de texto (p. ej., VCF). Algunas de estas interfaces son web y otras interfaces están basadas en línea de comandos. La Tabla 12 lista las herramientas que implementan este tipo de interfaces de usuario.

Tabla 12 Herramientas de análisis de variaciones genéticas

Nombre	URL	Soporte
ICGC Data Portal	https://dcc.icgc.org/	Web
VCF Tools	https://vcftools.github.io/index.html	Consola
VCF Miner	https://github.com/Steven-N-Hart/vcf-miner	web
Galaxy	https://usegalaxy.org/	Web
BiERapp (Alonso et al., 2015)	http://www.babelomics.org/	Web

⁴ <https://www.ncbi.nlm.nih.gov/books/NBK3827/>

Gem.app (formalmente GENESIS 2.0)	https://www.genesis-app.com	Web
Sequence Variant Analyzer	(URL del proyecto) http://www.svaproject.org/	Escritorio
DraGnET	http://www.dragnet.cvm.iastate.edu/	Web

Las interfaces de análisis de secuencia proveen mecanismos para cargar muestras genéticas y a partir de estas, listar, filtrar y visualizar variaciones genómicas o genes para analizar estos datos integrando información contextual genómica de fuentes externas. En este tipo de interfaces, la tabla de datos es uno de los controles de interfaz de usuario más utilizados para presentar, especialmente, cantidades masivas de registros de variaciones genéticas que contienen una larga lista de propiedades.

En cierta medida, la característica “analítica” de estas interfaces hace que muchos de los controles, funcionalidades y características de los dos tipos de interfaces mencionados anteriormente (es decir, búsqueda de bibliografía y visualización del genoma) estén incluidas en este tipo de interfaces. Por ejemplo, BiERapp (Alemán et al., 2014), una herramienta que asiste en la priorización de variaciones genéticas, incluye en su interfaz de usuario una lista de literatura biomédica asociada a la variación bajo estudio y el navegador genómico Genome Maps para analizar una variación genética en su contexto genómico.

Análisis de datos masivos. – Partiendo de que manejar datos genómicos significa manejar datos masivos, hemos considerado incluir en el inventario las herramientas que tratan con datos masivos. Los tipos de interfaces de usuario hasta aquí mencionadas están presentes en herramientas que forman parte del panorama genómico. Sin embargo, existe un tipo de interfaces de usuario presente en herramientas que están fuera del dominio genómico y que tratan con grandes cantidades de datos. Este tipo de interfaces puede ser de mucha ayuda para abordar los problemas de interacción con los datos masivos en el dominio genómico. Las interfaces de usuario de este tipo permiten al usuario crear sus propias visualizaciones de datos e incluso arrastrar y soltar las visualizaciones para crear cuadros de mando o “dashboards” que ofrecen una vista resumida de los datos con la que el usuario puede interactuar para extraer conocimiento de los datos. Estas interfaces están implementadas particularmente en herramientas de inteligencia de negocios y son utilizadas generalmente para tratar con Big Data. Las herramientas que implementan este tipo de interfaces se listan en la Tabla 13 y son las herramientas de inteligencia de negocios reportadas en el cuadrante mágico de Gartner publicado en febrero del 2017 (Richardson et al., 2017).

Tabla 13 Herramientas de inteligencia de negocio y analítica de datos

Nombre	URL
ZoomData	https://www.zoomdata.com/
Cognos IBM	https://www.ibm.com/es-es/products/cognos-analytics
Tableau	https://www.tableau.com
Qlik	https://www.qlik.com/us/
Tibco	https://www.tibco.com
Oracle Big Data Discovery	https://cloud.oracle.com/bigdata_discovery_saas

6.2.5 Formato para describir los patrones

No existe una regla general para determinar el formato de documentación de patrones. Para documentar los patrones, hemos estructurado un formato basándonos en el trabajo de Kruschitz et al. (Kruschitz, 2010) que, a partir del estudio de varios catálogos de patrones existentes, delinea un conjunto mínimo de partes necesarias para documentar un patrón.

Tabla 14 Formato para documentar patrones GenomIU

Parte	Descripción
Nombre	Nombre corto para identificar fácilmente al patrón y que describa el problema de diseño que aborda el patrón y su solución.
Tipo	“Patrón” o “Proto-patrón”, dependiendo del número de ejemplos. (Ver paso 5 de la Sección 6.2.2)
Alias	Nombre del patrón asignado por otros autores.
Imagen	Una captura de pantalla o esquema arquetípica que sintetiza la idea principal del patrón permitiendo al lector comprender de manera rápida la solución.
Problema	Describe lo que hace o lo que trata de solucionar el patrón en no más de dos oraciones o preguntas.
Contexto	Describe dónde y cómo ocurre el problema. El contexto permite al diseñador seleccionar el patrón adecuado a su realidad estableciendo analogías entre su situación particular y la situación descrita. Describe cuándo el patrón es aplicable.
Solución	Describe la manera en la que puede ser resuelto el problema mencionando los elementos esenciales que intervienen en la solución. La descripción debe ser suficientemente genérica, capturando las propiedades constantes y comunes a todas las interfaces de usuario donde se ha tenido éxito en solucionar el problema, para permitir al lector adaptar la solución a su contexto particular.
Esquema	Esta parte es exclusiva de los patrones específicos (patrones genómicos) y presenta cómo la solución del patrón se relaciona con los conceptos del modelo conceptual del genoma. Para esto, el esquema presenta la lógica (Controlador) y los datos (Modelo) involucrados en la solución en términos del patrón de arquitectura de software Modelo-Vista-Controlador (MVC); un patrón comúnmente usado en las interfaces de usuario. Las relaciones son útiles al momento de instanciar el patrón.
Patrones relacionados	Lista otros patrones relacionados con el patrón actual, siguiendo la semántica de relaciones definida en el modelo conceptual del catálogo (Sección 6.2.1): asociación, especialización, agregación.
Ejemplos	Muestra usos conocidos del patrón a través de capturas de interfaces de usuario en las que aparece implementado el patrón.

Consideramos apropiada la propuesta de Kruschitz et. al. ya que ha sido utilizada sin variaciones por las librerías de patrones más populares entre los diseñadores de interfaces como son (Tidwell, 2010) o (Toxboe, 2007). Adicional a las partes mínimas definidas por Kruschitz et. al., agregamos la parte “Esquema” para presentar el modelo conceptual que hace explícita la relación del patrón con los objetos del MCGH. El formato para documentar los patrones GenomIU se describe en la Tabla 14.

6.2.6 Patrones Genómicos

En esta sección describimos cada uno de los patrones genómicos siguiendo el formato descrito en la Sección 6.2.5. Es importante mencionar que la descripción de los patrones incluye capturas de pantalla de interfaces de usuario de las herramientas del “inventario de herramientas” conformado por la Tabla 10, Tabla 11, Tabla 12 y Tabla 13. Para evitar la sobrecarga de texto y hacer la descripción de los patrones lo más simple posible, hemos indicado en la descripción el nombre de la herramienta que implementa la interfaz de usuario, pero hemos suprimido su referencia. No obstante, el lector puede visitar las tablas que conforman el inventario de herramientas para conocer la referencia de cada herramienta. Estos patrones están también disponibles en línea a través de: <http://ciniguez.github.io/Genomic-UI-Patterns>.

Cargar Datos

Alias: Ninguno

Tipo: Patrón

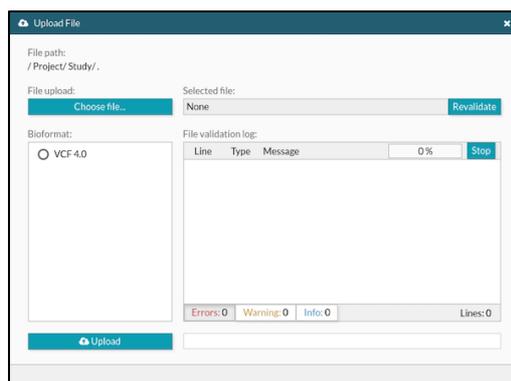


Imagen 6.5 Formulario de carga de datos de la aplicación Bier.app

Problema

Cargar conjuntos de datos genéticos disponibles en diferentes fuentes y formatos.

Contexto

Previo a la ejecución de consultas y operaciones sobre los datos, el usuario (novato o experto) necesitan importar dichos datos. Los datos pueden estar disponibles localmente o en recursos externos, así como en varios formatos (p. ej., BAM, VCF, PED).

Solución

Proveer un formulario de carga de datos que le permita al usuario realizar la siguiente secuencia de acciones:

- 1) Seleccionar el origen/fuente de datos (p. ej., recurso en línea, fichero de texto, texto). Permitir varias opciones de ingreso de datos. Por ejemplo, si el conjunto de datos se encuentra en uno o varios ficheros locales, permitir al usuario navegar por el repositorio local y seleccionar los ficheros; si el conjunto de datos es accesible por la web, permitir al

usuario indicar la URL del recurso; si el conjunto de datos es textual, permitir al usuario copiar y pegar el contenido en el formulario.

2) Opcionalmente, indicar el formato del contenido (p. ej., VCF);

3) Validar el contenido; los resultados de la validación del contenido realizada internamente por el sistema deben ser mostrados en pantalla.

4) Importar el contenido si los datos son válidos. Las acciones para importar los datos deben incluir “iniciar” y “cancelar”. Variaciones en las acciones de importación incluyen la acción “pausar” para interrumpir temporalmente el proceso. Es recomendable informar en tiempo real al usuario sobre el estado del procesamiento de los datos indicando como, por ejemplo, el número de ficheros cargados hasta el momento o el porcentaje de avance en la lectura del contenido.

Esquema

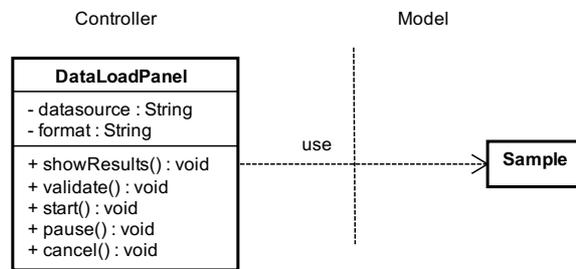


Imagen 6.6 Esquema del patrón Cargar Datos

El patrón se crea instanciando la clase “DataLoadPanel”. Los atributos *fFuente de datos* (“datasource”) y *formato* (“format”) son internos de la clase. Los métodos de la clase se corresponden con las posibles acciones que puede realizar el usuario. El resultado de la importación se almacena como una muestra genómica (clase “Sample”) del modelo conceptual del genoma.

Patrones relacionados

Asociación: relacionado a **Operador de Conjuntos**.

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

La Imagen 6.7 muestra el formulario de carga de datos de la aplicación ICGC Data Portal donde el usuario copia y pega los datos a importar o escoge un fichero a cargar.

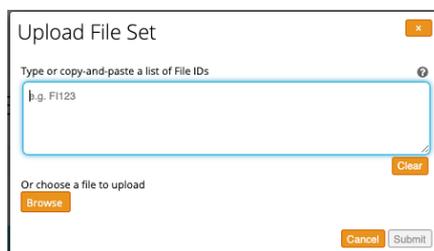


Imagen 6.7 Formulario de carga de datos. Fuente: ICGC Data Portal

Otras interfaces de aplicaciones donde se ha implementado el patrón incluyen: VCF Miner, Galaxy, Bier.app.

Operación de Conjuntos

Alias: ninguno

Tipo: Proto-patrón

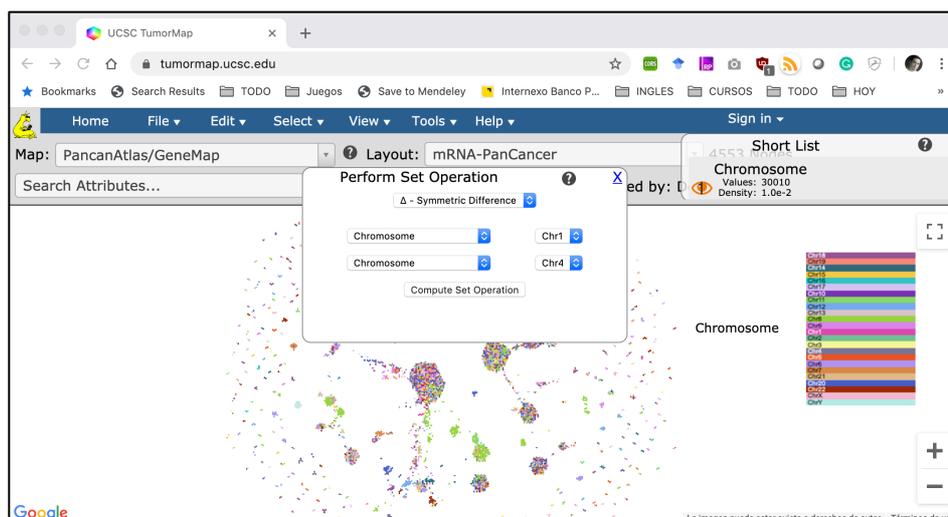


Imagen 6.8 Operación de conjuntos entre cromosomas. Fuente: TumorMap

Problema

Analizar diferencias o similitudes entre conjuntos de datos genómicos.

Contexto

El usuario necesita identificar diferencias o similitudes entre conjuntos de datos genómicos del mismo tipo (p. ej. cromosomas con cromosomas, genes con genes) respecto a un elemento en común (p. ej., variaciones genéticas). Los usuarios pueden utilizar esta solución independientemente de su nivel de experiencia.

Solución

Permitir al usuario: 1) seleccionar, de un listado de datos disponibles, los datos o grupos de datos para la operación. Al menos dos datos o grupos de datos del mismo tipo (p. ej. dos genes o dos grupos de genes) deben ser seleccionados para ejecutar la operación. La lista de datos disponible puede contener datos individuales o grupos de datos creados por el usuario a partir de los datos individuales; 2) seleccionar la operación de conjuntos a ejecutar. Las operaciones incluyen, entre otras, la unión, intersección y diferencia; 3) visualizar el resultado de la operación de varias maneras tales como un resumen cuantitativo, el detalle de los datos que cumplieron la operación o el resumen acompañado del detalle.

Una variación en la secuencia del punto 2 y 3, que se presenta en la Imagen 6.10, puede ser: 2) ejecutar todas las posibles operaciones entre los datos y mostrar los subconjuntos resultantes de las operaciones; 3) permitir al usuario seleccionar el o los subconjuntos de datos resultantes para ver su detalle.

El resultado de una operación es un nuevo conjunto de datos que puede ser almacenado en el listado de datos disponibles.

Esquema

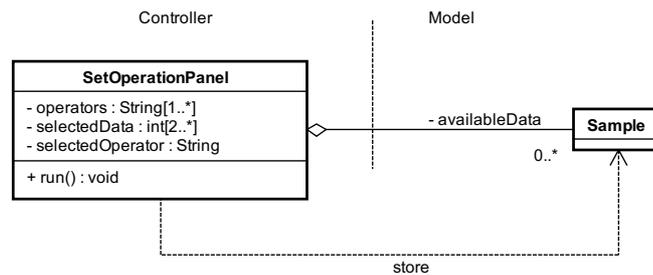


Imagen 6.9 Esquema del patrón Operación de Conjuntos

El patrón operación de conjuntos se crea instanciando la clase “SetOperationPanel”. El atributo “operators” contiene las posibles operaciones que se pueden realizar entre los datos. El atributo “selectedOperator” es la operación seleccionada por el usuario. Para realizar la operación de conjuntos, el sistema provee de una lista de conjuntos de datos de los cuales el usuario selecciona al menos dos datos o grupos de datos para ejecutar la operación. La lista de conjuntos de datos proviene de la clase “Sample” que representa a elementos del modelo conceptual que son sujetos a comparación (p. ej., genes, cromosomas, especies). El evento “run” procesa la operación de conjuntos.

Patrones relacionados

Asociación: relacionado a **Cargador de datos**, **Secuencia de acciones**, **Filtro Genómico**.

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

La Imagen 6.10 es una sección de la interfaz de usuario del sitio ICGC Data Portal que muestra a través de un diagrama de Venn los resultados de las posibles operaciones de conjuntos entre tres muestras de mutaciones (S_1 , S_2 , S_3). Las muestras son parte de las muestras disponibles en el sistema y han sido previamente seleccionadas por el usuario para realizar las operaciones. Cada sector del gráfico es un subconjunto de datos e indica el número de mutaciones. Por ejemplo, el número 17 en el centro del diagrama indica el número de mutaciones del subconjunto resultante de la intersección de las tres muestras ($S_1 \cap S_2 \cap S_3$). El usuario puede seleccionar el o los subconjuntos de resultados que utilizará para un posterior análisis.

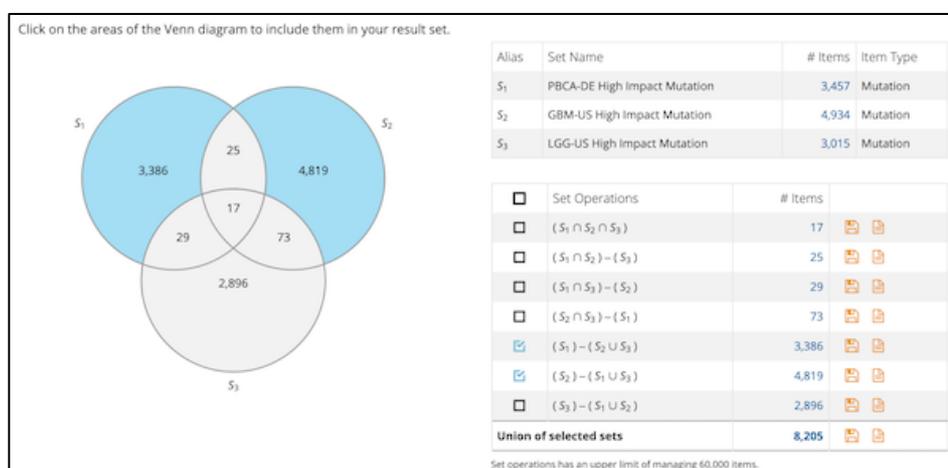


Imagen 6.10 Operación de conjuntos de muestras. Fuente: ICGC Portal

Secuencia de Acciones

Alias: *Command History* (Tidwell, 2010)

Tipo: Patrón

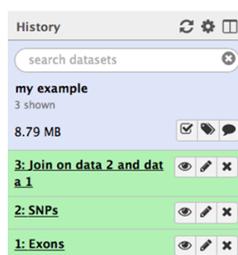


Imagen 6.11 Secuencia de acciones de Galaxy. Fuente: Tutorial Galaxy⁵

⁵ <https://galaxyproject.org/tutorials/g101/>

Problema

Mostrar las acciones ejecutadas sobre un conjunto de datos genómicos, ordenadas por su secuencia de ejecución.

Contexto

El usuario realiza acciones sobre los datos (p.ej., filtros) para obtener un conjunto de datos relevante que cumpla con tales condiciones. La cantidad de acciones es larga y su secuencia de ejecución es difícil de recordar para un usuario que desea repetir o modificar la secuencia de condiciones sobre los datos.

Solución

Presentar un listado de ítems donde cada ítem representa una acción realizada por el usuario en el análisis de datos. Por cada acción realizada, un ítem es agregado a la lista de acciones. Así, el orden de acciones se conserva. Cada ítem de la lista puede ser editado o eliminado por el usuario. Si aplica, cada ítem debe indicar el número de datos resultantes después de aplicar la acción.

Esta solución evita la sobrecarga de cognitiva del usuario de memorizar cada acción realizada sobre el conjunto de datos. El usuario puede revisar las acciones realizadas en el análisis de datos genómico e incluso repetir la secuencia de acciones. La lista de acciones es de mucha ayuda si se necesita evidenciar los resultados.

Esquema

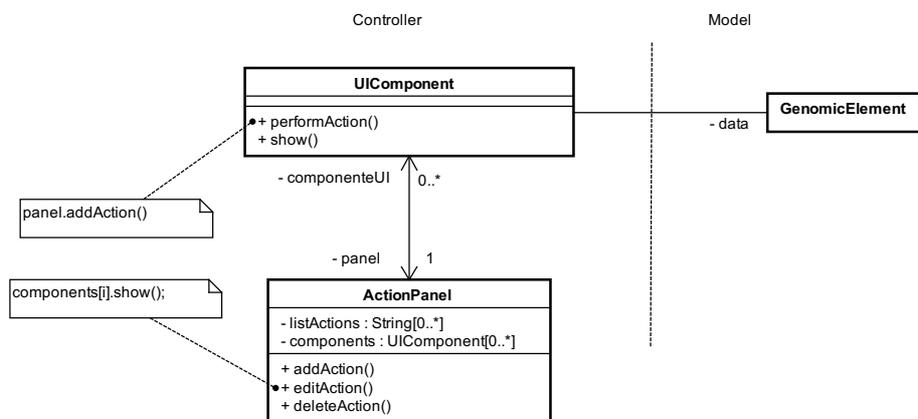


Imagen 6.12 Esquema del patrón Secuencia de Acciones

El patrón se crea instanciando la clase “ActionPanel” que representa a un panel de la interfaz de usuario que presenta la secuencia de acciones realizadas sobre los componentes de la interfaz de usuario (clase “UIComponent”). En efecto, el atributo “ActionPanel.components” contiene las referencias a los componentes que están conectados al panel de acciones. Cuando el usuario realiza una acción sobre el componente (“UIComponent.performAction()”), esta acción se agrega (“ActionPanel.addAction()”) a la

lista de acciones (“ActionPanel.listActions”) del panel, respetando el orden de ejecución de las acciones.

Si el usuario edita una acción registrada en el panel de acciones, el evento “ActionPanel.editAction()” se ejecuta y despliega el UIComponent para editar los valores de la acción. Si el usuario elimina una acción, la acción es eliminada de la lista de acciones del panel.

Patrones relacionados

Asociación: relacionado a **Operador de Conjuntos, Filtro Genómico**.

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

En la Imagen 6.13, la sección “Included Filters” (etiqueta 1) muestra la secuencia de cuatro filtros (Projects, Clinical Defaults, Clinical Significance y GENESIS Allele Counts) aplicados por el usuario a un conjunto de datos de secuencia genética. La etiqueta 1 resalta el filtro “Clinical Significance” el cual se ha pulsado sobre el ícono de lápiz para ser editado. El formulario de edición del filtro “Clinical Significance” se muestran en la etiqueta 2. El formulario en la etiqueta 2 es diferente para cada uno de los ítems de la lista de filtros seleccionados.

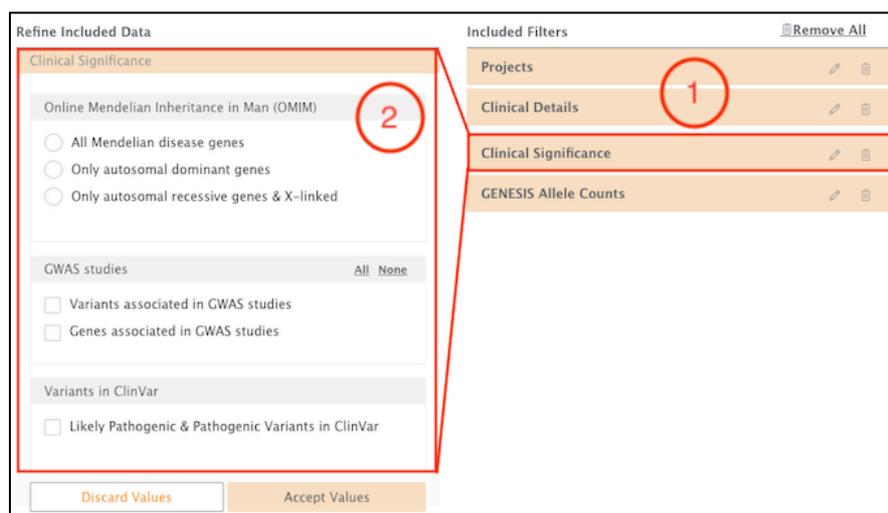


Imagen 6.13 Porción de interfaz de la aplicación Genesis 2.0

Buscar

Alias: ninguno

Tipo: Patrón



Imagen 6.14 Patrón buscar. Fuente: aplicación BioGPS

Problema

Buscar información de elementos genómicos.

Contexto

El usuario explora la información genómica partiendo de una búsqueda directa y sencilla de elementos genómicos. El usuario tiene poco conocimiento sobre la estructura de la información genómica y las propiedades intrínsecas. Por lo tanto, el usuario necesita guía sobre qué buscar y cómo buscar en el conjunto de datos genómico.

Solución

Permitir al usuario ingresar un término de búsqueda relacionado un elemento específico del modelo (p. ej., variación, gene, fenotipo) y que es de interés para el usuario. Para guiar la búsqueda, se recomienda mostrar sugerencias (o ejemplos) sobre términos de búsqueda que pueden ser ingresados y que producen resultados reales al ser consultados.

El término de búsqueda debe soportar varias alternativas de búsqueda relacionadas a las propiedades del elemento genómico. Por ejemplo, para buscar información de un gen, el usuario puede indicar como término de búsqueda cualquiera de las dos opciones: el símbolo del gen (es decir, la abreviación de su nombre) o el identificador del gen en la base de conocimiento Gen Ontology. Es recomendable flexibilizar la búsqueda permitiendo que el término de búsqueda a) pueda ser una lista de términos (p. ej., una lista de símbolos de genes) o b) permita el uso de comodines (* ó ?) par encontrar coincidencias.

Esquema

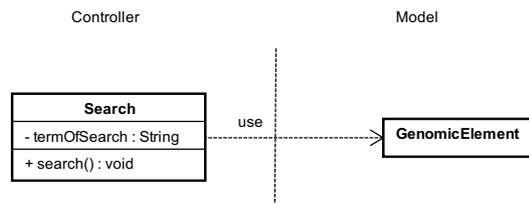


Imagen 6.15 Esquema del patrón Buscar

El patrón se crea instanciando la clase “Search”. El atributo “termOfSearch” captura el término de búsqueda ingresado por el usuario. El método “search()”, ejecutado cuando el

usuario solicita buscar, construye la cadena de consulta utilizando el término de búsqueda y realiza la consulta de datos al elemento específico del modelo (clase “GenomicElement”).

Patrones relacionados

Asociación: relacionado a **Columns ocultas**.

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

La Imagen 6.16 es una porción de la interfaz de usuario principal del sitio Mastermind que presenta tres cajas de texto, una por cada alternativa de búsqueda (enfermedad, gen, variación genética). El texto en cada caja de texto sugiere la información que se puede buscar y la forma de ingresar el término de búsqueda. Así, una variación genética puede ser buscada indicando su identificador expresado en nomenclatura HGVS (p. ej., c.179T>A) o el cambio que produce en la proteína (p. ej., V600E).

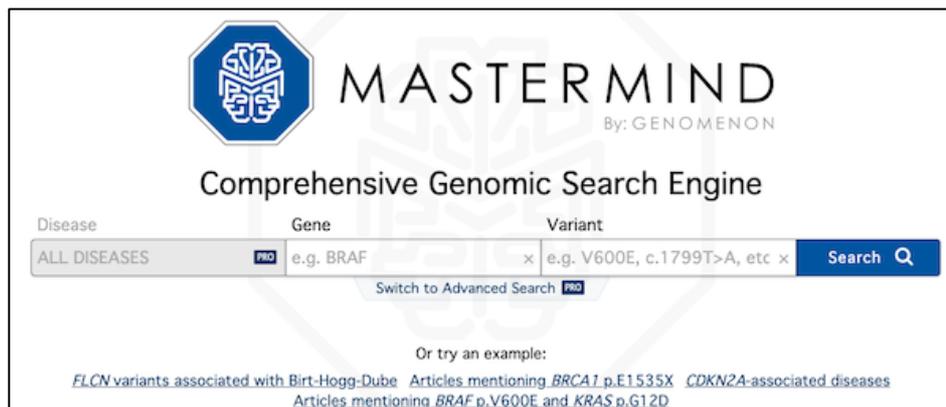


Imagen 6.16 Pantalla principal de Mastermind

La imagen de la portada (Imagen 6.14) presenta el diseño del panel de búsqueda de la interfaz principal del sitio BioGPS. Junto al cuadro de ingreso de términos de búsqueda se muestran ejemplos de las posibles alternativas de búsqueda. Una de ellas es Gene Ontology y sugiere al usuario buscar por el identificador del gen en la base de datos Gene Ontology, como lo muestra el término GO:0006275 en el cuadro de ingreso de términos de búsqueda. Un tercer ejemplo de implementación de este patrón puede ser vista en la interfaz de usuario de ICGC Data Portal.

Filtro Genómico

Alias: ninguno

Tipo: Patrón

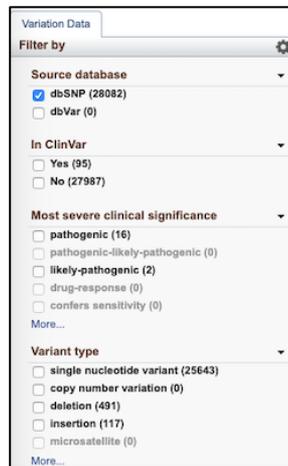


Imagen 6.17 Filtros genómico. Fuente: Sequence Viewer

Problema

Reducir el conjunto de datos genómico aplicando condiciones a las propiedades de los datos.

Contexto

A partir de un conjunto numeroso de variaciones o genes de una o varias muestras, el usuario quiere ocultar los datos que no cumplen con cierto criterio establecido para el estudio genético.

Solución

Presentar al usuario una lista de filtros aplicables al elemento genómico bajo estudio (p. ej., variación genética, gen, muestra). Cualquier propiedad del elemento genómico se puede convertir en filtro. Por ejemplo, tipo de variación, significancia clínica, impacto funcional son propiedades de la variación genética que pueden convertirse en candidatos para filtros. Por lo tanto, para crear un filtro, considerar una propiedad del elemento genómico y expresar los posibles valores por los cuales se puede filtrar a través de cuadros de control como cajas de texto, cuadros de verificación, barras de deslizamiento, etc.

En el caso de que el área de análisis permita colocar componentes visuales, es recomendable permitir al usuario agregar los filtros al área de análisis y colocarlos libremente en cualquier lugar del área. Esto facilita al análisis visual ya que los filtros pueden ser colocados junto a las visualizaciones. Así, el usuario puede ver la relación causa-efecto de aplicar el filtro sin necesidad de regresar al sector de la interfaz de usuario que contiene las opciones de filtro.

Esquema

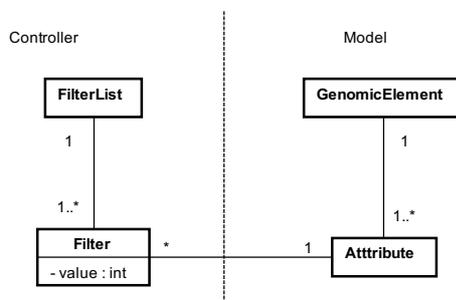


Imagen 6.18 Esquema del patrón Filtro Genómico

Patrones relacionados

Asociación: relacionado a **Gráfico, Secuencia de acciones, Operador de conjuntos.**

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

Imagen 6.19 Filtros de la aplicación GENESIS 2.0

La Imagen 6.19 es una porción de la interfaz de la aplicación GENESIS 2.0. que muestra dos filtros. El primero, “Sequence quality”, es un filtro aplicado a la calidad de las lecturas de la secuencia donde los posibles valores del filtro son profundidad (DP), calidad asignada al genotipo (GQ) y calidad (QUAL) y que son explicados en la especificación del fichero

VCF. El segundo, *Variant Type*, filtra las variaciones genéticas por su clase (p. ej., SNV, inserción, eliminación, etc).

La portada en la Imagen 6.17, es otro ejemplo de implementación del patrón Filtro Genómico en la interfaz de usuario de la aplicación Sequence Viewer.

Selector de Regiones

Alias: ninguno

Tipo: Patrón

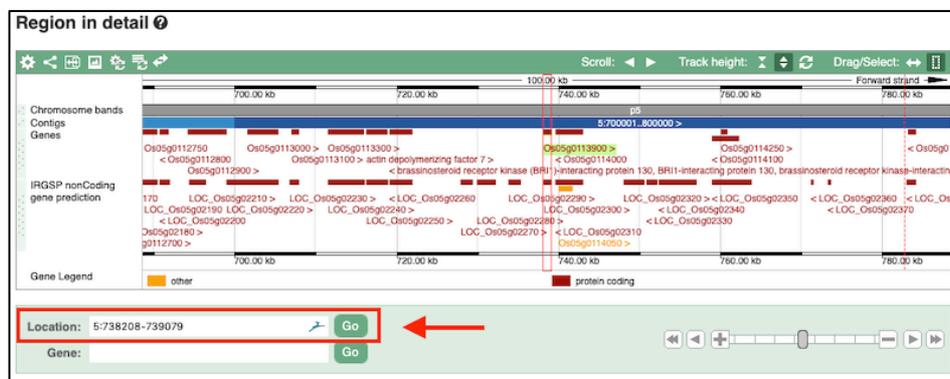


Imagen 6.20 Selección de una región del cromosoma 5. Fuente: Gramene

Problema

Especificar regiones de interés en la secuencia genética.

Contexto

El usuario analiza o explora el genoma enfocándose en regiones concretas de la secuencia genómica. El usuario interesado en este análisis generalmente conoce las posiciones o coordenadas exactas (locus del genoma) que quiere revisar dentro de la secuencia.

Solución

Permitir al usuario especificar el número del cromosoma o cromosomas que desea analizar y la coordenada de inicio (locus inicial) y fin (locus final) de la región de interés. Si se seleccionan varios cromosomas, las coordenadas de inicio y fin se aplican a las secuencias particulares de cada cromosoma.

Esquema

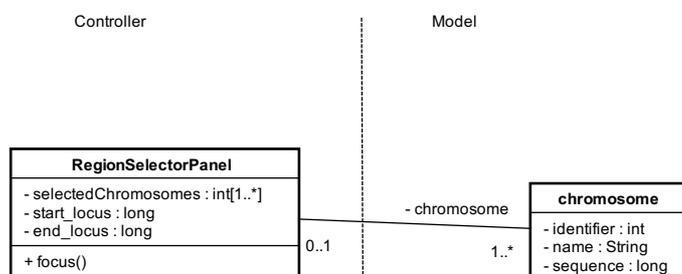


Imagen 6.21 Esquema del patrón Selector de Regiones

Patrones relacionados

- Asociación: relacionado a **Idiograma**.
- Especialización: es un **Filtro genómico**.
- Agregación: ninguno.

Ejemplos

El formulario de la Imagen 6.22 muestra la implementación del patrón en la que se permite al usuario seleccionar varias cromosomas y las posiciones de inicio y fin son aplicadas a todos los cromosomas seleccionados.

Otras implementaciones de este patrón se pueden encontrar en la mayoría de los navegadores genéticos: Genome Maps, UTGB, Genome Data Viewer.

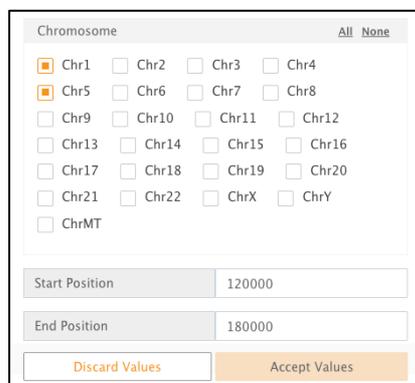


Imagen 6.22 Selector de regiones cromosómicas. Fuente: GENESIS 2.0

Gráfico

- Alias:** ninguno
- Tipo:** Patrón

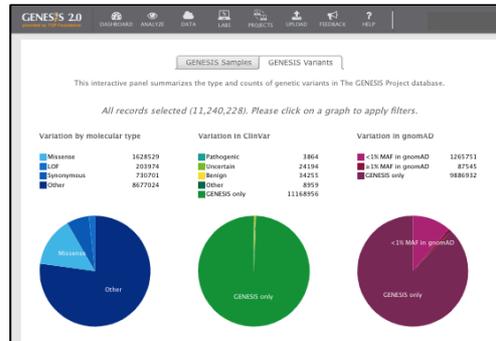


Imagen 6.23 Patrón gráfico. Fuente: GENESIS 2.0.

Problema

Presentar visualmente el comportamiento del conjunto de datos genómico.

Contexto

El usuario quiere tener una representación visual bidimensional del conjunto de datos que le permita comprender el comportamiento de los datos.

Solución

Crear un gráfico bidimensional indicando las dimensiones (largo y ancho), el tipo de gráfico (p. ej., barras, torta, dispersión) y la fuente de datos. La fuente de datos es una “vista de datos” derivada de operaciones de agregación (contar, sumar, etc.) aplicadas sobre las entidades del Modelo Conceptual del Genoma que se quiere representar. La vista de datos debe codificarse en formato bidimensional para ser consumida por el gráfico.

Para mejorar el análisis de los datos, el gráfico creado puede ser utilizado como un filtro visual y dinámico. El gráfico puede ser conectado a otros gráficos de tal manera que al pulsar un sector de datos del gráfico (p. ej., una barra, en el gráfico de barras), los otros gráficos conectados actualicen automáticamente su visualización de acuerdo con el filtro pulsado.

Esquema

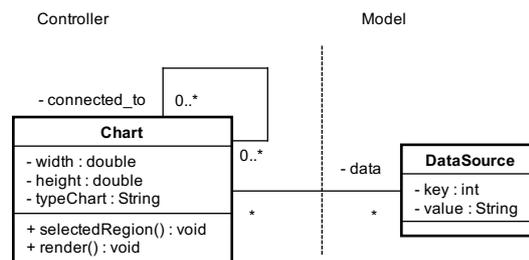


Imagen 6.24 Esquema del patrón Gráfico

El patrón se implementa instanciando la clase “Chart” que representa un componente visual de la interfaz de usuario y que al ser representado muestra un gráfico de dos

dimensiones. La instanciación del patrón supone la especificación del ancho (“width”), alto (“height”), y el tipo de gráfico (“typeChart”) que puede ser, por ejemplo, gráfico de barras, gráfico de torta, etc. Los datos para representar el gráfico son provistos por la clase “DataSource”. Esta clase no es parte del MCGH, sin embargo, ha sido incluida en el esquema para representar a la “vista de datos” resultante de realizar operaciones de agregación sobre los datos originales del MCGH y que está expresada en formato clave-valor.

La relación recursiva en la clase “Chart” indica las conexiones con otros gráficos permitiendo formar un ambiente gráfico total o parcialmente conectado. La operación “selectedRegion” es ejecutada cuando un sector de datos del gráfico se selecciona enviando mensajes a los otros gráficos que estén conectados y que están registrados en el atributo “connected_to”.

Patrones relacionados

Asociación: relacionado a **Filtro Genómico, Idiograma**

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

La Imagen 6.25 presenta cuatro gráficos estadísticos acerca de la distribución de las mutaciones (variaciones genéticas) disponibles en el repositorio de datos del sitio ICGC Data Portal. El primer gráfico representa mediante barras la frecuencia de variaciones en función del tipo de consecuencia que causan. Los tres gráficos son gráficos de sectores que presentan el porcentaje de mutaciones respecto a: la plataforma utilizada para su secuenciación (p. ej., el sector de mayor área corresponde a mutaciones secuenciadas con la plataforma Illumina HiSeq), el estado de verificación (p. ej., el sector de mayor área corresponde a mutaciones no probadas) y el tipo de variación (p. ej., el sector de mayor área corresponde a variaciones del tipo “sustitución”).

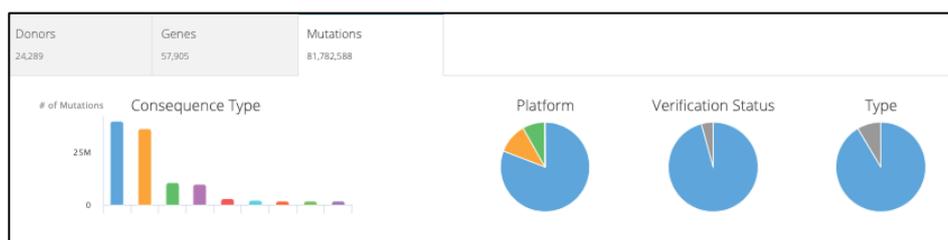


Imagen 6.25 Gráficos de distribución de mutaciones. Fuente: ICGC Data Portal.

Otros ejemplos de implementación del patrón se encuentran en interfaces de aplicaciones como Galaxy o Genesis (como lo muestra la imagen de portada de este patrón).

Idiograma

Alias: Ninguno

Tipo: Patrón

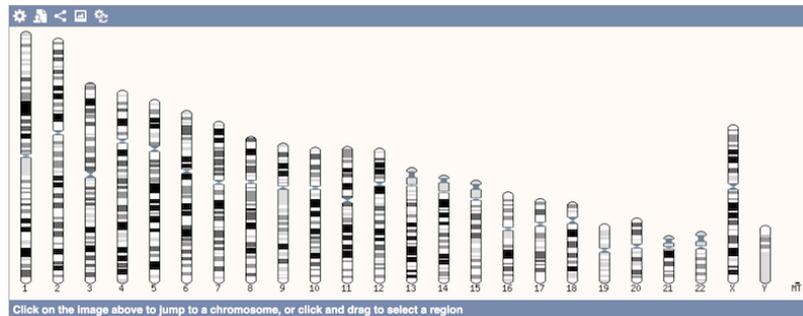


Imagen 6.26 Idiograma de la especie humana (GRCh38.p). Fuente: e!Ensembl

Problema

¿Cómo visualizar el conjunto de cromosomas de una especie?

Contexto

El investigador quiere visualizar el *cariotipo* (la organización del conjunto cromosómico de un organismo) para identificar anomalías cromosómicas causantes de malformaciones o de enfermedades. Este tipo de análisis se realiza especialmente en estudios citogenéticos.

Solución

Mostrar un *Idiograma* (conocido como *ideograma* o *cariograma*). Un Idiograma es una representación gráfica del cariotipo que expresa la organización del conjunto de cromosomas de una especie (Oliva et al., 2013).

Para crear un Idiograma:

- Agregar un componente visual (p. ej., panel) que sirva como contenedor de los cromosomas. Definir el *tamaño* (ancho y largo) y *orientación* del contenedor. El tamaño del contenedor define la *escala* de dibujo. La orientación define la posición del Idiograma, pudiendo ser *horizontal* o *vertical*.
- En el contenedor, dibujar los cromosomas ordenados ascendentemente por su número, dejando los cromosomas XY, si se incluyen, al final del conjunto de cromosomas. Respetar la escala de dibujo de tal manera que la diferencia de tamaño entre los cromosomas sea visible.
- Centrómero. – (opcional) Si los cromosomas definen el centrómero, brazo largo (q) y brazo corto (p), entonces dibujar los cromosomas con el brazo corto hacia arriba y el brazo largo hacia abajo. Si la orientación en el paso 1 es vertical, entonces el brazo largo debe ser dibujado hacia abajo. Si la orientación es horizontal, entonces el brazo largo debe ser dibujado hacia la derecha.
- Dibujar el *patrón de bandas* en cada cromosoma (ver “Patrón de Bandas”, más abajo).

Patrón de bandas:

Las bandas cromosómicas ayudan a identificar a cada cromosoma y localizar los *loci* genéticos (lugares dentro de un cromosoma) asociados a patologías relevantes. Las bandas son el resultado de la observación por microscopio al teñir el cromosoma con productos químicos y su visualización está definida por sectores representados en escala de grises. La

nomenclatura de las bandas está dada por el *International System for Chromosome Nomenclature* (ISCN) y distingue jerárquicamente la delimitación por cromosoma, brazo, región, banda y sub-banda. Las regiones, bandas y sub-bandas se enumeran ascendentemente empezando por 1, siendo 1 la banda más cercana al *centrómero* (el punto centro del cromosoma).

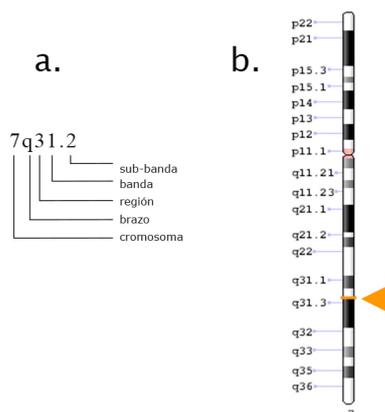


Imagen 6.27 a) Ejemplo de nomenclatura ISCN y b) representación de la posición 7q31.2 obtenida con Genome Decoration Page

Por ejemplo, la Imagen 6.27a muestra la nomenclatura 7q31.2 que indica que se hace referencia a un sector dentro del cromosoma siete, brazo largo (*q*), región tres, banda uno, y separado por el punto, la sub-banda 2. La Imagen 6.27b, obtenida con Genome Decoration Page (www.ncbi.nlm.nih.gov/genome/tools/gdp), representa gráficamente la posición citogenética 7q31.2 que corresponde a la región de pares de bases desde 117,479,963 a 117,668,665 en el cromosoma 7 (utilizando la versión actualizada 109.20190607 de la especie *Homo sapiens*, GRCh38.p13).

Los datos del patrón de bandas pueden ser obtenidos del NCBI a través de ficheros disponibles en su cuenta de FTP (<ftp://ftp.ncbi.nlm.nih.gov/pub/gdp/>) y que su nombre inicia con el formato `ideogram_<tax_id>`, donde “`tax_id`” es el identificador de la taxonomía del organismo que en el caso del *Homo Sapiens* es 9606.a

Esquema

El patrón Idiograma se crea instanciando la clase “`IIdiogramPanel`” (Imagen 6.28) la cual representa el panel gráfico en la interfaz de usuario donde se muestran los cromosomas dibujados. El idiograma se compone de uno o varios cromosomas dibujados (clase “`ChromosomeGraph`”) y cada cromosoma puede presentar su patrón de bandas (clase “`Band`”). Si el idiograma se compone de varios cromosomas, su visualización será similar a la Imagen 6.26. Si el idiograma se compone de un solo cromosoma, su visualización será similar al de la Imagen 6.29b.

La clase “`ChromosomeGraph`” está relacionada a la clase “`chromosome`” desde donde obtiene los datos para graficar los cromosomas. Cada cromosoma se grafica considerando los atributos de la clase: ancho (“`width`”), alto (“`height`”) y etiqueta (“`label`”). El cromosoma permite interacciones representadas por operaciones como seleccionar (“`select`”), desplazar

(“pan”), acercar y alejar (zoomIn y zoomOut, respectivamente). Además, el cromosoma puede resaltar posiciones de ciertos elementos del genoma como son, por ejemplo, variaciones genéticas o genes. A estas posiciones se las ha denominado “anotaciones” y provienen de la clase “AnnotationElement”. La clase “AnnotationElement” no es propia del Modelo Conceptual del Genoma, sino que ha sido introducida en este diagrama para generalizar el tipo de anotaciones que pueden representarse sobre la secuencia del cromosoma.

En el panel del Idiograma (clase “IdiogramPanel”), los Idiogramas se presentan ordenados por el número de cromosoma (“identifier”).

La clase “Band” tiene una relación recursiva para representar a la sección, banda y sub-banda. Estos datos son suministrados por la clase “cytoband” del modelo conceptual.

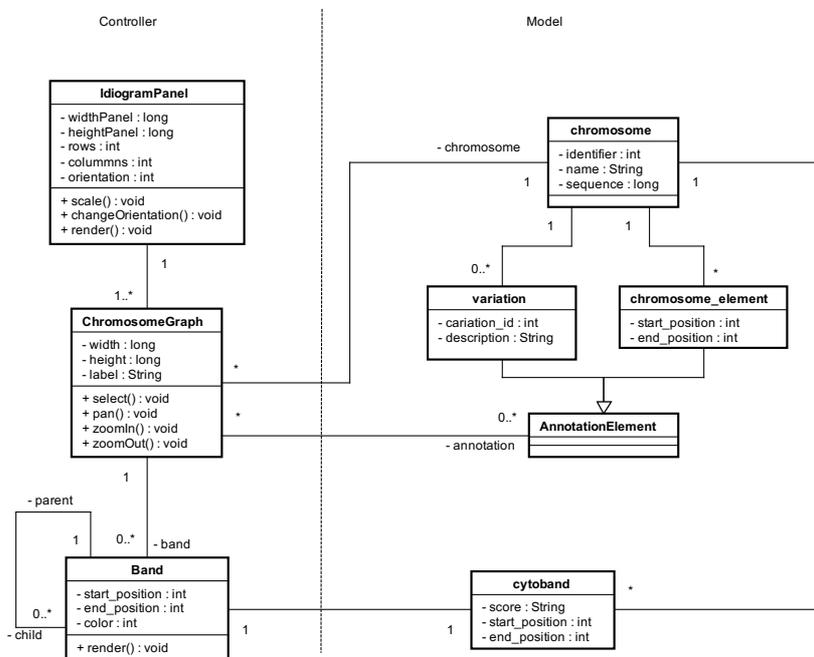


Imagen 6.28 Esquema para el patrón Idiograma

Patrones relacionados

Asociación: relacionado a **Selector de regiones**, **Gráfico**

Especialización: ninguno.

Agregación: puede tener un **Gráfico**.

Ejemplos

La Imagen 6.29 muestra una porción de la interfaz de usuario de Genome Maps, donde se muestra dos implementaciones del patrón Idiograma. La primera implementación (Imagen 6.29a) muestra el Idiograma dispuesto horizontalmente y en una sola fila, con 23

pares de cromosomas de la especie humana: 22 pares de cromosomas autosomas y un par de cromosomas gonosomas (X y o Y, siendo XX para el sexo femenino, y XY para el sexo masculino). La segunda (Imagen 6.29b) muestra un solo cromosoma dispuesto horizontalmente con bandas etiquetadas con la nomenclatura de las bandas. Las dos implementaciones están conectadas, cuando un cromosoma es seleccionado en la primera implementación, la segunda implementación muestra el cromosoma seleccionado.

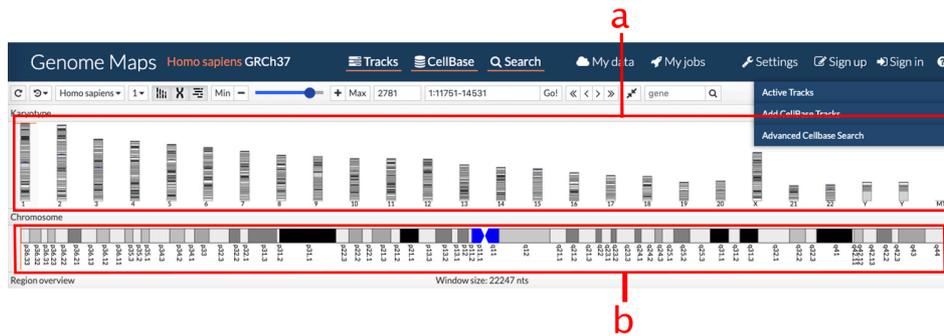


Imagen 6.29 a) Idiograma total del genoma e b) idiograma particular de un cromosoma.
Fuente: Genome Maps

Otras implementaciones del patrón se encuentran en las interfaces de las aplicaciones: G-compass, Genome Data Viewer.

Pista Comparativa

Alias: Ninguno
Tipo: Patrón

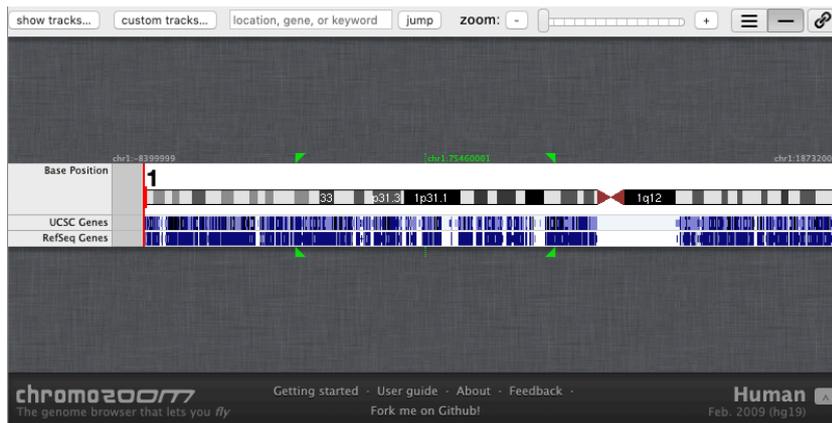


Imagen 6.30 Pistas con anotaciones genómicas (UCSC Genes, RefSeq Genes). Fuente: ChromoZoom

Problema

Comparar visualmente varias secuencias genómicas respecto a una posición (locus) específica y común para todas las secuencias

Contexto

El estudio genómico incluye comparar muestras de variaciones genéticas (p. ej., muestras de integrantes de una familia) con anotaciones genómicas provenientes de diferentes fuentes (p. ej., dbSNP, ENCODE, UCSC, ClinVar, KEGG), para interpretar el significado de las variaciones de las muestras. La comparación a diferentes niveles de detalle (desde bases genéticas hasta genes o cromosomas) supone alta carga cognitiva al usuario.

Solución

Presentar las secuencias genómicas a través de un sistema de pistas que permita la comparación visual entre ellas. Una pista es una representación unidimensional (es decir, una sola dimensión: la posición) del contenido de una muestra o secuencia genética. Usar el patrón Cargar Muestra para configurar la secuencia de datos que se quiere visualizar en la pista. La presentación y disposición de las pistas debe cumplir: i) adyacencia (las pistas se sitúan una junto a la otra, ya sea en una orientación vertical u horizontal); ii) alineación (todas las pistas deben presentar las secuencias alineadas respecto a una escala de posiciones de referencia); iii) escala (los controles de alejamiento o acercamiento modifican la escala de visualización definiendo un rango visible de bases de ADN, por lo tanto, todas las pistas deben presentar su contenido considerando el número de bases visible).

La orientación del sistema de pistas es horizontal, ocupando en gran medida el espacio de la interfaz de usuario. Opcionalmente, el orden de las pistas puede ser modificado por el usuario.

El sistema de pistas sincroniza la navegación unidimensional por todas las pistas. Es decir, si el usuario se desplaza a una posición específica de una secuencia contenida en una pista, todo el conjunto de pistas se desplaza a la misma posición y al mismo tiempo. Los controles del nivel de acercamiento o alejamiento permiten explorar las secuencias desde una perspectiva global o detallada.

Esta solución es aplicable cuando se dispone de un espacio amplio (vertical u horizontal) en la interfaz de usuario.

Esquema

El patrón se crea instanciando la clase “ComparativeTracks” (Imagen 6.31) la cual contiene un idiograma (es decir, el patrón Idiograma representado por la clase “Idiograma”) y varias pistas (clase “Track”). El patrón Idiograma muestra un único cromosoma (generalmente dispuesto horizontalmente) que sirve como marco de referencia de navegación ya que establece el rango de coordenadas en el que las pistas deben mostrar sus datos. Por ejemplo, si en el idiograma se ejecuta una acción de “acercamiento”, todas las pistas automáticamente cambian su estado de visualización a las nuevas coordenadas establecidas por el acercamiento en el idiograma.

Las pistas, ubicadas generalmente bajo el patrón Idiograma, muestran datos de una fuente de datos externa (clase “ExternalDataSource”) como por ejemplo dbSNP. Cada pista muestra el nombre de la fuente de datos (“name”) y los datos entre una posición inicial (start_locus) y final (end_locus). La clase ComparativeTracks controla las coordenadas de

visualización y configura dichas coordenadas a todas las pistas, de tal manera que todas las pistas sincronicen su visualización al mismo rango de coordenadas.

Las interacciones de navegación se realizan desde el idiograma o desde las acciones propias de la clase `ComparativeTracks`. Por ejemplo, si una acción de acercamiento es realizado mediante el idiograma (ver patrón Idiograma), las acciones son detectadas por la clase `ComparativeTracks` y esta ejecuta su operación “`zoomIn()`” que redimensiona la escala de todas las pistas e indica los nuevos valores de inicio (`start_locus`) y fin (`end_locus`) para todas las pistas.

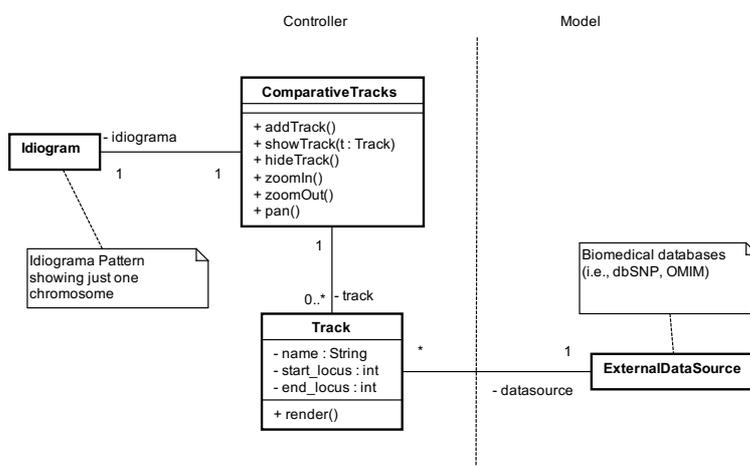


Imagen 6.31 Esquema del patrón Pista Comparativa

Patrones relacionados

Asociación: ninguno.

Especialización: ninguno.

Agregación: tiene un **Cargar Datos**.

Ejemplos

La Imagen 6.32 muestra una parte de la interfaz del navegador genómico GBrowse con varias pistas horizontales alineadas verticalmente respecto a la posición indicada por la escala entre las 660 kilobases y las 760 kilobases (primera pista).

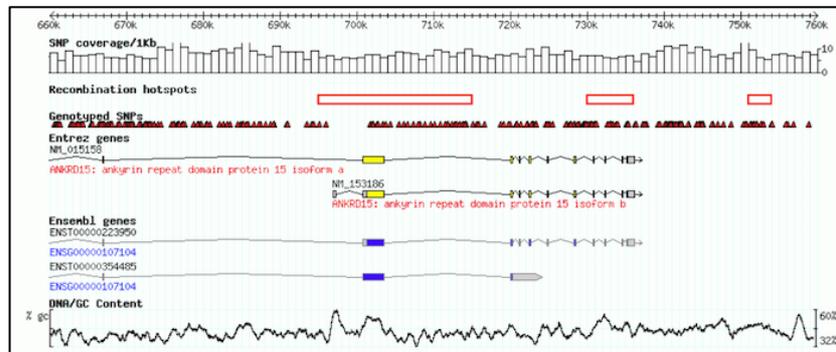


Imagen 6.32 Pistas de secuencias de datos de GBrowse

Este patrón está presente en la mayoría de browser genéticos como por ejemplo UCSC Genome Browser, Vega, Pileup, Biodalliance, IGV LooqSeq, JBrowse, Genome Maps, ChromoZoom.

Columnas Ocultas

Alias: Ninguno
 Tipo: Proto-patrón

The screenshot shows the 'Genes and regulation' configuration panel in Gramene. It includes a 'Show/hide columns' section with a table of columns and their visibility status. The table is as follows:

Column	Visible
Gene	Yes
Transcript (strand)	Yes
Allele (Tr. allele)	Yes
Consequence Type	Yes
Position in transcript	Yes
Position in CDS	Yes
Position in protein	Yes
AA	No
Codons	No
SIFT	No
Detail	No

The table also displays data for two variants: a missense variant (G) at position 1252 and a 5 prime UTR variant (C) at position 1131.

Imagen 6.33 Configuración de columnas visibles/ocultas. Fuente: Gramene

Problema

Visualizar el número de columnas relevantes para el análisis.

Contexto

La presentación tabular de los datos genéticos requiere multitud de columnas debido a la cantidad de características de cada elemento genético (p.ej., variación, gen, etc.) y a las relaciones entre ellos. La cantidad de columnas sobrecarga el espacio de análisis y convierte a la navegación y comparación de información en tareas complejas. El usuario generalmente está interesado en no más de siete columnas para llevar a cabo su análisis.

Solución

Agregar a la tabla de datos la opción para mostrar la lista de columnas que integran la tabla. A partir de la lista, el usuario configura/indica cual columna debe ser visible y cual oculta. Una vez guardada la configuración, la tabla debe refrescar su estado ocultando las columnas que se configuraron como ocultas.

Esquema

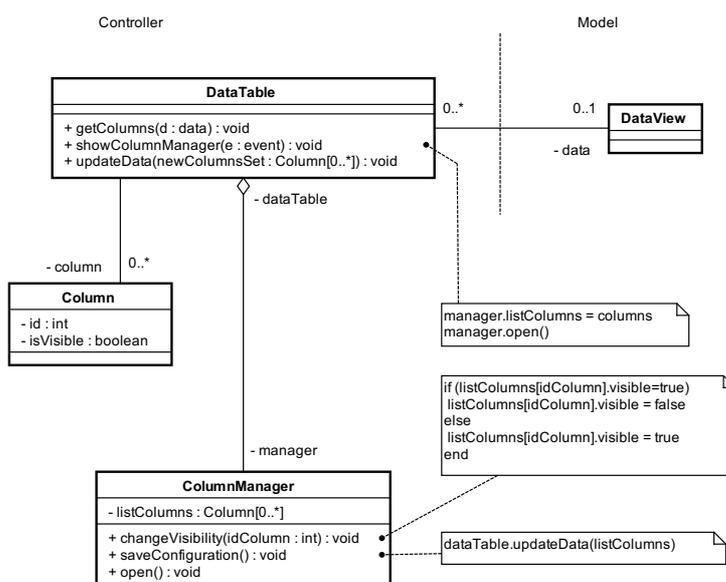


Imagen 6.34 Esquema del patrón Columnas Ocultas

El patrón se crea instanciando un *administrador de columnas* (clase “ColumnManager”). La clase “DataTable” representa a una tabla (un componente visual de la interfaz de usuario que contiene filas y columnas para mostrar los datos) la cual contendrá el administrador de columnas para ocultar/mostrar columnas. La tabla es poblada con datos provenientes de la clase “DataView”; una vista creada con los datos de las entidades del Modelo Conceptual del Genoma que se quiere listar en la tabla. El conjunto de columnas que forman la tabla es obtenido a partir de los datos de la vista y cada columna es una instancia de la clase “Column”, donde los atributos “id” y “isVisible” representan el identificador de la columna y su estado binario de visibilidad (es decir, visible, no visible).

El administrador de columnas es un panel visual que muestra las columnas de la tabla tanto las visibles como las no visibles y permite al usuario cambiar el estado visible de cada columna. El panel se muestra una vez que el usuario solicita su apertura (evento “showColumnManager”) en el componente tabla. En el panel visual, el usuario cambia el estado de visualización de las columnas que desea ocultar o visualizar y cada cambio es capturado por el administrador (evento “changeVisibility”). El conjunto de cambios se convierte en la nueva configuración de columnas. Cuando el usuario guarda esta

configuración (evento “saveConfiguration”), el administrador solicita a la tabla actualizar sus datos (evento “updateData”) con la nueva configuración de columnas (“listColumns”).

Patrones relacionados

Asociación: relacionado a **Navegación continua, Navegador genómico, Buscar.**

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

Ejemplos de implementación de este patrón se encuentran en las interfaces de las aplicaciones: GENESIS 2.0.

Círculo cromosómico

Alias: Radial Table (Tidwell, 2010)

Tipo: Patrón

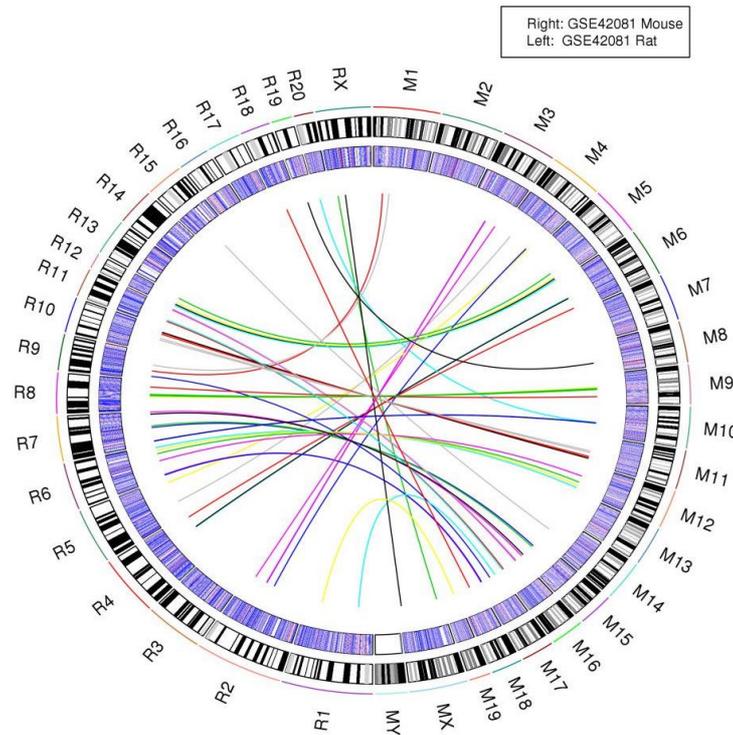


Imagen 6.35 Idiogramas del ratón y la rata, mapas de calor y líneas de enlace. Fuente: RCircos (Zhang, 2013)

Problema

Visualizar relaciones entre elementos o posiciones de múltiples secuencias genómicas.

Contexto

El usuario quiere comparar dos o más secuencias genómicas (común en estudios de genómica comparativa para estudiar la conservación o *sintenia*), identificando relaciones de elementos o posiciones entre ellas. Las secuencias son cantidades extensas de datos y la representación tabular no facilita la identificación de relaciones entre los datos.

Solución

Representar la secuencia genómica en una distribución circular de tal manera que se pueda indicar gráficamente los enlaces entre los elementos relacionados de la secuencia, como lo muestra la Imagen 6.36. La secuencia genómica está representada por idiogramas (1) dispuestos en una composición circular. Los elementos cromosómicos (2) que se desea visualizar dentro del idiograma, son indicados respetando la posición dentro de la secuencia del cromosoma. Las relaciones entre elementos cromosómicos (2), o sectores de secuencia (3) están indicados a través de enlaces (4)(5) que conectan las posiciones de los idiogramas asociados. En el caso de existir categorías de enlaces, pintar los enlaces con diferentes colores para distinguir entre las diferentes categorías (p.ej., el color de los enlaces que conectan genes que participan en una determinada ruta metabólica difiere del color de los enlaces que conectan grupos de regiones similares en la secuencia). Para contextualizar la información genómica y mejorar el análisis de la secuencia, agregar anotaciones sobre (o por debajo) de los idiogramas a través del patrón “Pista Comparativa” (6).

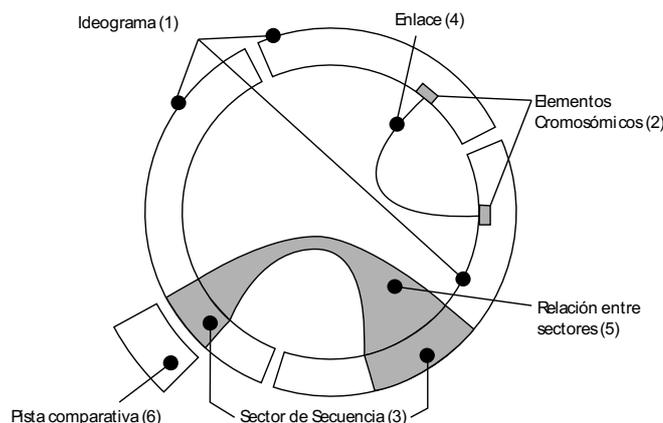


Imagen 6.36 Estructura del Círculo Cromosómico

Esquema

El patrón Círculo Cromosómico se crea instanciando la clase “ChromosomicCircle” (Imagen 6.37). La solución de este patrón incluye las soluciones de los patrones Idiograma (grupo “Patron Idiograma” y Pistas Comparativas (grupo “Patrón Pista Comparativa”). El “patrón Idiograma” muestra los cromosomas de una o varias especies (clase “species”, del modelo conceptual del genoma) en una representación circular (*IdiogramPanel.orientation = “circle”*). Los enlaces entre los elementos cromosómicos (clase “Connector”) se representan con líneas de color y conectan una posición origen de un cromosoma con una posición destino en otro cromosoma. El color de los enlaces varía según las categorías de relaciones

existentes (p. ej., un color para enlaces entre genes participantes en una cierta ruta metabólica o *pathway* y otro color para enlaces entre variaciones causantes de una cierta enfermedad),

Las anotaciones, provenientes de fuentes locales (clase “AnnotationElement”) o externas (clase “ExternalDataSource”), que contextualizan la información en los cromosomas del Idiograma, se muestra a través del patrón Pistas Comparativas (clase “ComparativeTracks”).

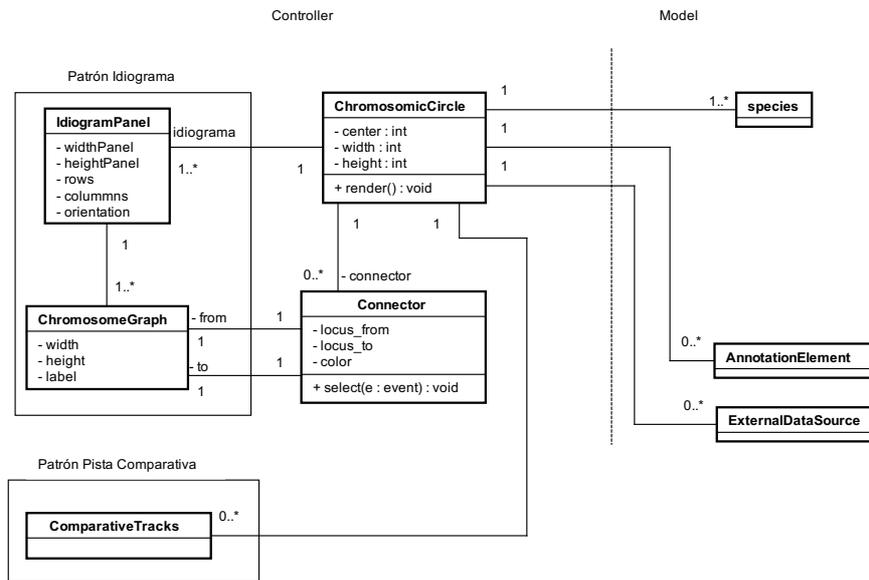


Imagen 6.37 Esquema para el patrón Círculo Cromosómico

Patrones relacionados

Asociación: ninguno.

Especialización: es un **Navegador genómico**.

Agregación: tiene un **Idiograma** y **Pista comparativa**.

Ejemplos

Este patrón se encuentra implementado en las interfaces de usuario de las aplicaciones: Circos, J-Circos, ChromoWheel. Librerías que implementan el patrón: RCircos (como lo muestra la portada Imagen 6.35), BioCircos.js.

Navegador Genómico

Alias: Ninguno

Tipo: Patrón



Imagen 6.38 Interfaz de Genome Data Viewer

Problema

Visualizar y explorar elementos cromosómicos en su contexto genómico.

Contexto

El usuario necesita navegar y explorar los elementos de la secuencia genómica con respecto a un genoma de referencia. Los usuarios necesitan enfocarse en una determinada región del genoma y analizarla en su contexto genómico (información anotada asociada a dicha región disponible localmente o de forma remota).

Solución

Proveer un ambiente exploratorio compuesto de dos vistas:

- Vista general. - Contiene controles capaces de navegar por el genoma completo y el patrón Cargar Datos para seleccionar los datos a ser visualizados; el patrón Idiograma puede incluirse para facilitar la orientación a lo largo de la secuencia genética. Si se incluye, el patrón Idiograma facilita tanto la visualización y navegación por los cromosomas de la especie como la navegación por la secuencia de un cromosoma específico.
- Vista detallada. – Debe incluir el patrón Pista Comparativa configurado de manera horizontal y ocupando el mayor espacio de la interfaz de usuario. Permitir al usuario configurar el contenido de cada pista utilizando el patrón Cargar Datos. Una pista debe ser dedicada a los datos genómicos que se quiere explorar. Otras pistas pueden configurarse con anotaciones provenientes de diferentes fuentes (p. ej., Gene, 1000G, dbSNP, OMIM, ClinVar), que agregan información de contexto como información de genes, población, enfermedad, estadísticas que facilitan la interpretación y exploración de la primera pista. El patrón Pista Comparativa incorpora controles de desplazamiento, alejamiento y acercamiento para explorar y visualizar las pistas a más bajo detalle.

Esquema

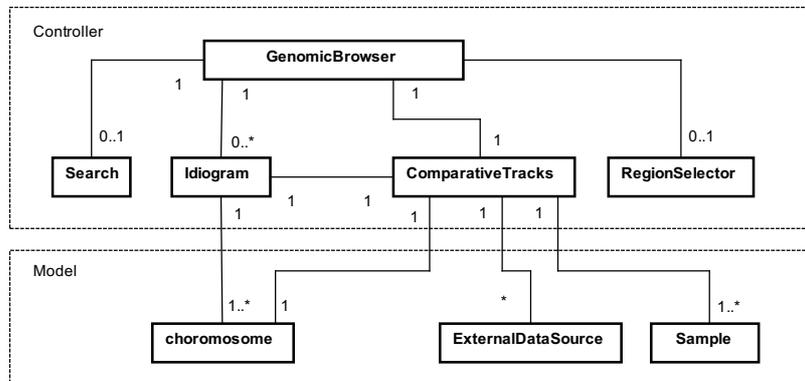


Imagen 6.39 Esquema del patrón Navegador Genómico

El patrón se crea instanciando la clase “GenomicBrowser”. La instancia del patrón inicializa sus componentes, es decir, los patrones: Idiograma (clase “Idiogram”), Buscar (clase “Search”), Pista Comparativa (clase “ComparativeTracks”), Selector de Regiones (clase “RegionSelector”). Las interacciones que ofrece el patrón Navegador Genómico están dadas por las interacciones de cada uno de sus componentes.

En este esquema, las clases correspondientes a los patrones que componen el navegador genómico son representaciones generalizadas de cada uno de los patrones. La estructura detallada de cada uno de ellos puede ser consultada en la documentación de cada patrón.

Patrones relacionados

Asociación: relacionado a **Columnas ocultas**.

Especialización: ninguno.

Agregación: tiene un **Buscar, Selector de regiones, Idiograma, Pista comparativa**.

Ejemplos

En la interfaz de usuario de Genome Data Viewer mostrada en la portada (Imagen 6.38) se puede apreciar algunas de los patrones que conforman el patrón Navegador Genómico: a) Idiograma b) Pista Comparativa c) Selector de Regiones y Buscar y d) acceso al patrón Cargar Datos como parte del patrón Pista Comparativa. Otras implementaciones de este patrón son Navegadores Genéticos y Visores genéticos implementados en plataforma web y escritorio.

Basados en web: Vega, Sequence Viewer, GBrowse, G-compass, JBrowse, LookSeq, UTGB, Ensembl Gente Browser, UCSC Xena, IGV. Basados en escritorio: Integrative Genomics Viewer, SnapGene Viewer

Navegación continua

Alias: Continuous Scrolling, Infinite Scrolling (UI-Patterns (Toxboe, 2007))

Tipo: Patrón

Big Data Grid							EXPORT TO ...	DELETE	DONE
	Id	Name	Verified	Ratings		Date of Birth			
				Avg	All				
Support (67)									
▼	496	372253	Dave White	<input type="checkbox"/>	5.10		13-07-1986		V
▼	497	134435	Dave Robinson	<input type="checkbox"/>	5.80		16-05-1962		V
▼	498	147531	David Davis	<input type="checkbox"/>	4.70		10-04-1987		V
▼	499	471615	Jay Davis	<input checked="" type="checkbox"/>	5.90		17-06-1979		V
▼	500	765725	Jamie Elias	<input type="checkbox"/>	5.10		20-07-1989		V
					5.18				
					5.05				

Imagen 6.40 Solución de navegación continua. Fuente: Librería Ext JS (Sencha, 2016)

Problema

¿Cómo mejorar la fluidez en la navegación y exploración de grandes volúmenes de datos?

Contexto

El usuario quiere navegar o explorar los datos de manera fluida (sin interrupciones). Sin embargo, los componentes de interfaz de usuario que visualizan grandes volúmenes de datos y permiten navegar por ellos (p. ej., tablas, gráficos, mapas) presentan problemas de rendimiento que afectan a la fluidez de la navegación. La navegación por segmentos (p. ej., paginación), donde el usuario interrumpe la navegación para especificar el segmento de datos al que quiere navegar, es una estrategia para navegar por grandes cantidades de datos. Sin embargo, esta estrategia sacrifica la fluidez de navegación.

Solución

Cargar automáticamente el siguiente conjunto o segmento de contenido cuando el usuario llegue al final del segmento actual. Para esto:

Capturar el o los eventos producidos por el usuario sobre los componentes visuales de la interfaz y que suponen recuperación de datos (p. ej., alejar, acercar, desplazar, etc.). Los datos recuperados en cada evento deben ser almacenados por segmentos en una “memoria intermedia” y entregados progresivamente al componente visual. La memoria intermedia almacena los datos de manera temporal y por segmentos. Cada vez que la memoria llena su capacidad con una porción de los datos recuperados, envía al componente visual el segmento de datos y vuelve a llenar su almacenamiento con otra porción de datos del conjunto total de datos. De esta manera, la entrega continua de datos realizada por la memoria intermedia evita que la visualización o navegación de datos en el componente visual se detenga hasta recuperar el conjunto de datos completo.

Este patrón agrega fluidez a la navegación y exploración de datos, presentando los datos a medida que el usuario se desplaza por el conjunto de datos. El usuario no indica de manera

explicita su intención de visitar un determinado segmento del conjunto de datos. La intención es implícita y es capturada automáticamente.

Esquema

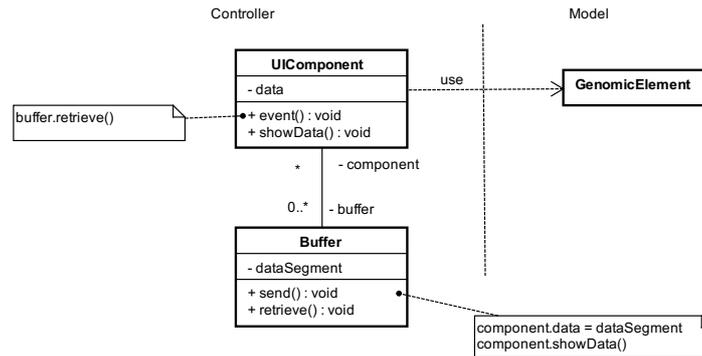


Imagen 6.41 Esquema del patrón Entrega Continua

Este patrón se crea instanciando la clase “Buffer” que representa la memoria intermedia que entrega de manera continua los datos al componente visual (clase “UIComponent”).

El evento “event()” del componente es ejecutado cada vez que el usuario realiza una interacción que implica la recuperación de datos. La memoria recupera una porción de datos del modelo (método “retrieve()”) que almacena internamente (atributo “dataSegment”). Una vez que la memoria llena su espacio de almacenamiento, los datos son enviados (método “send()”) al componente, el cual muestra el segmento de datos (método “showData()”).

Patrones relacionados

Asociación: relacionado a **Columnas ocultas**.

Especialización: ninguno.

Agregación: ninguno.

Ejemplos

La imagen de la portada (Imagen 6.40) es un ejemplo de implementación de la librería *Sencha Ext Js* (<https://www.sencha.com/products/extjs/>), donde la tabla datos con 500 registros y sin control de paginación puede ser navegada por el usuario libremente con la rueda de desplazamiento del ratón. Los datos se cargan conforme el usuario se desplaza por la tabla. A diferencia del control de paginación común empleado para navegar por las tablas de datos y donde el usuario hace explícita su intención de navegar a un segmento (página) de contenido, la implementación del patrón permite al usuario navegar libremente por una tabla de datos completa sin experimentar interrupciones por los cambios de página. Otros ejemplos de este patrón pueden ser revisados en los ejemplos indicados en el patrón Continuous Scrolling (<http://ui-patterns.com/patterns/ContinuousScrolling>).

6.2.7 Interrelaciones entre patrones genómicos

Aunque las relaciones definidas en el modelo conceptual del catálogo GenomIU (Sección 6.2.1) se aplican a todos los patrones dentro del catálogo GenomIU, en esta sección resumimos las relaciones entre los patrones genómicos ya que éstos son el principal aporte de esta Tesis.

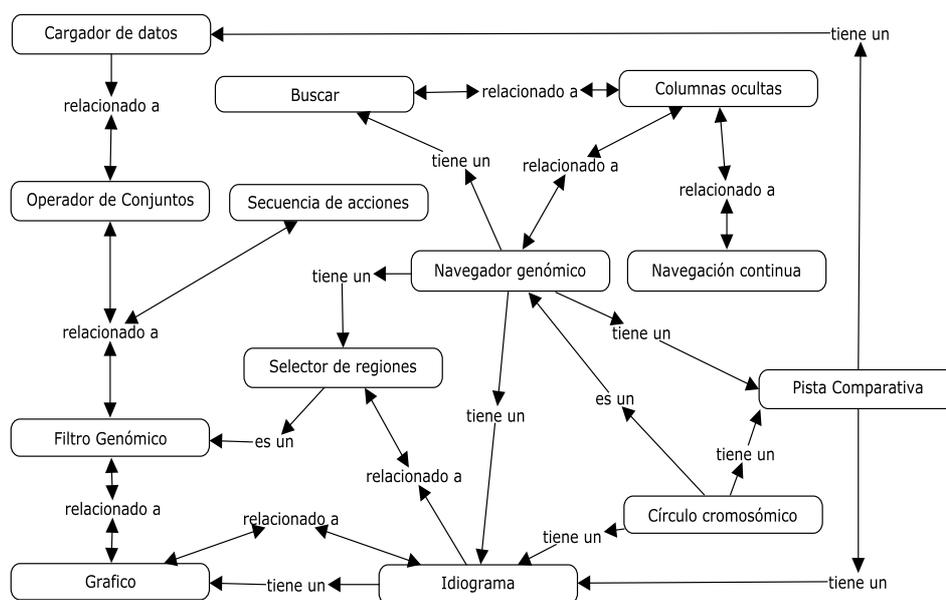


Imagen 6.42 Relaciones entre patrones genómicos

La Imagen 6.42 presenta las relaciones internas entre los patrones genómicos que hemos encontrado en las interfaces de usuario de aplicaciones genómicas estudiadas, y a continuación, describimos ejemplos de los tipos de relaciones:

Asociación. - El patrón Filtro Genómico **está relacionado** con el patrón Gráfico ya que, en un ambiente de exploración de datos, el patrón Filtro Genómico puede ser usado para aplicar varias condiciones de filtro a los datos y el resultado (los datos que cumplen con las condiciones) ser visualizado de manera consolidada a través del patrón Gráfico. Un ejemplo concreto de esta situación es la interfaz de ICGC Data Portal (Imagen 6.43) donde se muestra el patrón Filtro Genómico instanciado a través de tres filtros: *Consequence Type*, *Type*, *Function impact*. Al seleccionar uno o varios valores del filtro, los gráficos, que son instancias del patrón Gráfico, refrescan su contenido para mostrar las variaciones genéticas que han cumplido con las condiciones seleccionadas.



Imagen 6.43 Ejemplo de asociación entre patrones (tomado y adaptado de ICGC Data Portal)

Especialización. - El Círculo Cromosómico es un Navegador Genómico. El Círculo Cromosómico es básicamente un Navegador Genómico especializado porque además de contener las características del Navegador Genómico (p. ej., Pista comparativa, Idiograma), contiene una característica especial que no tiene el Navegador Genómico: *las relaciones visuales entre los elementos*. Por ejemplo, la Imagen 6.35 de la portada del patrón Círculo Cromosómico permite visualizar las relaciones entre los elementos de las secuencias del ratón y la rata, y para esto, el patrón Círculo Cromosómico incluye los patrones Pista Comparativa e Idiograma (propios del Navegador Genómico) y los presenta en forma circular con el fin de trazar las relaciones entre las dos secuencias.

Agregación. - El patrón Navegador Genómico provee un ambiente donde el usuario puede explorar el genoma enfocándose en zonas particulares de la secuencia genómica y analizar uno o varios elementos genómicos contextualizados con anotaciones provenientes de otras fuentes de datos. Para cumplir con todo esto, el patrón Navegador Genómico dentro de su solución **tiene un patrón**: Pista Comparativa, Cargar Datos, Buscar, Selector de regiones y un Idiograma. El navegador genómico Genome Maps (Imagen 6.44) es un ejemplo de esto donde el patrón Idiograma visualiza la secuencia genómica del cromosoma 1 con su patrón de bandas; el patrón Selector de regiones centra la visualización en un rango de interés, específicamente entre las coordenadas 11870 y 14412 del cromosoma 1; el patrón Pista Comparativa cubre el análisis de elementos en contexto genómico proveyendo información de la base de datos “Gene”, “Sequence” y “SNP” dentro del rango de coordenadas indicando por el patrón Selector de regiones; el patrón Cargar Datos cubre el comportamiento de cargar información de las fuentes que se muestran en el patrón Pista Comparativa; el patrón Buscar permite buscar e identificar en la visualización diferentes características sobre la pistas del patrón Pista Comparativa.

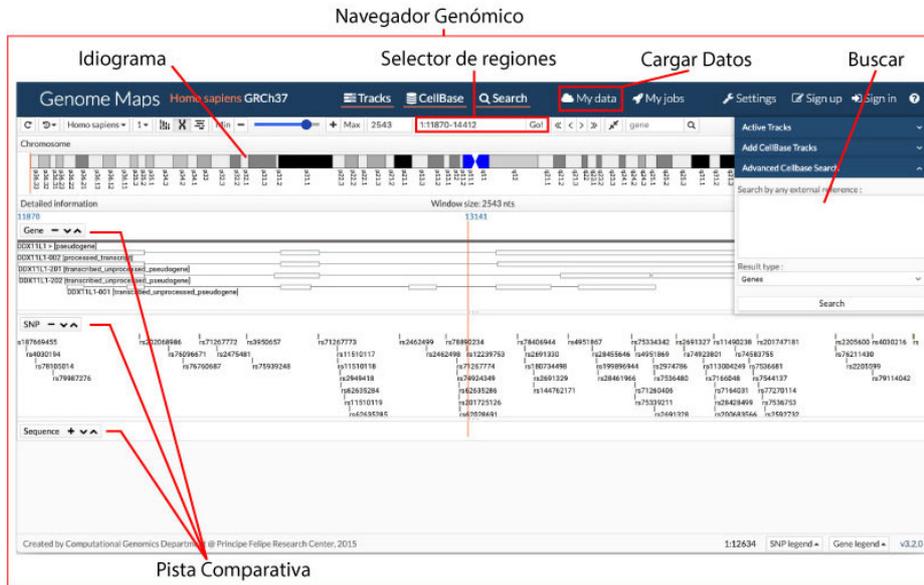


Imagen 6.44 Ejemplo de composición entre patrones. Tomado y adaptado de Genome Maps

6.3 PASOS DEL PROCESO DE DISEÑO

6.3.1 Diseño arquitectónico

Un sistema interactivo se forma de varias interfaces de usuario conectadas de una manera lógica. El diseño arquitectónico tiene como objetivo definir la arquitectura de la información del sistema interactivo, especificando las interfaces de usuario que conformarán el sistema de análisis genómico y el flujo de navegación entre ellas.

La especificación de la interfaz de usuario involucra considerar aspectos complejos de estética (p. ej., gráficos, posicionamiento de los elementos, fuentes, imágenes y color) y de comportamiento (p. ej., funcionalidades, navegación, dinamismo de los elementos). Para simplificar la complejidad en la especificación, adoptamos el enfoque basado en patrones de Molina (Molina Moreno, 2003) quien considera a una interfaz de usuario como una *unidad de interacción* (UI). Una UI es una “unidad de presentación donde no sólo se abstrae la presentación sino también el comportamiento de la unidad por medio de tipificación según patrones”. Una ventana, o una página web pueden ser ejemplos de una UI la cual está formada por una combinación de patrones. En consecuencia, las UIs que definimos en este paso del proceso serán tipificadas con patrones en los pasos posteriores del proceso de diseño.

Para diseñar la arquitectura de información del sitio interactivo, el desarrollador usa patrones de información del catálogo GenomIU los cuales expresan la arquitectura de información a través de nodos y enlaces. Por ejemplo, el desarrollador puede usar el patrón

“Secuencial” si la idea es guiar al usuario a través de una serie de pasos en el análisis genómico. Este patrón es un patrón publicado en la página 121 del libro *Web Style Guide* (Lynch et al., 2016) donde el autor indica que: “la forma más fácil y familiar para organizar la información es ubicarla en una secuencia...Las secuencias rectas son la organización más adecuada para sitios de capacitación o educación, por ejemplo, en los cuales se espera que el usuario avance a través de un conjunto fijo de material y los únicos enlaces de navegación son aquellos que soportan la navegación lineal”. La Imagen 6.45 resume la definición del patrón secuencial, donde el nodo con la letra “H” (“home”) representa a la primera UI de la secuencia.



Imagen 6.45 Patrón secuencial. Tomado y modificado de (Lynch, 2016)

El resultado de este paso es un modelo de la arquitectura de información del sistema genómico que incluye las UIs como nodos y los enlaces entre ellas, similar a la Imagen 6.45.

6.3.2 Diseño estructural

En el paso anterior se definieron las UIs y cómo estas están relacionadas en el sistema interactivo de acceso a datos genómicos. Ahora, la atención se centra en establecer la estructura interna de cada UI definiendo la organización de los elementos de contenido de la interfaz. Para esto, el desarrollador se apoya de los patrones de página del catálogo. Por ejemplo, el patrón “Pestaña de módulo” es un patrón de página usado cuando el objetivo es presentar gran cantidad de contenido heterogéneo que puede incluir bloques de texto, formularios, imágenes, gráficos, etc. Este patrón está documentado bajo el nombre *Module Tabs* en la página 155 del libro *Designing interfaces: Patterns for effective interaction design* (Tidwell, 2010) donde la instrucción descrita por la solución del patrón es: “Colocar los módulos de contenido en un área pequeña con pestañas para que solo un módulo sea visible a la vez. El usuario hace clic en las pestañas para llevar diferentes módulos al frente”. La Imagen 6.46 muestra un ejemplo del patrón Pestaña de módulo aplicado en una UI para estructurar el contenido en tres pestañas.

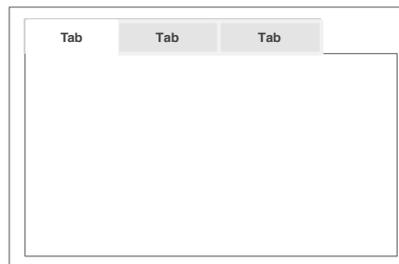


Imagen 6.46 Modelo de diseño estructural de la UI aplicando el patrón Pestaña de módulo

Este diseño de estructura puede visualizarse en sitios web de datos genómicos como, por ejemplo, el sitio ICGC Data Portal (Imagen 6.47), un sitio especializado en datos genómicos sobre el Cáncer, que ha implementado el patrón Pestaña de módulo para estructurar el contenido en tres pestañas de las cuales la pestaña *Saved Sets* es la pestaña activa, es decir, la que presenta contenido.

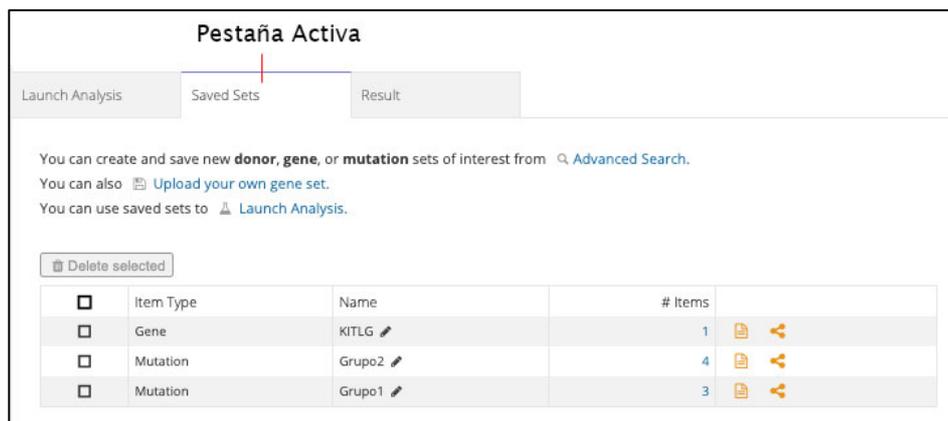


Imagen 6.47 Patrón Pestaña de módulo. Tomado y adaptado de ICGC Data Portal (<https://dcc.icgc.org/analysis/sets>).

El resultado de este paso es el modelo estructural de cada UI, similar al ilustrado en la Imagen 6.46, es decir, el esbozo de la estructura interna de la UI basado en los patrones de página seleccionados por el desarrollador.

6.3.3 Diseño de contenido

Este paso consiste en indicar los elementos de contenido que se mostrarán en la organización definida por los patrones de página. Aquí, “indicar” se refiere a señalar los elementos de contenido que integrarán la UI, mas no a definir o mostrar los detalles visuales de cada elemento. Estos detalles son abordados más adelante.

Para indicar los elementos de contenido, el desarrollador se apoya de los *patrones genómicos* y los *patrones contenedores de información*. En el caso de los patrones genómicos (Sección 6.2.5), por ejemplo, para indicar que en la primera pestaña del ejemplo anterior (Imagen 6.46) se requiere presentar gráficamente el complejo cromosómico de una especie bajo estudio (p. ej., la especie humana), el desarrollador dibujará un recuadro con el nombre *Idiograma* (como lo muestra la Imagen 6.48) haciendo referencia al patrón genómico “Idiograma” el cual cumple con la mencionada necesidad. Si, además, se requiere indicar condiciones de filtro que afecten a la información presentada en el Idiograma, el desarrollador dibujará recuadros con el texto *Filtro* haciendo referencia al patrón genómico “Filtro Genómico”. Nótese que, los detalles de presentación del Idiograma y de los Filtros Genómicos no son importantes, más sí saber qué patrones serán parte de la primera pestaña.

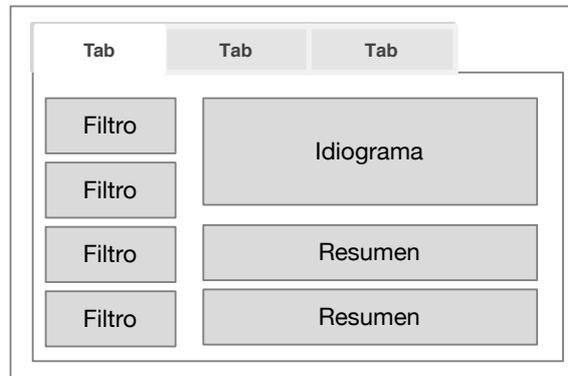


Imagen 6.48 Diseño de contenido ilustrado con el patrón Idiograma y Filtro Genómico

Los patrones contenedores de información son complementarios a los patrones genómicos y son útiles para potenciar el contenido de la UI. Por ejemplo, la información presentada en el Idiograma de la Imagen 6.48 puede tener mucha información relacionada en otras páginas del sitio y su contenido puede ser extenso. Para evitar que el usuario dedique mayor tiempo a leer la extensa cantidad de información, una buena decisión de diseño puede ser mostrar al usuario la información de manera resumida. El patrón *Resumen*, definido en (Seffah et al., 2012) bajo el nombre “Execuctive Summary Pattern” y utilizado en trabajos relevantes de patrones (Javahery, 2007), es adecuado para este propósito. Así, el desarrollador indica varios recuadros con el nombre *Resumen* bajo el Idiograma.

El resultado de este paso es un modelo de la UI similar al de la Imagen 6.48 con los patrones de contenido (genómicos y contenedores de información) indicados en la estructura de la UI.

6.3.4 Refinamiento

El paso final del proceso de diseño consiste en “instanciar” cada patrón de contenido. Aquí, *instanciación* significa adaptar la solución descrita por el patrón al contexto particular del sistema de análisis genómico que se está diseñando, indicando los detalles de visualización que tendrá el patrón en función de los datos reales y particulares del sistema. Por ejemplo, la instanciación del patrón Idiograma indicado en la Imagen 6.48 del ejemplo anterior, resultaría en la representación gráfica y ordenada de 23 pares de cromosomas (si los datos son de la especie humana) cada uno desplegando su patrón de bandas como muestra el ejemplo de este patrón instanciado en el sitio e!Ensembl (Imagen 6.26).

El resultado de este paso es el diseño conceptual de la interfaz de usuario con los patrones de contenido instanciados. Este paso marca la finalización del proceso de diseño del concepto de la interfaz de usuario y puede ser evaluado por los usuarios. Es importante considerar que el proceso es iterativo y el desarrollador puede repetir varias veces la secuencia de pasos del proceso hasta conseguir un diseño conceptual ajustado a las necesidades del usuario. El diseño conceptual, obtenido tras varias iteraciones del proceso de diseño, se considera un artefacto listo para ser implementado. Sin embargo, la etapa de

“construcción” del proceso de diseño total de la interfaz de usuario incluyendo los detalles de implementación están fuera del alcance de esta Tesis.

6.4 CONCLUSIONES

Este capítulo se ha descrito el método para el diseño de interfaces de usuario de acceso a datos genómicos: GenomIUm. El método aborda el diseño conceptual de la interfaz de usuario de un sistema interactivo de acceso a datos genómicos y consiste en un catálogo de patrones, denominado GenomIU, y un proceso de diseño en el cual se aplica el catálogo.

El proceso de diseño guía al desarrollador en el diseño del concepto de la interfaz de usuario a través de cuatro pasos (diseño arquitectónico, diseño estructural, diseño de contenido y refinamiento) que cubren desde la organización del espacio de información del sistema genómico en interfaces de usuario hasta el refinamiento de cada uno de los elementos de la interfaz de usuario.

El catálogo GenomIU está diseñado para soportar al proceso de diseño y se presenta en forma de estructura de árbol donde cada rama corresponde a una categoría de patrones que apoya a un paso específico del proceso de diseño. La estructura de árbol facilita también la búsqueda de patrones a los desarrolladores durante el proceso de diseño de la interfaz. Para soportar el proceso de diseño, el catálogo GenomIU reúne, por un lado, patrones definidos por otros autores en librerías de patrones externas y a los cuales hemos denominado “patrones genéricos”, y por otro, patrones definidos por nosotros en esta Tesis y a los cuales hemos denominado “patrones específicos”. Mientras los patrones genéricos abordan problemas de diseño generales de la IU, los patrones específicos abordan problemas de diseño específicos en el acceso a datos genómicos. El principal aporte de esta Tesis está en la definición de los patrones específicos a los cuales hemos denominado “patrones genómicos”. El catálogo GenomIU es el resultado de la captura de soluciones de diseño a partir del estudio sistemático y continuo de las interfaces de usuario implementadas por herramientas existentes en el dominio genómico y BigData. El conjunto de herramientas estudiadas forma un inventario de herramientas que hemos utilizado repetidamente durante la identificación de los patrones.

El proceso de diseño junto con la definición de los patrones genómicos que conforman el catálogo GenomIU se convierten en las contribuciones relevantes de esta Tesis. Complementario a estas contribuciones principales, en este capítulo también se describieron los instrumentos que fueron necesarios para la creación del catálogo de patrones tales como: el formato para documentar los patrones y el proceso sistemático para la identificación de los patrones y el inventario de herramientas.

El siguiente capítulo de la Tesis, aborda la validación del método descrito en este capítulo.

PARTE IV. VALIDACIÓN

Capítulo 7

Validación

7.1 INTRODUCCIÓN

El siguiente paso en el ciclo de la metodología DS es validar el diseño de GenomIUm. Nuestro objetivo es evaluar las partes esenciales del método (patrones y proceso de diseño). Por lo tanto, en este capítulo presentamos dos experiencias empíricas:

- Un *experimento controlado de comparación estadística* donde evaluamos el impacto de los patrones en la usabilidad de las interfaces.
- Un *experimento de mecanismo simple* que evalúa la percepción de desarrolladores de interfaces respecto al proceso de aplicación de los patrones.

Los dos tipos de experimentos son parte del conjunto de *métodos de investigación* descritos en (Wieringa, 2014) como métodos aplicables en la metodología DS para validar una solución. Finalmente, para ilustrar la aplicación del método en un ambiente real, detallamos una tercera experiencia:

- Un caso de estudio donde aplicamos GenomIUm en un proyecto real.

7.2 VALIDACIÓN DE INTERFACES IMPLEMENTADAS CON PATRONES

Como un primer paso en la evolución de los patrones genómicos, identificamos un conjunto de cuatro patrones que fueron descritos en la Sección 6.2.3. En esta sección presentamos la validación de dichos patrones identificados en las primeras etapas de la investigación. Por la temporalidad en la que fue realizada esta validación, este estudio no cubre el conjunto total de patrones genómicos. Sin embargo, este estudio aporta con evidencia del impacto de este primer conjunto de patrones y establece la estructura y procedimiento para validar el conjunto total de patrones.

7.2.1 Objetivo y preguntas de investigación

El propósito de este estudio es saber hasta qué punto las interfaces diseñadas con patrones genómicos permiten a los usuarios lograr sus objetivos.

El contexto escogido para realizar el estudio es la búsqueda de literatura biomédica, una actividad frecuente en el trabajo de los investigadores (genetistas, médicos, estudiantes, etc.)

inmersos en el dominio genómico. En este contexto, nos enfocamos en el escenario de “la identificación de variaciones genéticas asociadas a una o varias enfermedades genéticas”, donde los investigadores usan las bases de datos de información genómica del NCBI para buscar literatura biomédica asociada a la relación enfermedad-variación.

El estudio está basado en un experimento de comparación estadística, donde nuestra idea es i) rediseñar la interfaz original de búsqueda de literatura biomédica del NCBI con los patrones GenomIU, ii) solicitar a un conjunto de usuarios realizar varias tareas de búsqueda de literatura con las interfaces rediseñadas y originales, iii) evaluar el impacto de las interfaces rediseñadas con patrones comparando la usabilidad de estas interfaces con la usabilidad de las interfaces originales usando métodos estadísticos aplicados a las respuestas de los usuarios.

La norma ISO/IEC 25010 define a la *usabilidad en el uso* como “el grado en que los usuarios específicos pueden lograr objetivos específicos con efectividad en el uso, eficiencia en el uso y satisfacción en el uso en un contexto de uso específico”. Para simplificar la redacción, nos referiremos a la efectividad en el uso, eficiencia en el uso y satisfacción en el uso con términos simples como efectividad, eficiencia y satisfacción, respectivamente. Por lo tanto, basado en la plantilla de Meta/Pregunta/Medida (MPM) (Van Solingen et al., 2002), el objetivo de este estudio empírico es:

Analizar	<i>Las interfaces de usuario diseñadas con los patrones GenomIU</i>
Con el propósito de	<i>Evaluar la usabilidad en el uso</i>
Con respecto a	<i>La efectividad, eficiencia y satisfacción</i>
Desde la perspectiva de	<i>El investigador</i>
En el contexto de	<i>Analistas buscando documentos</i>

A partir del objetivo de investigación, derivamos varias preguntas de investigación (PI). Dado que el objetivo está basado en la norma ISO/IEC 25010 que define la usabilidad en el uso en términos de tres atributos (efectividad, eficiencia y satisfacción), las PI y las hipótesis nulas correspondientes (H_0) están formuladas respecto a estos tres atributos:

PI1: ¿El uso de patrones GenomIU afecta la efectividad del usuario en la interacción con la aplicación? La hipótesis derivada de la pregunta es:

H_{01} : La efectividad interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la efectividad observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones.

PI2: ¿El uso de patrones GenomIU afecta la eficiencia del usuario en la interacción con la aplicación? La hipótesis derivada de la pregunta es:

H_{02} : La eficiencia interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la eficiencia observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones.

PI3: ¿El uso de patrones GenomIU afecta la satisfacción del usuario en la interacción con la aplicación? La hipótesis derivada de las preguntas es:

H_{03} : La satisfacción interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la satisfacción observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones.

7.2.2 Planeamiento

Factores y tratamientos

Nuestro experimento estudia un factor con dos niveles (también conocido como tratamientos o alternativas). La selección del factor depende del propósito del experimento (Kuehl, 2001); por lo tanto, en nuestro experimento, el factor es el “diseño de interfaces de usuario para analizar literatura”. Los niveles o tratamientos son las alternativas de los factores que nos ayudan a responder a las hipótesis de investigación. Trabajamos con dos tratamientos a los cuales hemos nombrado en función de si los patrones son o no son empleados:

- **T.sin-patrones** o Tratamiento sin patrones: este tratamiento representa a las interfaces de usuario para el análisis de literatura que no han sido diseñadas con los patrones GenomIU. Este tratamiento se ha operacionalizado a través de interfaces de usuario reales existentes que no incorporan los patrones del método. T.sin-patrones se conoce como *tratamiento control*.
- **T.con-patrones** o Tratamiento con patrones: este tratamiento representa a las interfaces de usuario similares al tratamiento control, pero su diseño se ha modificado aplicando los patrones GenomIU. En otras palabras, las interfaces de usuario del tratamiento control fueron rediseñadas incorporando los patrones, cambiando las características visuales, pero manteniendo la funcionalidad de la interfaz de usuario original.

Estos dos tratamientos se convierten en variables independientes y son los causantes de los efectos producidos en las variables de respuesta que queremos medir en el experimento. Las variables de respuesta se describen en la siguiente sección.

Variables de respuesta y métricas

Las variables de respuesta (también conocidas como variables dependientes o variables explicativas) son las características que deben medirse en el experimento y su valor depende de la respuesta de los sujetos experimentales cuando utilizan los tratamientos. Las variables de respuesta se derivan directamente de las PI definidas en 7.2.1.

La PI1 está relacionada con la variable de respuesta: *efectividad*. De acuerdo con la norma ISO/IEC-25010, la efectividad es *el grado en que usuarios específicos pueden lograr objetivos específicos con precisión y completitud en un contexto de uso específico*. Para medir la efectividad se consideró la *completitud*, esto es, la proporción de las tareas completadas por los sujetos experimentales, por lo cual, se solicitó a los usuarios que realicen varias tareas para resolver un problema experimental. Las tareas se expresan en forma de preguntas abiertas y para responderlas el usuario debe, en ciertos casos, hacer clic en la interfaz de usuario o, en otros casos, seleccionar una respuesta de una lista de respuestas. Dependiendo del caso, la respuesta del usuario registra el lugar donde el usuario hizo clic o la opción de respuesta seleccionada de la lista. Cada respuesta se guarda y es evaluada asignando uno de los dos valores: éxito (resultado=1) o falla (resultado = 0). De esta manera, la efectividad es calculada dividiendo el porcentaje de tareas completadas con éxito para el número total de tareas propuestas en

el problema experimental. Cuanto mayor sea el porcentaje obtenido, mayor será la efectividad.

La PI2 está relacionada con la variable de respuesta: *eficiencia*. De acuerdo con la norma ISO/IEC-25010, la eficiencia es *el grado en el cual usuarios específicos emplean la cantidad adecuada de recursos en relación con la efectividad lograda en un contexto de uso específico*. De esta manera, la eficiencia puede ser entendida como la cantidad de tiempo (segundos) que un sujeto emplea para completar una tarea. Durante el experimento, el tiempo empleado por el sujeto experimental en responder a cada tarea (desde el inicio de la tarea hasta que esta es respondida) se cronometra automáticamente en segundos. Por lo tanto, la eficiencia de cada sujeto experimental se calcula como la efectividad lograda por el sujeto experimental dividida por el tiempo consumido para completar la tarea. Cuanto menor sea el número de segundos, mayor será la eficiencia.

La PI3 está relacionada con la variable de respuesta: *satisfacción*. Según ISO/IEC-25010, la satisfacción se define como *el grado en que los usuarios están satisfechos en un contexto de uso específico*. Para medir esta variable, utilizamos el cuestionario CSUQ de IBM (Lewis, 1995) ampliamente aplicado para evaluar la satisfacción del usuario en el contexto de los estudios de usabilidad debido a su fiabilidad aceptable (el coeficiente alfa de Cronbach excede a 0,89 (Lewis, 1995)). El cuestionario contiene 19 preguntas que se evalúan individualmente de manera cualitativamente a través de una escala Likert de 7 puntos desde “totalmente en desacuerdo” (1) a “totalmente de acuerdo” (7). Además, el cuestionario contiene tres preguntas abiertas que solicitan información sobre aspectos negativos, positivos y recomendaciones sobre las interfaces de usuario evaluadas. De esta manera, la satisfacción total de un sujeto experimental está dada por la suma de todos los puntos de la escala seleccionados por el sujeto experimental. Por lo tanto, si una encuesta estuviera marcada como “totalmente de acuerdo” para las 19 preguntas, dicha encuesta alcanzaría una puntuación máxima de 133 (7 puntos de la escala de Likert multiplicado por 19 preguntas del cuestionario). Por lo tanto, cuanto mayor sea la suma total, mayor será la satisfacción. Los sujetos experimentales completan el cuestionario CSUQ después de completar todas las tareas propuestas para un problema.

La Tabla 15 resume las variables de respuesta, las métricas para medirlas y cómo se relacionan con cada PI y las hipótesis.

Tabla 15 Resumen de las Preguntas de Investigación, hipótesis, variables de respuesta y métricas.

PI	Hipótesis	Variable Respuesta	Métrica	Comentarios
PI1	H ₀₁	Efectividad	$\frac{\sum_{i=1}^N n_i}{N} \times 100\%$	N = número total de tareas. n _i = el resultado de la tarea i. t _i = el número de segundos utilizados para realizar la tarea i.
PI2	H ₀₂	Eficiencia (tarea/seg.)	$\frac{\sum_{i=1}^N \frac{n_i}{t_i}}{N}$	
PI3	H ₀₃	Satisfacción	Cuestionario CSUQ	

Sujetos experimentales

El objetivo de este estudio es analizar los efectos de los patrones en interfaces de búsqueda de literatura biomédica, por cuanto los sujetos experimentales (las personas que participan en el experimento) son personas que tienen conocimiento del dominio genómico y que consumen información biomédica. Sin embargo, hemos pensado que sería interesante estudiar si el conocimiento del usuario acerca del dominio influye en los efectos producidos por los tratamientos. Es por eso por lo que, hemos ampliado el conjunto de sujetos experimentales para abarcar personas no inmersas en el dominio genómico y que están enfocadas en la búsqueda de literatura en otras áreas.

El experimento se compone de dos réplicas: la Réplica 1 (R1) se llevó a cabo con 24 estudiantes de pregrado del grado de Información y Documentación de la Universidad de Valencia y la Réplica 2 (R2) se llevó a cabo con ocho investigadores del Grupo de Genómica del PROS (siete estudiantes del grado y máster de bioinformática y un PhD en informática) de la Universidad Politécnica de Valencia. El perfil académico de los sujetos experimentales se resume en la Tabla 16.

Tabla 16 Perfil académico de los sujetos experimentales por réplica.

Réplicas	Pregrado	Master	Ph.D.	Total
R1	24	0	0	24
R2	3	4	1	8

Los sujetos de R1 trabajan frecuentemente en bibliotecas y centros de información y documentación realizando con frecuencia tareas de búsqueda, organización y evaluación de información de literatura científica utilizando el sitio web de búsqueda Scopus⁶. Los sujetos de R2 trabajan como investigadores en la academia que se dedican a estudiar el genoma humano, específicamente en la identificación de las variaciones genéticas del ADN que causan enfermedades y una de las tareas más frecuentes en su trabajo es buscar evidencia clínica utilizando el sitio web PubMed⁷ (base de datos del NCBI conteniendo información de literatura biomédica).

Tabla 17 Frecuencia de uso de sitios web de búsqueda utilizados por los sujetos de la réplica 1 (R1) y replica 2 (R2).

Replica	Sitio web	Siempre	Regularmente	Ocasionalmente	Nunca
R1	PubMed	0	0	2	22
	Scopus	23	1	0	0
R2	PubMed	8	0	0	0
	Scopus	0	0	1	7

⁶ <https://www.scopus.com>

⁷ <https://www.ncbi.nlm.nih.gov/pubmed/>

La Tabla 17 muestra la frecuencia de uso de los sitios web de búsqueda para los sujetos experimentales en cada réplica (R1 y R2). Respecto a la frecuencia de uso de Scopus para buscar información bibliográfica, 23 sujetos de R1 informaron que siempre usan Scopus para buscar información y solo 1 sujeto usa Scopus regularmente, mientras que para los sujetos de R2 el uso de Scopus es ocasional o nunca. Respecto a la frecuencia de uso de PubMed para buscar bibliografía biomédica, todos los sujetos de R2 informaron que siempre usan PubMed como sitio web de búsqueda. Esto es consistente con el hecho de que los sujetos en R2 se enfocan en estudiar temas biológicos y médicos utilizando en gran medida PubMed en lugar de Scopus.

Objetos experimentales

Para observar los efectos producidos por los dos tratamientos (es decir, T.con-patrones y T.sin-patrones), definimos dos problemas, a los que nos referimos como P.scopus y P.pubmed. Estos problemas implican analizar bibliografía, en dos escenarios: “literatura científica” y “literatura biomédica”. Seleccionamos estos escenarios ya que los sujetos experimentales de cada replicación son expertos en uno de estos escenarios (ver Tabla 17). Los sujetos experimentales de R1 son expertos en literatura científica, mientras que los sujetos de R2 son expertos en literatura biomédica, por lo que es posible analizar si los efectos producidos por los tratamientos se ven afectados por el conocimiento previo del sujeto experimental en el dominio.

El problema P.pubmed consiste en analizar la bibliografía clínica mediante el uso de PubMed; una de las bases de datos más importantes de la Biblioteca Nacional de Medicina de los EEUU, que es administrada por el NCBI y que contiene citas de artículos científicos y médicos completos almacenados en la base de datos MEDLINE. El objetivo de este problema consiste en *buscar en la bibliografía la información sobre la enfermedad genética denominada “Cri-du-chat” y analizar los documentos clínicos resultantes para identificar las mutaciones y los genes que causan la enfermedad.* Para lograr este objetivo, se solicita al sujeto que realice varias tareas que se han expresado como preguntas para su mejor comprensión, como se muestra en la Tabla 18. Para responder a las preguntas, se proporcionan dos prototipos de interfaz de usuario, uno para cada tratamiento.

Tabla 18 Preguntas para el problema PubMed.

N°	Pregunta
1	Buscar información de la enfermedad “ <i>Cri du chat Syndrome</i> ”. ¿Cuántos documentos fueron encontrados?
2	¿En qué intervalo de años fueron publicados los documentos encontrados?
3	¿Cuántos documentos fueron publicados en los 10 últimos años?
4	¿Cuántos genes en la base de datos “Gene” están relacionados a la lista de artículos clínicos de PubMed?
5	¿Cuántas variaciones genéticas de “ClinVar” están relacionadas con artículos de PubMed?

El prototipo de interfaz de usuario relacionado con T.sin-patrones consiste en interfaces de usuario que simulan las interfaces de usuario de PubMed a través de imágenes de las

interfaces existentes, como se muestra en el ejemplo “sin patrones” de la Imagen 7.1, número 2.

El prototipo de interfaz de usuario relacionado con T.con-patrones consiste en imágenes de interfaces de usuario análogas a las interfaces existentes de PubMed, pero diseñadas con los patrones, como se muestra el ejemplo “con patrones” de la Imagen 7.1, número 2. Las partes de la interfaz de usuario donde se han aplicado los patrones se indican con etiquetas de literales: a) visualizar, conectar, filtrar b) Entrega implícita, c) Filtro móvil y d) Caja Operaciones. Es importante notar que la aplicación de los patrones cambia el aspecto visual de las interfaces de usuario originales de PubMed, pero la funcionalidad sigue siendo la misma que presta el sitio original.

El problema de P.scopus se centra en analizar la literatura mediante el uso de Scopus, que es una base de datos bibliográfica de artículos científicos. Actualizado a la fecha, Scopus contiene⁸ aproximadamente más de 22,000 títulos indexados de más de 5000 editores. El objetivo de este problema es *buscar y analizar bibliografía relacionada con el tema “Diseño de interacción” y analizar la información que se muestra en la pantalla*. Para lograr esto, se solicita al sujeto experimental que realice varias tareas indicadas en la Tabla 19 y que se han expresado como preguntas para mejor comprensión del usuario.

El problema P.scopus contiene más preguntas que el problema P.pubmed; sin embargo, el problema P.pubmed implica un análisis más complejo, ya que implica tratar los términos biológicos y clínicos e identificar los documentos clínicos relacionados con los elementos genéticos (es decir, genes y variaciones genéticas).

Al igual que en el problema P.pubmed, construimos dos prototipos de interfaz de usuario basados en imágenes, uno para cada tratamiento. El prototipo de interfaz de usuario relacionado con T.sin-patrones consiste en una interfaz de usuario real de Scopus y el prototipo de interfaz de usuario relacionado con T.con-patrones contiene interfaces de usuario diseñadas con los patrones GenomIU.

Tabla 19 Preguntas para el problema Scopus.

Nº	Pregunta
1	Buscar información sobre el tema: “Interaction Design”. ¿Cuántos resultados fueron obtenidos?
2	¿En qué intervalo de años fueron publicados los resultados?
3	¿Cuáles son los dos años que reportaron el mayor número de documentos publicados?
4	¿Cuál es el autor con el mayor número de documentos publicados?
5	¿En qué país se concentra el mayor número de documentos publicados?
6	¿Que proporción del total de documentos representa el conjunto de documentos entre 2008 y 2018?
7	¿Cuántos documentos del tipo “artículo” existen?
8	¿Cuántas citas tiene el artículo más citado?
9	¿De qué país son los autores que tienen el mayor número de artículos citados?

Es necesario considerar que los prototipos de interfaz de usuario desarrollados son páginas web basadas en imágenes enriquecidas con controles interactivos de interfaz de usuario (p. ej. listas, botones) colocados sobre la imagen. El comportamiento de cada

⁸ <https://www.elsevier.com/solutions/scopus/how-scopus-works/content>

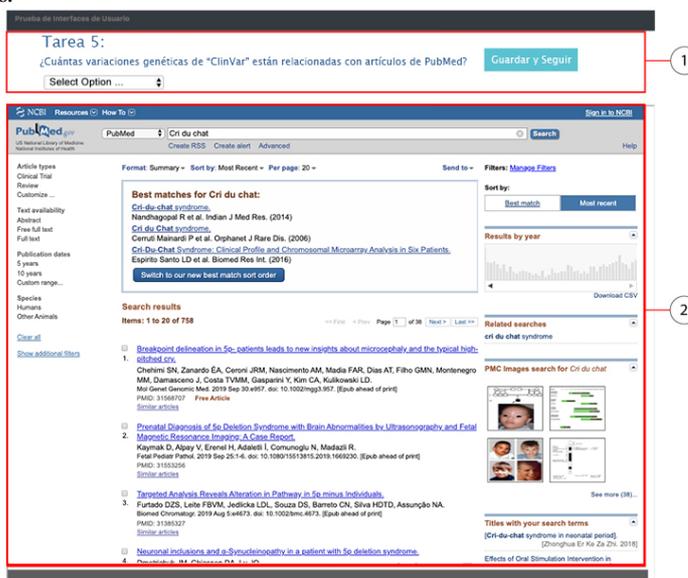
control interactivo está implementado a través de JavaScript. Por ejemplo, el prototipo de interfaz de PubMed con patrones (Imagen 7.1, “con patrones”, número 2), consta de una imagen de fondo y controles de interfaz con los que el usuario interactúa para responder a la pregunta propuesta en el número 1. Cuando el usuario hace clic en el círculo de mayor área (correspondiente al documento más citado) en el cuadro de “Documentos (Matching and citations)”, instrucciones de JavaScript se ejecutan y presentan automáticamente el resumen del artículo seleccionado en el sector inferior del listado de documentos. Por lo tanto, los prototipos de interfaz de usuario desarrollados contienen todos los mecanismos interactivos para responder a las preguntas propuestas.

Plataforma de evaluación. - Para llevar a cabo el experimento se construyó una plataforma de evaluación en línea que hace posible estructurar las tareas de cada problema e integrar a las tareas los prototipos de interfaz de usuario correspondientes para realizar la tarea.

A través de esta plataforma, el usuario puede realizar el experimento de principio a fin, mientras que sus respuestas a las tareas se capturan automáticamente, así como el tiempo en segundos que se tarda en realizarlas.

La plataforma muestra, en una única interfaz, tanto la secuencia de tareas como el prototipo de interfaz de usuario correspondiente para realizar la tarea, como se muestra en cualquiera de los prototipos “con patrones” o “sin patrones” de la Imagen 7.1. El número 1 es el área destinada a mostrar las tareas a realizar, mientras que el número 2 está destinada a presentar el prototipo de interfaz de usuario para realizar la tarea. Al final de cada tratamiento, la plataforma de evaluación muestra el cuestionario CSUQ y recopila los resultados de cada sujeto.

Sin patrones:



Con patrones:

The image displays two versions of a user interface for evaluating PubMed search results. The top version, labeled '1', shows a task prompt: "Tarea 5: ¿Cuántas variaciones genéticas de 'ClinVar' están relacionadas con artículos de PubMed?" with a "Guardar y Seguir" button. The bottom version, labeled '2', shows search results for "Cidu chat Syndrome" with a network graph of citations and a list of related databases. The network graph shows a central node "Cidu chat Syndrome" connected to several other nodes, with a "22%" label. The list of related databases includes "Genetic Variations Related (ClinVar)", "Chromosomal Abnormalities", "Pathogenic", and "Likely pathogenic".

Imagen 7.1 Interfaz de usuario de la plataforma de evaluación mostrando el prototipo de PubMed sin patrones (imagen superior) y con patrones (imagen inferior).

Diseño del experimento

El diseño del experimento consiste en definir la asignación de los tratamientos a los sujetos experimentales. Vegas et. al. (Vegas et al., 2016) describen tres tipos de diseño experimental: *medidas independientes*, *medidas repetidas* y *pares emparejados* (grupos formados por individuos de similares características). Para este experimento se ha elegido el diseño de *medidas repetidas* (también conocido como *diseño experimental intra-sujetos*) ya que nuestra intención es que todos los sujetos experimentales prueben ambos tratamientos (es decir, ambos tratamientos son asignados a todos los sujetos). Como variable bloqueada, hemos utilizado el problema. No estamos interesados en estudiar si en un problema específico se obtienen mejores resultados que en otros problemas, es por eso por lo que hemos bloqueado el problema aplicando ambos problemas a ambos tratamientos.

Los beneficios de usar el diseño de medidas repetidas están relacionados con a) la maximización del tamaño de la muestra y b) contrarrestar los efectos del aprendizaje y la fatiga (Vegas, 2016). El tamaño de la muestra se maximiza ya que cada sujeto experimental aplica todos los tratamientos, por lo tanto, habrá medidas repetidas para un mismo sujeto experimental (uno para cada tratamiento). El efecto de aprendizaje se contrarresta ya que los sujetos experimentales aplican los tratamientos en diferente orden, por lo que el conocimiento aprendido en un primer tratamiento no se aplica ni se transfiere a los tratamientos posteriores. El efecto de la fatiga también se contrarresta aplicando los tratamientos en diferente orden. De esta manera, el problema y las tareas a realizar con un tratamiento son diferentes de las de los tratamientos posteriores, asegurando que el sujeto

experimental no se cansa ni se aburra haciendo las mismas tareas. Además, este diseño evita que los sujetos experimentales confundan el tratamiento con el problema debido a que el orden de secuencia en el que deben realizarse tanto el problema como el tratamiento está claramente definido.

La Tabla 20 muestra la configuración de medidas repetidas para el experimento. El diseño se estructura en dos sesiones (dos períodos de tiempo) dentro del mismo día; el segundo comienza cuando el primero termina. En cada sesión, un sujeto experimental aplica un tratamiento para resolver un problema. Se definen cuatro grupos de sujetos experimentales equilibrados al azar (es decir, grupos del mismo número de sujetos experimentales asignados al azar) para asignarlos a las cuatro posibles secuencias de tratamiento-problema. Cada secuencia se compone de la sesión 1 (primer problema, primer tratamiento) y la sesión 2 (segundo problema, segundo tratamiento), como se muestra en la Tabla 20. En consecuencia, cada grupo tiene un orden de secuencia diferente para aplicar el tratamiento en un problema. Por ejemplo, un sujeto experimental de G1 ejecutará primero el P.scopus utilizando T.sin-patrones (Sesión 1) y luego, ejecutará el P.pubmed utilizando T.con-patrones (Sesión 2). De esta manera, el experimento está diseñado de forma que todos los sujetos experimentales aplican cada tratamiento a un problema dentro de cada sesión.

Tabla 20 Diseño del experimento.

Grupo	Sesión 1		Sesión 2	
	1er Problema	1er Tratamiento	2do Problema	2do Tratamiento
G1	P.scopus	T.sin-patrones	P.pubmed	T.con-patrones
G2	P.pubmed	T.con-patrones	P.scopus	T.sin-patrones
G3	P.scopus	T.con-patrones	P.pubmed	T.sin-patrones
G4	P.pubmed	T.sin-patrones	P.scopus	T.con-patrones

Procedimiento del experimento

El diagrama de flujo de la Imagen 7.2 resume el procedimiento que seguimos para llevar a cabo el experimento. El procedimiento se basa estrictamente en la configuración de diseño del experimento que se explica en la Tabla 20. El procedimiento se ha etiquetado claramente con números para explicar cada paso.

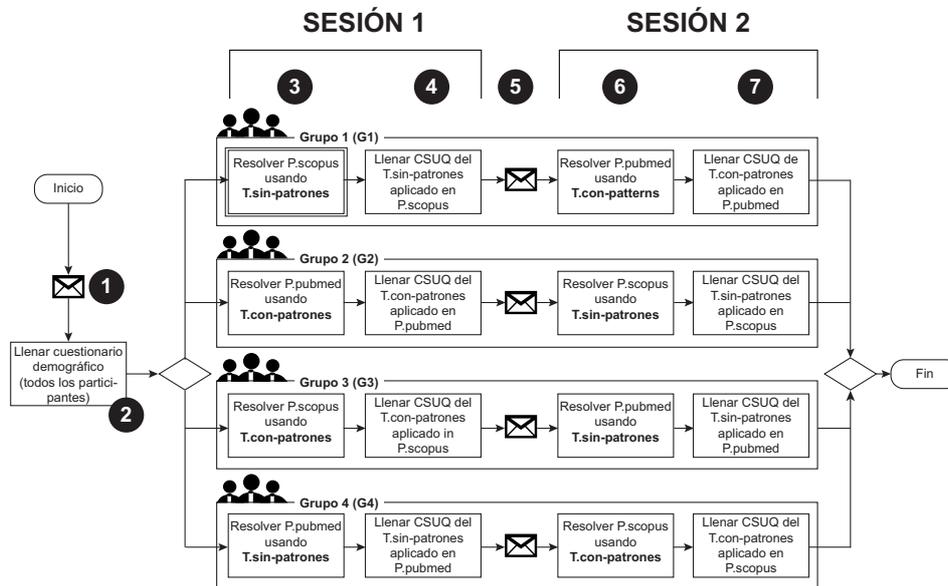


Imagen 7.2 Procedimiento del experimento

Antes de comenzar el experimento, explicamos los objetivos del experimento a los sujetos experimentales, así como el papel que desempeñan en él. Además, creamos al azar los cuatro grupos de sujetos (es decir, G1, G2, G3 y G4) y nos aseguramos de que el número de sujetos entre los grupos sea equilibrado.

Paso 1.- El experimentador envió a cada sujeto experimental la URL de la plataforma de evaluación correspondiente al tratamiento y problema a resolver en la sesión 1 (en total cuatro URLs diferentes, una por cada grupo). Por ejemplo, la URL recibida por los sujetos del grupo G1 en la sesión 1 (rectángulo dibujado en doble línea únicamente para facilitar la explicación), les permitirá acceder al tratamiento T.sin-patrones para resolver el problema P.scopus.

Paso 2.- Una vez que los sujetos experimentales acceden a la plataforma de evaluación, responden un cuestionario demográfico. El cuestionario recopila información sobre el sujeto experimental, incluidas preguntas sobre el perfil académico, el conocimiento de los motores de búsqueda de bibliografía y la experiencia en su uso. Aunque las URLs de la plataforma son diferentes para cada grupo, las preguntas del cuestionario demográfico son las mismas para todos los sujetos experimentales independientemente de su grupo. El cuestionario se automatizó mediante formularios web de Google que facilitan el almacenamiento automático de las respuestas y su posterior descarga como hoja de cálculo de Excel. El cuestionario demográfico se llena una única vez al comienzo del experimento.

Paso 3.- Una vez que se completa el cuestionario demográfico, los sujetos comienzan la sesión 1. Según el grupo al que pertenezca el sujeto experimental, la plataforma de evaluación lo redirige automáticamente al tratamiento y problema correspondiente dentro de la sesión 1 (en el caso del G1 al T.sin-patrones y el problema P.scopus, como lo muestra

el rectángulo dibujado en doble línea). En primer lugar, la plataforma muestra i) una breve explicación sobre el problema a resolver, ii) una breve descripción del prototipo de la interfaz de usuario y iii) las instrucciones para realizar las tareas mediante el uso del prototipo de interfaz. Una vez leídas las instrucciones, el sujeto comienza la secuencia de tareas especificadas para el correspondiente problema (Tabla 18 para P.pubmed y Tabla 19 para P.scopus). La respuesta de cada sujeto, los clics y el tiempo (en segundos) dedicado a responder cada tarea se almacenan automáticamente a través de la plataforma de evaluación.

Paso 4. - Una vez que los sujetos terminan el problema de la sesión 1, la plataforma de evaluación los redirige al cuestionario CSUQ. Aquí, los sujetos evalúan la satisfacción de usar el tratamiento asignado. Este cuestionario CSUQ se creó como una página web utilizando los formularios web de Google desde donde es posible descargar las respuestas como una hoja de cálculo de Excel.

La sesión 1 termina cuando se completa el cuestionario correspondiente.

Paso 5. - Entre las dos sesiones, tenemos un espacio de 3 minutos durante el cual el evaluador envía a los sujetos un correo electrónico con la URL de la plataforma de evaluación para la Sesión 2. Por ejemplo, los sujetos experimentales del G1 de la Sesión 2 recibirán la URL para acceder al T.sin-patronos para resolver el P.pubmed.

Paso 6.- Este paso es similar al Paso 3, cada sujeto interactúa con un problema y un tratamiento diferentes en esta Sesión 2.

Paso 7.- De la misma manera que en el Paso 4, los sujetos llenan el cuestionario CSUQ para medir la satisfacción de usar el tratamiento aplicado en la Sesión 2.

Una vez que el experimento ha terminado, uno de los experimentadores analiza los datos para extraer conclusiones.

Evaluación de amenazas

El análisis de las amenazas a la validez protege los resultados del experimento, evitando la selección inadecuada de pruebas estadísticas, tamaños de muestra y otros temas que influyen en la veracidad de los resultados. Para evaluar la validez, hay 4 categorías de validez (Cook et al., 1979): Conclusión, Interna, Constructo y Externa. A continuación, se describen las amenazas identificadas en el experimento y las consideraciones tomadas para intentar mitigarlas.

Validez de Conclusión. - Las amenazas que afectan a la capacidad de sacar la conclusión correcta sobre el resultado del tratamiento. Estas amenazas son:

Bajo poder estadístico. - Esta amenaza está relacionada con el poder de la prueba estadística, la cual se define como la probabilidad de rechazar la hipótesis nula cuando es falsa (Pituch et al., 2015), es decir, la probabilidad de tomar una decisión correcta. El poder de una prueba estadística depende fuertemente del tamaño de la muestra y el tamaño del efecto: un bajo poder de la prueba estadística impide rechazar la hipótesis nula por tener una muestra pequeña (pocos sujetos experimentales), aún cuando existan diferencias significativas entre los tratamientos. Utilizamos el modelo lineal mixto (MLM) como prueba estadística, y no es posible calcular la potencia estadística para esta prueba. Sin embargo, para enfrentar esta amenaza, hemos estimado a través de G*Power⁹ cuántos sujetos y cuánto *tamaño de efecto*

⁹ <http://www.gpower.hhu.de/>

necesitábamos para obtener un poder del 95%. Encontramos que una muestra de 16 sujetos y un tamaño de efecto *medio* ($d = 0,50$) dan como resultado un 95% de potencia. Por lo tanto, el tamaño de nuestra muestra experimental ($R1 + R2 = 24 + 8 = 32$) es más que suficiente para garantizar efectos significativos y producir una potencia del 95%, lo que reduce la probabilidad de sacar conclusiones erróneas.

Supuestos vulnerados de pruebas estadísticas. - Esta amenaza se refiere a las condiciones previas que deben cumplirse antes de aplicar una prueba estadística. Aplicamos la prueba estadística MLM y el supuesto de que se cumple es la normalidad del residuo. Antes de aplicar el MLM, aplicamos la prueba W de Shapiro-Wilk para verificar la normalidad del residuo. Todos los residuos tuvieron una distribución normal.

Fiabilidad de las medidas. - Esta amenaza está asociada a la confiabilidad de las medidas. Las medidas objetivas son más confiables que las subjetivas, ya que no hay intervención del juicio humano. En nuestro experimento tenemos tres variables de respuesta: dos que se miden objetivamente (“Efectividad” y “Eficiencia”) y una variable de respuesta que se mide subjetivamente (la “Satisfacción”). Las dos primeras variables no dependen de la interpretación del sujeto experimental y se miden automáticamente por la plataforma de evaluación, mientras que la tercera variable depende de las respuestas cualitativas del sujeto. Para minimizar la amenaza de fiabilidad de la medida de la tercera variable (“Satisfacción”) hemos utilizado una escala Likert de 7 puntos para que el sujeto experimental evalúe cualitativamente la satisfacción y hemos sumado las respuestas de los sujetos experimentales para convertir las respuestas cualitativas en respuestas cuantitativas que pueden ser computadas (Jamieson, 2004).

Heterogeneidad aleatoria de los sujetos. - Esta amenaza se refiere al riesgo de tener un grupo de sujetos que sea “muy” heterogéneo o “muy” homogéneo. Mientras que en grupos muy heterogéneos las diferencias individuales pueden ser mayores que las diferencias producidas por los tratamientos, en grupos muy homogéneos el grupo de sujetos puede no ser representativo de la población del estudio. Evitamos esta amenaza haciendo dos réplicas del mismo experimento donde el grupo de cada réplica es homogéneo (con conocimientos y antecedentes similares), pero los grupos entre las réplicas son heterogéneos (cada grupo pertenece a dominios diferentes).

Validez interna. - Las amenazas que, sin el conocimiento del investigador, pueden afectar las conclusiones sobre una posible relación causal entre el tratamiento y el resultado. Las amenazas son:

Maduración. - Se refiere a las causas que hacen que el sujeto se comporte diferente a medida que pasa el tiempo. Una causa negativa puede ser que el sujeto se sienta cansado o aburrido. Para minimizar esta amenaza, el experimento se diseñó con dos problemas diferentes que se pueden resolver en poco tiempo, evitando (o al menos minimizando la posibilidad de) que el sujeto se canse o frustre. Sin embargo, dado que la prueba se realiza en línea, el usuario puede distraerse o verse tentado a hacer otras cosas diferentes a la prueba. Para evitar estos efectos, solicitamos a los sujetos que asignen un tiempo específico y las condiciones adecuadas para realizar exclusivamente la prueba. Además, los experimentadores estaban en la misma sala para asegurarse de que los sujetos participaran en el experimento todo el tiempo.

Pruebas. - Esta amenaza aparece cuando el experimento se repite con diferentes sujetos. Por lo tanto, los sujetos tienen un conocimiento previo sobre cómo se lleva a cabo. Nuestro experimento no sufre esto ya que cada repetición tenía diferentes temas.

Instrumentación. - Significa que los objetos experimentales (formas, cuestionarios, prototipos de IU) que no están diseñados adecuadamente pueden afectar negativamente el experimento. Para evitar esta amenaza, la plataforma de evaluación junto con los cuestionarios integrados, los prototipos de interfaz de usuario y las tareas fueron evaluadas previamente por un investigador independiente del grupo de experimentadores. Como resultado de la evaluación previa, se resolvieron problemas de ambigüedad y se cambió el orden de las tareas para mejorar la coherencia, antes de realizar el experimento.

Selección. - Esta amenaza trata con los efectos derivados de la forma en que se seleccionan los sujetos experimentales. El experimento no sufre esta amenaza ya que los sujetos fueron seleccionados por conveniencia, considerando su conocimiento y experiencia en el manejo de motores de búsqueda de bibliografía.

Validez del constructo. - Las amenazas relacionadas con la generalización de los resultados del experimento para justificar el concepto o teoría detrás del experimento:

Sesgo mono-operatorio. - Esta amenaza significa que cuando se usa una única variable independiente, sujeto o tratamiento, el experimento puede representar de forma insuficiente o subestimar el constructo y, por lo tanto, no ofrecer una imagen completa de la teoría. Evitamos esta amenaza utilizando dos tratamientos (T. con-patrones y T.sin-patrones) en dos contextos de búsqueda diferentes y relevantes (es decir, PubMed y Scopus). Por lo tanto, los resultados obtenidos se pueden generalizar para sujetos con un perfil similar al de los sujetos que reclutamos en nuestras réplicas.

Interacción de diferentes tratamientos. - Esta amenaza significa que no hay forma de concluir si el efecto es debido a los tratamientos o debido a una combinación de ellos. No sufrimos esta amenaza ya que en cada sesión se aplica un solo tratamiento por grupo. Una combinación de tratamientos no es posible en el mismo grupo por sesión. Por lo tanto, los efectos son producidos directamente por los tratamientos.

Suposición de hipótesis. - Esta amenaza ocurre cuando los sujetos experimentales determinan el propósito y el resultado esperado del experimento. Para evitar esta amenaza, no explicamos a los sujetos experimentales los objetivos específicos del experimento ni las variables de respuesta que pretendíamos medir.

Temor a la evaluación. - Esta amenaza ocurre cuando los sujetos experimentales tienen miedo de ser evaluados y tratan de estar más atentos y enfocados en la evaluación; un comportamiento que afecta el resultado del experimento. Para evitar esta amenaza, antes de comenzar el experimento, explicamos a los sujetos experimentales que todas las respuestas son importantes, independientemente de que sean positivas o negativas. En las instrucciones de las tareas, teníamos el siguiente texto: “Importante: esta no es una prueba de habilidad (no hay respuestas correctas o incorrectas)”.

Expectativas del experimentador. - Esta amenaza ocurre cuando el experimentador tiende a sesgar consciente o inconscientemente los resultados en base a lo que espera del experimento. Desafortunadamente, este experimento sufre esta amenaza porque el experimentador es el mismo que creó la plataforma de evaluación, asignó al azar tratamientos a los sujetos experimentales y realizó el análisis de los datos.

Validez externa. - Las amenazas concernientes a las limitaciones para generalizar los resultados del experimento a la práctica industrial:

Interacción de selección y tratamiento. - Esta amenaza se refiere al efecto de no tener una muestra representativa de la población objetivo que queremos generalizar. Para reducir esta amenaza, nos aseguramos de que los sujetos experimentales pertenezcan a un entorno en el que la “búsqueda bibliográfica” sea una actividad predominante, como es el caso de la “investigación científica” y la “gestión de bibliotecas”.

Interacción de ambientación y tratamiento. - Esta amenaza se refiere al efecto de no tener el material representativo (p. ej., herramientas) para llevar a cabo el estudio. Nuestro experimento sufre esta amenaza ya que utilizamos prototipos de interfaz en lugar de interfaces reales. Sin embargo, para minimizar esta amenaza, diseñamos y utilizamos prototipos de alta fidelidad que, en el caso del tratamiento de control, lucen y funcionan de manera similar a las interfaces de usuario reales. Las interfaces de usuario del prototipo incorporan el mismo diseño visual, contenido e interacciones que las interfaces reales. Para hacer eso, cada interfaz se diseñó sobre una imagen de la interfaz original en la que agregamos la interacción necesaria para realizar las tareas específicas del experimento.

7.2.3 Análisis de Datos e Interpretación de Resultados

Para contrastar las hipótesis y tomar decisiones sobre si rechazar o no las hipótesis nulas en este estudio, definimos el nivel de significación en 5%, que se interpreta como el 5% de probabilidad de rechazar la hipótesis nula (H_0), cuando ésta es verdadera (Pituch, 2015) (también conocido como error de tipo I o “falso positivo”). En este estudio, cada variable de respuesta se analiza siguiendo dos pasos: 1) análisis descriptivo y 2) análisis analítico.

Primer paso. - Aplicamos estadística descriptiva para observar cómo se comportan las variables de respuesta entre los tratamientos en la muestra de los sujetos. Para esto, compararemos la distribución de datos de cada tratamiento utilizando diagramas de caja y bigotes y explicaremos las diferencias o similitudes en los datos observados en términos de medidas descriptivas como la mediana, la media, la desviación estándar (DE) y los cuartiles (Q1, Q2, Q3, Q4). Esto nos provee una visión general del comportamiento de las variables de respuesta con respecto a los tratamientos.

Segundo paso. - Aplicamos estadísticas inferenciales para determinar si las observaciones basadas en muestras en el primer paso reflejan parámetros a nivel de la población. Para esto, aplicamos el método estadístico “modelo lineal mixto” (MLM), también conocido como *modelo multinivel*, *modelo de jerarquía lineal* o *modelo de coeficiente aleatorio* (Field, 2013). Elegimos este método estadístico por dos razones: 1) este método trata con datos correlacionados producidos por las medidas repetidas y 2) el método es adecuado para analizar más de una medida repetida (Seltman, 2012) como es el caso de nuestro experimento donde tenemos dos medidas repetidas: el tratamiento y el problema.

Para aplicar la prueba estadística MLM, debe cumplirse la suposición de normalidad del residuo. Por lo tanto, aplicamos la prueba de Shapiro-Wilk (W) a los residuos obtenidos con el MLM y correspondientes a cada variable de respuesta y réplica, como se muestra en la Tabla 21. La prueba W es ampliamente recomendada para evaluar si la distribución de datos sigue una distribución normal (Elliott et al., 2006). La mayoría de los residuos cumplieron con el supuesto de normalidad, pero para aquellos que no (específicamente los residuos

correspondientes a la variable de satisfacción en R1), transformamos los residuos aplicando el enfoque de “Dos Pasos” (Templeton, 2011) para lograr la normalidad del residuo. En consecuencia, la prueba W aplicada a los residuos transformados confirmó su normalidad.

Tabla 21 P-valores resultantes de la prueba de normalidad de Shapiro-Wilk aplicado al valor de residuo del Modelo Lineal Mixto (MLM) para la Replica 1(R1) y Replica 2 (R2)

Tratamiento	R1		R2	
	T.con-patrones	T.sin-patrones	T.con-patrones	T.sin-patrones
Efectividad	0.997	0.521	0.932	0.932
Eficiencia	0.422	0.552	0.111	0.194
Satisfacción	0.999(*)	1.000(*)	0.939	0.214

(*) p-valor del residuo previamente transformado usando el enfoque “Dos Pasos” (Templeton, 2011).

Los p-valores, resultantes de aplicar el MLM, nos ayudan a analizar si los tratamientos producen o no efectos diferentes en los sujetos. Si el p-valor es menor o igual al nivel de significación (p-valor ≤ 0.05), concluimos que los tratamientos producen “efectos diferentes” (es decir, hay diferencias significativas entre los tratamientos). De lo contrario, los tratamientos producen los “mismos efectos”.

Si los efectos son diferentes, nuestro objetivo es medir qué tan grande o pequeña es la diferencia, ya que esta medida nos permitirá conocer la magnitud del efecto producido por las interfaces de usuario basadas en patrones. Para ello, estimamos el *tamaño del efecto* (TE), que se define como la *magnitud cuantitativa del efecto en la muestra estudiada*, utilizando la *d* de Cohen (o índice de Cohen) porque este índice es muy recomendable cuando se evalúan las diferencias entre los tratamientos (Ellis, 2010). El valor *d* resultante nos permite definir si el tamaño del efecto es *pequeño* ($d=0.20$), *medio* ($d=0.50$) o *grande* ($d=0.80$), de acuerdo con las convenciones establecidas por Cohen (Cohen, 1988).

Utilizamos el programa estadístico SPSS para llevar a cabo el Paso 1 y el Paso 2. Después de analizar los efectos producidos en cada variable de respuesta, extraemos conclusiones considerando los efectos en: a) cada réplica por separado; y b) la muestra total, combinando los datos de los participantes individuales de ambas réplicas. Este es un enfoque cada vez más popular para sintetizar datos en disciplinas como la medicina, donde se ha demostrado que tiene ventajas potenciales (Riley et al., 2010).

Combinamos las réplicas agregando una variable moderadora que modera la relación entre el tratamiento y la variable de respuesta. Por lo tanto, analizamos el impacto que causa la variable moderadora en la relación entre los tratamientos y la variable de respuesta bajo estudio.

Efectividad

En esta sección, tratamos la pregunta de investigación PI1 y H_{01} : *La efectividad interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la efectividad observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones*. El análisis de efectividad para ambas réplicas (R1 y R2) se describe a continuación.

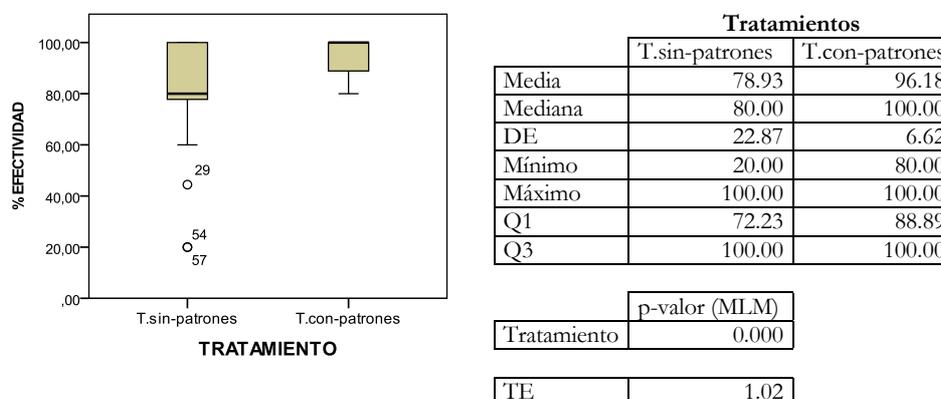


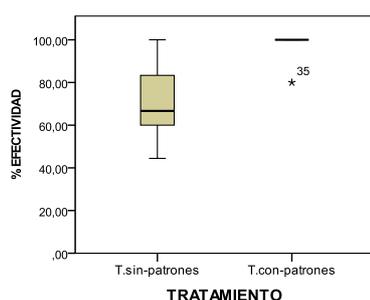
Imagen 7.3 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del Modelo lineal mixto (MLM) y el índice de tamaño del efecto (TE) para la variable efectividad en la Réplica 1 (R1)

El diagrama de caja de la Imagen 7.3 compara la distribución de los porcentajes de efectividad entre los dos tratamientos para R1. La diferencia entre las medianas indica que T.con-patrones tiene mejores valores de efectividad que T.sin-patrones. El valor medio de la efectividad para T.con-patrones es 96.18% con DE=6.62 y 78.93% con DE=22.87 para T.sin-patrones. Esto significa que los tratamientos producen diferentes efectos en la efectividad siendo el T.con-patrones el tratamiento con el que se obtiene un mejor promedio de efectividad ya que supera al T.sin-patrones en un 17.25% (96.18% - 78.93%) de efectividad.

La baja efectividad obtenida con T.sin-patrones puede ser causada por los sujetos representados como valores atípicos que afectan negativamente la efectividad total (puntos 29, 54 y 57 en la gráfica). Dos de ellos alcanzaron el 20% de efectividad (puntos 54 y 57) y el otro alcanzó el 44% de efectividad (punto 29). Con respecto al valor de la mediana y al valor máximo del porcentaje de efectividad, se puede decir que al usar T.sin-patrones, el 50% de los sujetos obtuvieron valores de efectividad que oscilan entre el 80% y el 100% (rango entre mediana y máxima) mientras que al usar T.con-patrones, el mismo porcentaje de sujetos alcanzó exactamente el 100% de efectividad. De los sujetos que usaron el T.sin-patrones, solo unos pocos pudieron conseguir una efectividad del 100%.

Un detalle para considerar en el diagrama de cajas es la “asimetría de las cajas” que se puede ver por la diferencia de áreas entre Q2-Q1 y Q3-Q2. En ambos tratamientos, la distribución de la efectividad es asimétrica, ya que la longitud del rango Q2-Q1 es diferente de la longitud del rango Q3-Q2. Sin embargo, esta asimetría muestra mejores resultados de efectividad cuando se usa T.con-patrones ya que con T.con-patrones el valor de efectividad más frecuente es el 100% y es alcanzada por aproximadamente el 50% de los sujetos, mientras que con T.sin-patrones los valores más frecuentes de efectividad oscilan entre 72.23% y 80% y son alcanzados por aproximadamente el 25% de los sujetos (rango Q1-Mediana).

Se aplicó el método estadístico MLM para buscar diferencias significativas. El $p\text{-valor}=0.000<0.05$ y el tamaño del efecto de 1.02 para R1 indican que hay una diferencia significativa entre los tratamientos y que esta diferencia es grande en magnitud, siendo el T.con-patrones con el que se obtienen mejores resultados de efectividad. Estos resultados concuerdan con el análisis descriptivo.



	Tratamientos	
	T.sin-patrones	T.con-patrones
Media	71.11	97.14
Mediana	66.67	100.00
DE	19.03	7.56
Mínimo	44.44	80.00
Máximo	100.00	100.00
Q1	60.00	100.00
Q3	88.89	100.00

	p-valor (MLM)
Tratamiento	0.025
TE	1.79

Imagen 7.4 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable efectividad en la Réplica 2 (R2)

El diagrama de caja de la Imagen 7.4 compara la efectividad entre los dos tratamientos para R2. La diferencia entre las medianas indica que los sujetos experimentales lograron mejores porcentajes de efectividad cuando usaron T.con-patrones. La comparación de las medias indica que los sujetos que utilizaron T.con-patrones alcanzaron un promedio de efectividad más alto (97.14%, DE=7.56) que los sujetos que usaron T.sin-patrones (71.11%, DE=19.03). En resumen, los sujetos que usaron las interfaces de usuario diseñadas con patrones reportaron 26.03% más de efectividad que los sujetos que usaron las interfaces de usuario convencionales (diseñadas sin patrones).

La distribución de datos del T.sin-patrones indica que el 50% de los sujetos experimentales que utilizaron el tratamiento tuvieron una efectividad de entre el 60% y el 88.89% (Q1 y Q3, respectivamente). De este rango, la mayoría de los sujetos experimentales obtuvieron porcentajes de efectividad entre 60% y 66.67% (Q1 y mediana, respectivamente), mientras que pocos sujetos obtuvieron valores de efectividad entre 66.67% y 88.9% (mediana y Q3, respectivamente). La distribución de datos de T.con-patrones indica que todos los sujetos experimentales alcanzaron el 100% de efectividad, excepto un sujeto (punto 35 en el diagrama de caja de la Imagen 7.4) que alcanzó el 80% de efectividad.

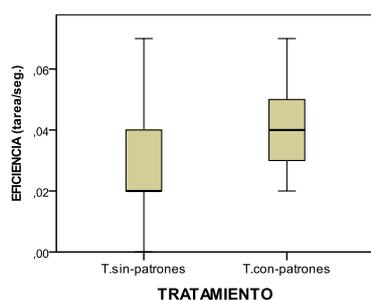
El $p\text{-valor}=0.025<0.05$ de MLM y el tamaño del efecto ($d=1.79$) para la variable de efectividad en R2 indican que existe una diferencia significativa entre los tratamientos y que ésta diferencia es grande en magnitud, siendo T.con-patrones el tratamiento con el que se ha obtenido mejores resultados (un promedio de efectividad del 97.14%).

Sobre la base de los resultados analíticos y descriptivos, concluimos que H_{01} se rechaza tanto para R1 como para R2. Esto significa que, en ambas réplicas, la efectividad cuando se interactúa con las interfaces de usuario diseñadas con los patrones GenomIU es diferente de la efectividad que se observa cuando las interfaces de usuario no incorporan los patrones, siendo las interfaces de usuario diseñadas con patrones aquellas que permiten alcanzar los más altos valores de efectividad. La diferencia de efectividad entre los tratamientos es *grande* en magnitud en ambas réplicas.

Eficiencia

En esta sección, respondemos a la pregunta de investigación PI2 mediante el análisis de H_{02} : *La eficiencia interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la eficiencia observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones.*

Usando la fórmula de eficiencia indicada en la Tabla 15, calculamos el número promedio de tareas por segundo realizadas por cada sujeto experimental usando los tratamientos.



	Tratamientos	
	T.sin-patrones	T.con-patrones
Media	0.03	0.04
Mediana	0.02	0.04
DE	0.02	0.01
Mínimo	0.00	0.02
Máximo	0.07	0.07
Q1	0.02	0.03
Q3	0.04	0.05

	p-valor (MLM)
Tratamiento	0.002
TE	0.92

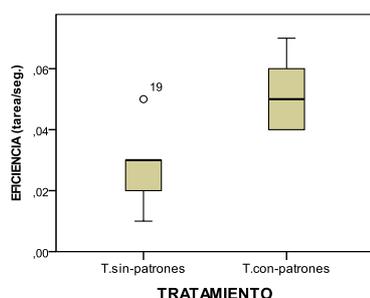
Imagen 7.5 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable eficiencia en la Réplica 1 (R1)

El diagrama de caja y bigotes en la Imagen 7.5 compara la eficiencia (medida en tareas por segundo) obtenida para cada tratamiento. La comparación de medianas muestra que T.con-patrones tiene mejor puntuación (0.04 tareas/seg.) que T.sin-patrones (0.02 tareas/seg.). Esta observación está totalmente respaldada por el valor de las medias donde los sujetos que usan T.con-patrones realizan aproximadamente dos tareas y media por minuto (media=0.04 tareas/seg. =2.4 tareas/min, DE = 0.01), mientras que los sujetos que utilizan T.sin-patrones realizan aproximadamente dos tareas (media=0.03 tareas/seg. = 1.8 tareas/min., DE = 0.02).

T.con-patrones logra puntuaciones más concentradas y de mayor eficiencia que las obtenidas con T.sin-patrones. Al analizar el rango entre los valores máximo y mínimo, podemos decir que la distribución de T.sin-patrones es bastante dispersa con valores de

eficiencia que van desde 0 tareas/min. a 4.2 tareas/min. (mínimo = 0 tareas/segundo y máximo = 0.07 tareas/segundo, respectivamente), mientras que la distribución de resultados de T.con-patrones está más concentrada con valores que van desde 1.2 tareas/min. a 4.2 tareas/min. (mínimo=0.02 tareas/segundo y máximo=0.07 tareas/segundo, respectivamente). Aunque ambos tratamientos alcanzan un valor máximo de 4,2 tareas/min. (máximo=0.07 tareas/seg.), la diferencia entre los tratamientos está en el valor mínimo. Esta diferencia indica que, en el peor de los casos, los sujetos que utilizaron T.con-patrones lograron al menos 1 tarea en un minuto (Mínimo=0.02 tareas/seg. =1.2 tareas/min.), mientras que, en la misma cantidad de tiempo de un minuto, los sujetos que usaron el T.sin-patrones no lograron ninguna tarea (mínimo=0.0 tareas/seg. =0 tareas/min.).

Desde la perspectiva del análisis analítico, el p-valor=0.002<0.05 calculado con el MLM para la variable de eficiencia en R1 significa que existe una diferencia estadísticamente significativa entre los tratamientos y que esta diferencia es *grande* ($d = 0,92$) en magnitud.



	Tratamientos	
	T.sin-patrones	T.con-patrones
Media	0.03	0.05
Mediana	0.03	0.05
DE	0.01	0.01
Mínimo	0.01	0.04
Máximo	0.05	0.07
Q1	0.02	0.04
Q3	0.03	0.06

	p-valor (MLM)
Tratamiento	0.001
TE	1.97

Imagen 7.6 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable eficiencia en la Réplica 2 (R2)

El diagrama de caja y bigotes en la Imagen 7.6 compara la eficiencia obtenida por cada tratamiento para R2. Existe una clara diferencia entre las medianas de los dos tratamientos, siendo T.con-patrones el tratamiento que ha obtenido los mejores resultados de eficiencia. Mientras que los sujetos que utilizaron T.sin-patrones, realizaron un promedio de 2 tareas por minuto (media=0.03 tareas/seg. = 1.8 tareas/min., DE=0.01) aproximadamente, los sujetos que utilizaron T.con-patrones realizaron 3 tareas por minuto (media=0.05 tareas/seg. = 3 tareas/min., DE = 0.01). Considerando el peor de los escenarios (es decir, los valores de eficiencia más bajos), los sujetos que usaron T.sin-patrones lograron al menos 1 tarea por minuto aproximadamente (mínimo=0.01 tareas/seg. =0.6 tareas/min.), mientras que los sujetos usaron T.con-patrones lograron al menos 2 tareas y medio por minuto aproximadamente (mínimo=0.04 tareas/seg. =2.4 tareas/min.). Considerando el mejor de los escenarios (es decir, los valores más altos de eficiencia), se puede decir que los sujetos que utilizaron T.con-patrones lograron aproximadamente cuatro tareas por minuto

(máximo=0,70 tareas/seg. = 4,2 tareas/min.) mientras que con el T.sin-patrones se lograron tres tareas por minuto (máximo = 0.05 tareas/seg. =3 tareas/min.) y que corresponden excepcionalmente al valor atípico 19 en el gráfico.

El p-valor=0.001<0.05 obtenido con el MLM indica que los tratamientos produjeron diferentes efectos y que esta diferencia es grande en magnitud ($d = 1.97$), siendo T.con-patrones el tratamiento con el que se obtuvieron mejores resultados de eficiencia.

En consecuencia, concluimos rechazar H_{02} para R1 y R2. Esto significa que, en ambas réplicas, la eficiencia cuando se interactúa con las interfaces de usuario que incorporan los patrones GenomIU es diferente a la eficiencia observada cuando las interfaces de usuario no incorporan los patrones. Las interfaces que incorporan los patrones permiten al sujeto lograr más tareas por minuto que las que se pueden lograr con las interfaces que no incorporan los patrones (un promedio de aproximadamente cuatro tareas por minuto en comparación con tres tareas por minuto, respectivamente). La diferencia en la eficiencia entre los tratamientos es *grande* en magnitud en ambas réplicas.

Satisfacción

En esta sección, respondemos a la pregunta de investigación PI3 analizando H_{03} : *La satisfacción interactuando con las interfaces de usuario diseñadas con los patrones GenomIU es similar a la satisfacción observada cuando se interactúa con las interfaces de usuario que no están diseñadas con los patrones*. La satisfacción se midió a través del cuestionario CSUQ, como se indica en la Tabla 15.

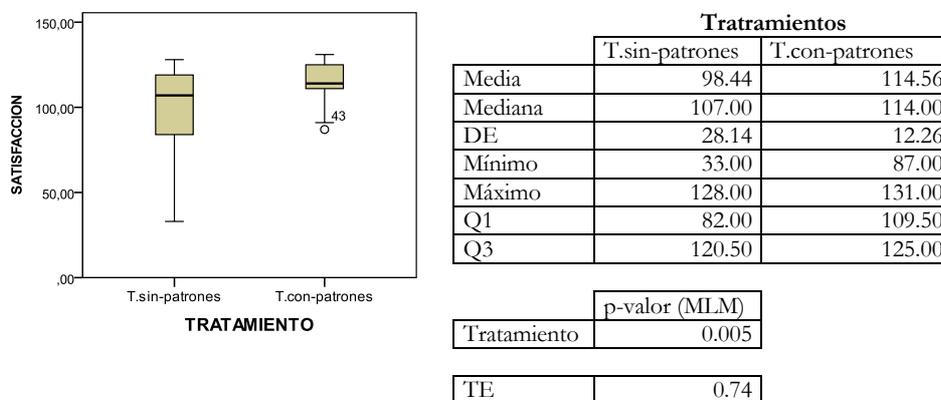


Imagen 7.7 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable satisfacción en la Réplica 1 (R1)

La Imagen 7.7 compara la distribución de datos de satisfacción percibida por los sujetos entre los tratamientos en R1. Existe una ligera diferencia visual entre las medianas. Sin embargo, los puntajes de satisfacción obtenidos con T.sin-patrones son más dispersos que los obtenidos con T.con-patrones, como puede verse al comparar el rango intercuartil (Q3-

Q1) entre los dos tratamientos. El análisis de las medias indica que los sujetos que utilizaron T.con-patrones percibieron una satisfacción promedio más alta (media=114.56, DE =12.26) que los sujetos que usaron T.sin-patrones (media=98.44, DE=28.14). Aunque con ninguno de los tratamientos se alcanza la puntuación de satisfacción máxima (133), los sujetos que usaron el T.con-patrones obtuvieron la puntuación de satisfacción más alta de 131 en comparación con la puntuación de 128 obtenida por los sujetos que usaron el T.sin-patrones. Si consideramos el 25% de los sujetos con los puntajes más bajos de satisfacción (Q1-mínimo) en ambos tratamientos, podemos decir que cualquier puntaje de satisfacción obtenido con el T.con-patrones es superior a cualquier puntaje obtenido con el T.sin-patrones.

El p -valor=0.005<0.05 del método MLM indica que existe una diferencia significativa entre los tratamientos, ratificando así la diferencia observada en la inspección visual. Esta diferencia es media en magnitud ($d = 0.74$), siendo el T.con-patrones es tratamiento con el que los sujetos perciben mayor satisfacción.

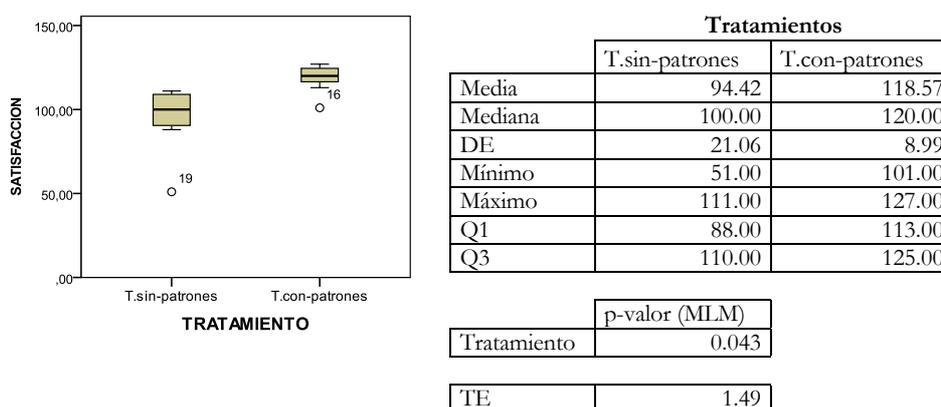


Imagen 7.8 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del modelo lineal mixto (MLM) y el tamaño del efecto para la variable satisfacción en la Réplica 2 (R2)

La Imagen 7.8 compara la distribución de datos de la satisfacción percibida por los sujetos de la R2 usando los dos tratamientos. Aunque con ninguno de los tratamientos los sujetos alcanzan el puntaje máximo de satisfacción (133), la diferencia visual entre las medianas y los valores medios muestra que con el T.con-patrones los sujetos obtuvieron puntajes de satisfacción más altos (media=118.57, DE=8.99) que con el T.sin-patrones (media=94.42, DE=21.06). Los puntajes de satisfacción obtenidos con el T.con-patrones (máximo-mínimo = 127-101 = 26) son menos variables que aquellos obtenidos con el T.sin-patrones (máximo-mínimo= 111-51=60). Por lo tanto, las puntuaciones de satisfacción son más altas y menos variables cuando se usa el T.con-patrones que cuando se usa el T.sin-patrones. El p -valor=0.043<0.05 obtenido del método MLM confirma que existe una diferencia significativa entre los tratamientos y que la magnitud de esta diferencia es grande

($d=1.49$), siendo el T.con-patrones el tratamiento con el que se obtiene mejores resultados de satisfacción.

En consecuencia, concluimos rechazar la H_{03} tanto para R1 como para R2. Esto significa que, la percepción de la satisfacción cuando se interactúa con las interfaces de usuario que incorporan los patrones GenomIU es diferente de la satisfacción percibida cuando se interactúa con interfaces de usuario que no incorporan los patrones, siendo las interfaces que incorporan patrones GenomIU las que producen los puntajes de satisfacción más altos. La diferencia en la satisfacción entre los tratamientos es más pronunciada en R2 que en R1, ya que la magnitud del efecto en R2 es *grande*, mientras que en R1 la magnitud es *media*.

Combinación de réplicas

Además de estudiar por separado el efecto producido por los tratamientos en cada réplica, analizamos los efectos producidos en toda la muestra. Para obtener un único conjunto de observaciones, combinamos las observaciones individuales para cada réplica agregando el perfil como una variable moderadora (Fritz et al., 2017) que contiene dos valores posibles: *científico* y *biomédico*. Esta variable representa las diferencias entre los sujetos experimentales de acuerdo con las preferencias de uso de los sitios web de búsqueda mostrados en la Tabla 17. El perfil científico representa a los sujetos de la R1 y el perfil biomédico representa a los sujetos de la R2.

Con respecto a la **efectividad** (Imagen 7.9a), la diferencia de las medianas entre los perfiles es más evidente en el T.sin-patrones que en el T.con-patrones:

- *En el T.sin-patrones*, la comparación de las medianas entre los perfiles indica que el perfil científico obtuvo un valor de efectividad más alto (80%) que el perfil biomédico (66.67%). El valor medio es 78.93% con $DE=22.87$ para el perfil científico y 71.11% con $DE=19.03$ para el valor biomédico. La comparación visual del tamaño del rango Q3-máximo indica que el perfil científico alcanzó exactamente el 100% de efectividad, mientras que el perfil biomédico varió de 88.89% a 100% de efectividad.
- *En el T.con-patrones*, la mediana de ambos perfiles es 100% y la media es 96.18% con $DE=6.62$ para el perfil científico (Imagen 7.3) y 97.14% con $DE=7.56$ para el perfil biomédico (Imagen 7.4). Estos valores promedio están cerca del 100% de efectividad, y la diferencia entre ellos es leve. Sin embargo, la variabilidad de los valores de efectividad es mayor para el perfil científico que para el perfil biomédico. De hecho, a diferencia del perfil biomédico, los valores de efectividad del perfil científico están más dispersos del valor medio, como se puede ver comparando visualmente el tamaño del rango Q3-Q1 entre los dos perfiles.

Además, el p -valor= $0.000 < 0.05$ correspondiente a la efectividad (Imagen 7.9d) muestra que hay una diferencia significativa entre los tratamientos y la magnitud de esta diferencia se considera *grande* ($d=1.77$), donde el uso del T.con-patrones reportó mejores valores de efectividad. El p -valor= $0.206 > 0.05$ correspondiente a la interacción entre el tratamiento y el perfil no es significativo, lo que indica que el perfil no influye en los efectos producidos por los tratamientos. **Por lo tanto, concluimos rechazar la H_{01} para las réplicas combinadas.**

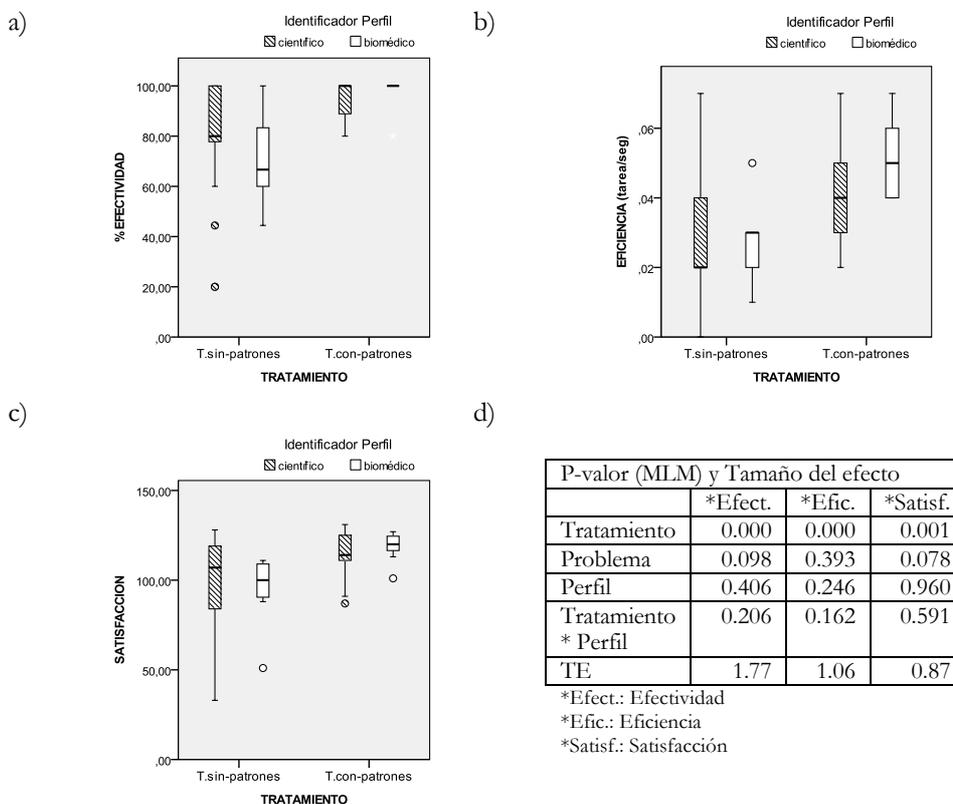


Imagen 7.9 Comparación de las variables (a) efectividad, (b) eficiencia y (c) satisfacción entre los tratamientos aplicados a la combinación de las réplicas considerando al “perfil” como variable moderadora. P-valores resultantes del MLM y tamaño de efecto para cada variable (d).

Con respecto a la **eficiencia** (Imagen 7.9b), el valor de la mediana del perfil biomédico es mayor que la mediana del perfil científico, en los dos tratamientos:

- *En el T.sin-patrones*, la comparación de las medianas entre los perfiles indica que el perfil biomédico obtuvo (0.03 tareas/seg., ver Imagen 7.6) un valor de eficiencia más alto que el perfil científico (0.02 tareas/seg., ver Imagen 7.5). El valor de la media es el mismo para ambos perfiles (0.03 tareas/seg.), pero con una ligera diferencia en el valor de la DE (0.02 para R1 y 0.01 para R2). Por otro lado, comparando el rango entre Q3 y el máximo, podemos decir que el perfil biomédico puede alcanzar un máximo de 1.8 tareas/min. (0.03 tareas/seg., ver Imagen 7.6), mientras que el perfil científico puede superar este valor, pudiendo alcanzar un valor de eficiencia máxima de 4.2 tareas/min. (0.07 tareas/seg., ver Imagen 7.5).
- *En el T.con-patrones*, la comparación de medianas indica que los valores de eficiencia para el perfil biomédico están por encima de 3 tareas/minuto (mediana=0.05 tareas/seg., ver Imagen 7.6), mientras que los valores del perfil científico están por

encima de 2.4 tareas/minuto (mediana=0.04 tareas/seg., ver Imagen 7.5). La diferencia en el tamaño del rango Q3-Q1 entre los perfiles muestra que la distribución de los datos del perfil biomédico está más concentrada que la distribución de datos del perfil científico.

Además, el p -valor=0.000<0.05 resultante del MLM aplicado a la variable de eficiencia (Imagen 7.9d) indica que existe una diferencia significativa entre los tratamientos y además la magnitud de esta diferencia es *grande* ($d=1.06$). De esta diferencia se sabe que por el análisis descriptivo se sabe que, los mejores resultados de eficiencia se consiguen con el T.con-patrones. El p -valor=0.162>0.05 correspondiente a la interacción entre el tratamiento y el perfil no es significativo, lo que indica que el perfil no influye en los efectos producidos por los tratamientos. **Por lo tanto, concluimos rechazar la H₀₂ para las réplicas combinadas.**

Con respecto a la **satisfacción** (Imagen 7.9c), la comparación de medianas en el T.sin-patrones reportó que el perfil científico obtuvo mejores puntuaciones de satisfacción que el perfil biomédico. Contrariamente, en el T.con-patrones, las puntuaciones de satisfacción del perfil científico fueron más bajas que el perfil biomédico. Además, el p -valor=0.001<0.05 resultante del MLM aplicado a la variable de satisfacción (Imagen 7.9d) significa que existe una diferencia significativa entre los tratamientos y que el tamaño de esta diferencia es *grande* ($d=0.87$), siendo el T.con-patrones el tratamiento con el que se obtienen mejores puntuaciones de satisfacción. El p -valor=0.551>0.05 correspondiente a la interacción entre el tratamiento y el perfil no es significativo, lo que indica que el perfil no influye en los efectos producidos por los tratamientos. **Por lo tanto, concluimos rechazar la H₀₃ para las réplicas combinadas.**

Discusión

El principal hallazgo de este estudio empírico es que: las interfaces de usuario diseñadas con los patrones GenomIU producen mejores puntajes de efectividad, eficiencia y satisfacción que las interfaces de usuario convencionales que no incorporan los patrones. En términos generales, los patrones propuestos proporcionan mejoras en la usabilidad, permitiendo al usuario lograr las tareas propuestas en menos tiempo y de una manera fácil, natural y simple. El significado de estos hallazgos se explica como respuesta de cada pregunta de investigación:

PI1: ¿Es la efectividad afectada por el uso de los patrones GenomIU en las interfaces de usuario? Los resultados del análisis de datos de ambas réplicas (R1 y R2) indicaron que las interfaces de usuario diseñadas con los patrones GenomIU alcanzaron mayor efectividad que las interfaces diseñadas sin patrones. Desde la perspectiva del perfil de los sujetos, es importante destacar que los sujetos con el perfil científico obtuvieron las mejores puntuaciones de efectividad. Esto puede explicarse por el hecho de que los sujetos de este perfil están más familiarizados con la búsqueda y recuperación de información debido a su capacitación profesional en documentación y gestión de bibliotecas.

PI2: ¿Es la eficiencia afectada por el uso de los patrones GenomIU en las interfaces de usuario? Los resultados del análisis de datos de ambas réplicas indican que el uso de los patrones GenomIU afecta la eficiencia, lo que permite a los usuarios realizar más tareas por minuto. Una justificación plausible para este resultado es que los patrones

aceleran y enriquecen la forma en que los usuarios visualizan y navegan a través de los datos, reemplazando la representación tradicional tabular, plana y desconectada de datos por una representación más dinámica y conectada.

PI3: ¿Es la satisfacción del usuario afectada por el uso de los patrones GenomIU en las interfaces de usuario? Los hallazgos indican que los usuarios que usaron las interfaces de usuario diseñadas con los patrones GenomIU obtuvieron puntuaciones de satisfacción más altas. La satisfacción percibida puede ser corroborada por comentarios de los usuarios como: *“la idea de tener toda la información en una única interfaz de usuario con contenido conectado y recomendaciones me ayudó a realizar las tareas”*. Sin embargo, un pequeño número de sujetos indicó que la sobrecarga de información hacía difícil identificar las respuestas correctas. Esta situación nos ha llevado a pensar que puede ser necesario diferenciar entre usuarios novatos y expertos y diseñar prototipos de interfaz para cada tipo de usuario.

7.2.4 Resumen

En este estudio empírico, valoramos el impacto causado por la aplicabilidad de los patrones GenomIU. El escenario que escogimos para realizar la valoración es la búsqueda de literatura biomédica. Para valorar el impacto, evaluamos la usabilidad midiendo la efectividad, eficiencia y satisfacción percibida por los usuarios cuando interactúan con interfaces de usuario de búsqueda de bibliografía diseñadas con los patrones GenomIU. Consideramos el impacto resultante como el resultado causado por el conjunto de patrones en lugar del resultado aislado causado por patrones individuales.

Con la intención de comprobar que el efecto producido por los patrones no está influenciado por el conocimiento del dominio genómico de los sujetos experimentales, realizamos el estudio considerando sujetos experimentales de un dominio diferente al genómico pero que comparten en común el escenario general de búsqueda de literatura: sujetos experimentales enfocados a buscar literatura científica. De esta manera, el estudio consistió en llevar a cabo un experimento, diseñado como medidas repetidas, considerando dos réplicas, una con 24 sujetos expertos experimentales en la búsqueda de literatura científica y otra con ocho sujetos expertos experimentales en la búsqueda de literatura biomédica. Los resultados de los experimentos nos permitieron probar tres hipótesis, una para cada medida de usabilidad, que se analizó utilizando el método estadístico de Modelo Lineal Mixto. Para cada hipótesis, evaluamos si las interfaces de usuario diseñadas con o sin patrones GenomIU producen efectos similares o diferentes en la medida de usabilidad.

Nuestros resultados del estudio empírico sugieren que las interfaces de usuario diseñadas con patrones GenomIU ayudan a los usuarios a lograr mejores resultados en efectividad, eficiencia y satisfacción. Los hallazgos relevantes de este estudio se pueden resumir de la siguiente manera:

- a) Los patrones GenomIU mejoran la usabilidad de las interfaces de usuario de búsqueda de literatura biomédica.
- b) Los resultados obtenidos en cada réplica y la combinación de estas indican que los efectos producidos por las interfaces de usuario que incorporan los patrones GenomIU no dependen del conocimiento y experiencia del sujeto.

c) El impacto en la efectividad, la eficiencia y la satisfacción producidos por el conjunto de patrones se puede traducir como usuarios más satisfechos que pueden completar más tareas en menos tiempo.

7.3 VALIDACIÓN DEL PROCESO DE DISEÑO

7.3.1 Objetivo y Preguntas de Investigación

El objetivo de este estudio empírico es analizar cualitativamente la percepción de desarrolladores de interfaces de usuario para conocer en qué medida ellos aceptarían o rechazarían el proceso de diseño del método GenomIUm. La aceptación de un artefacto software puede ser entendida como la intención del usuario en usar el artefacto y esta “intención de uso” puede ser explicada subjetivamente en términos de “utilidad percibida” y “facilidad de uso” (Davis et al., 1989). De acuerdo con la plantilla Meta/Pregunta/Medida (GQM, por sus siglas en inglés) (Van Solingen, 2002), el objetivo de este estudio es:

Analizar	el método GenomIUm
Con el propósito de	evaluar la aceptación de usar el método
Con respecto a la	percepción de facilidad de uso, utilidad percibida e intención de uso
Desde el punto de vista del	investigador
En el contexto de	desarrolladores/diseñadores de software diseñando interfaces de usuario

Las preguntas de investigación (PI) derivadas del objetivo y sus correspondientes hipótesis son:

- **PI 1:** ¿El método GenomIUm es percibido como útil para el diseño de interfaces de usuario? La hipótesis por evaluar es **H₀₁**: El método GenomIUm es percibido como útil en un nivel medio.
- **PI 2:** ¿El método GenomIUm es percibido como fácil de usar para el diseño de interfaces de usuario? La hipótesis por evaluar es **H₀₂**: El método GenomIUm es percibido como fácil de usar en un nivel medio.
- **PI 3:** ¿Existe intención de usar el método GenomIUm en futuros diseños de interfaces? La hipótesis por evaluar es **H₀₃**: Existe la intención, a un nivel medio, de usar el método GenomIUm en el futuro
- **PI 4:** ¿La utilidad percibida está influenciada por la facilidad de uso? La hipótesis por evaluar es **H₀₄**: La utilidad percibida es afectada por la facilidad de uso.
- **PI 5:** ¿La intención de uso está influenciada por la facilidad de uso? La hipótesis por evaluar es **H₀₅**: La intención de uso es afectada por la facilidad de uso.
- **PI 6:** ¿La intención de uso está influenciada por la utilidad percibida? La hipótesis por evaluar es **H₀₆**: La intención de uso es afectada por la utilidad percibida.

7.3.2 Planeamiento

Sujetos experimentales

Los participantes de este estudio fueron 16 desarrolladores de software con experiencia de 1 a 4 años en el diseño y desarrollo de interfaces de usuario en proyectos de software realizados tanto en la academia como en la industria. De los 16 desarrolladores, 12 fueron de Ecuador y cuatro de España. El método que utilizan los participantes para desarrollar interfaces de usuario en sus proyectos consiste en crear la interfaz de usuario al mismo tiempo que se codifica el sistema. A este método lo denominaremos “método tradicional” para efectos de este estudio.

Factores y variables independientes

El experimento estudia un factor con un tratamiento. El factor es el “proceso de diseño conceptual de la interfaz de usuario” y el tratamiento es el proceso de diseño del método GenomIUm (es decir, desarrolladores aplicando el método GenomIUm con el objetivo de diseñar interfaces de usuario).

No existe un proceso estándar para diseñar el concepto de la interfaz de usuario; por lo tanto, no podemos evaluar el proceso del método GenomIUm contra un proceso control. Por lo tanto, la variable independiente tendrá un único valor en una escala nominal. Esta consideración está basada en el trabajo de Abrahão et. al. (Abrahão et al., 2011), donde los autores evalúan su método propuesto verificando si los puntajes obtenidos por el uso de su método son mejores que el puntaje neutral (es decir, el puntaje de 3) de la escala tipo Likert de 5 puntos.

Cabe mencionar que el objetivo de este estudio es medir la percepción de los desarrolladores después de usar el proceso de diseño del método GenomIUm y al no existir un proceso control contra el cual evaluar, consideramos que el valor neutral de la escala tipo Likert es adecuado para nuestra variable independiente.

Variables dependientes

Las variables dependientes, asociadas a cada PI, son la *percepción de uso*, la *utilidad percibida* y la *intención de uso*. El instrumento utilizado para medir estas variables de percepción es una adaptación del Modelo de Aceptación de Tecnología (TAM, por sus siglas en inglés) (Davis, 1989). El TAM ha sido ampliamente utilizado bajo diferentes situaciones, factores de control y sujetos (Lee et al., 2003) con el fin de evaluar y predecir la aceptación de uso de un artefacto software donde la intención de uso y comportamiento de los usuarios son afectados por la facilidad de uso y utilidad percibida por el usuario. De acuerdo con el TAM, las variables dependientes de este estudio se definen como:

- **Utilidad percibida (UP)**, se refiere al grado en el cual el desarrollador considera que el método GenomIUm mejorará el rendimiento de su trabajo.
- **Facilidad de uso percibida (FUP)**, se refiere al grado en el cual el desarrollador considera que usando el método GenomIUm está libre de esfuerzo.
- **Intención de uso (IDU)**, que se refiere al grado de predisposición de los desarrolladores en usar el método GenomIUm.

Para operacionalizar la evaluación de las variables de percepción, el instrumento especifica un conjunto de 16 ítems. La Tabla 22 muestra los ítems asignados para cada variable y que han sido reformulados para adaptarse a la necesidad de nuestro experimento: 8 para UP, 6 para FUP y 2 para IDU. En total 16 ítems. Cada ítem es evaluado mediante una escala de Likert de 5 puntos, que va desde 1= “Nada probable” a 5= “Completamente probable”. La Tabla 23 detalla las variables de percepción, su métrica y su escala.

Tabla 22 Ítems para medir las variables de percepción: utilidad percibida (UP), facilidad de uso percibida (FUP) e intención de uso (IDU)

Ítem	Descripción
UP1	Creo que el proceso de diseño de GenomIUm reduciría el esfuerzo requerido para diseñar interfaces
UP2	Interfaces de usuario diseñadas con este proceso de diseño serían fáciles de entender para los usuarios.
UP3	Este proceso de diseño facilitaría a los usuarios verificar si las interfaces de usuario resultantes son usables
UP4	En general, encontré el proceso de diseño de GenomIUm útil
UP5	El uso de este proceso de diseño facilitaría el diseño de interfaces de usuario
UP6	En general, creo que este proceso de diseño proporciona una solución efectiva al problema de diseñar interfaces de usuario
UP7	En general, creo que este proceso de diseño es una mejora al desarrollar interfaces
UP8	El uso de este proceso de diseño facilitaría el diseño de interfaces de usuario usables
FUP1	Me resultó fácil de seguir el procedimiento para aplicar el proceso de diseño de GenomIUm
FUP2	En general, me resultó fácil de usar el proceso de diseño de GenomIUm
FUP3	Me resultó fácil aprender el proceso de diseño de GenomIUm
FUP4	Me resultó fácil aplicar el proceso de diseño de GenomIUm al ejemplo de diseño de interfaces
FUP5	Las reglas del proceso de diseño de GenomIUm me resultaron claras y fáciles de entender
FUP6	Confío en que ahora soy capaz de aplicar el proceso de diseño de GenomIUm en la práctica
IDU1	Definitivamente usaría este proceso de diseño para diseñar interfaces de usuario
IDU2	Tengo la intención de usar este proceso de diseño si tengo que diseñar interfaces en el futuro

Tabla 23 Resumen de variables dependientes, sus métricas y su escala de medida

Nombre	Métrica	Escala
Facilidad de uso percibida (FUP)	Calculado como el promedio de los ítems FUP1 a FUP8 obtenidos de la encuesta conformada por los ítems mostrados en Tabla 22.	Proporción
Utilidad percibida (UP)	Calculado como el promedio de los ítems UP1 a UP7 obtenidos de la encuesta conformada por los ítems mostrados en Tabla 22.	Proporción
Intención de Uso (IDU)	Calculado como el promedio de los ítems FUP1 a IDU2 obtenidos de la encuesta conformada por los ítems mostrados en Tabla 22.	Proporción

Para efectos de aplicación del cuestionario, los ítems sufrieron cambios en narración y estructura: 1) La mitad de los ítems fueron narrados en sentido negativo con el fin de

mantener alerta al encuestado en cada ítem (Hu et al., 1999). Por ejemplo, el ítem UP5 fue redactado en el cuestionario como: “El uso de este método **dificultaría** el diseño de interfaces de usuario”. Es importante considerar que, para el análisis de datos, las respuestas de los ítems negativos deben ser invertidas con el fin de que todos los ítems apunten en el mismo sentido; 2) Los ítems fueron organizados de manera aleatoria para reducir el “efecto techo y suelo” que puede inducir a respuestas monótonas a ítems que miden lo mismo (Hu, 1999).

Objetos experimentales y diseño del experimento

Para observar la percepción de los desarrolladores con relación al proceso de diseño, planteamos un problema de diseño de interfaz de usuario que consiste en diseñar una interfaz de usuario para analizar información genómica relacionada con una enfermedad genética. Para lograr este objetivo, solicitamos al desarrollador las siguientes tareas de diseño que han sido descritas como requisitos:

- Visualizar las relaciones de citación entre los reportes clínicos. Un determinado artículo, por ejemplo, puede haber sido citado por 15 artículos del conjunto de resultados.
- Filtrar los reportes clínicos por un determinado año de publicación o un rango de años y por género.
- Visualizar las tendencias en publicaciones a través del tiempo.

Considerando que el experimento tiene un factor y un tratamiento, el diseño del experimento es simple: cada sujeto resuelve el problema de diseño planteado con el método GenomIUm.

Procedimiento

El procedimiento del experimento consiste en 4 pasos:

Paso 1.- Encuesta preliminar. - Los sujetos experimentales completan una encuesta demográfica que recoge información personal, formación y experiencia profesional y experiencia con el diseño de interfaces de usuario. Este cuestionario fue implementado como un formulario de Google que es accesible en línea. El cuestionario es totalmente anónimo (no hay manera de identificar al usuario que registró la información).

Paso 2.- Presentación y entrenamiento. - Los sujetos experimentales reciben un documento digital que explica el propósito y uso del proceso de diseño del método GenomIUm y presenta un ejemplo de aplicación.

Paso 3.- Diseño de interfaz de usuario. - Los sujetos experimentales reciben el documento de instrucciones que contiene el problema a resolver y un enlace de la página web con el catálogo de patrones. El sujeto experimental aplica el proceso de diseño y los patrones para crear un prototipo de interfaz utilizando papel y lápiz.

Paso 4.- Cuestionario de percepción basado en TAM. - Los sujetos experimentales responden el cuestionario de 16 ítems mostrados en la Tabla 22 que recoge su percepción respecto al uso del proceso para diseñar la interfaz de usuario.

Evaluación de amenazas

En este apartado valoramos las amenazas que pueden afectar a la validez del experimento y describimos la manera en la que las hemos afrontado y/o evitado.

Validez de Conclusión:

Supuestos vulnerados de pruebas estadísticas. - Esta amenaza se refiere a las condiciones previas que deben cumplirse antes de aplicar una prueba estadística. Antes de aplicar las pruebas estadísticas se comprobó la normalidad aplicando la prueba W de Shapiro-Wilk. En el caso del método estadístico de Regresión Lineal aplicado, se verificó la normalidad del residuo. Todos los residuos tuvieron una distribución normal.

Fiabilidad de las medidas. - Esta amenaza está asociada a la confiabilidad de las medidas. Las medidas objetivas son más confiables que las subjetivas, ya que no hay intervención del juicio humano. Las variables del experimento son cualitativas y para minimizar la amenaza de fiabilidad de las medidas hemos utilizado una escala Likert de 5 puntos. Los números de la escala de Likert, correspondientes a las respuestas obtenidas han sido sumados para convertir las respuestas cualitativas en respuestas cuantitativas que pueden ser computadas.

Heterogeneidad aleatoria de los sujetos. - Esta amenaza se refiere al riesgo de tener un grupo de sujetos que sea “muy” heterogéneo o “muy” homogéneo. En el experimento, todos los participantes son desarrolladores de interfaces de usuario convirtiendo al conjunto de participantes en una muestra homogénea, sin embargo, la diferencia en años de experiencia en el desarrollo de interfaces de usuario entre los participantes hace que la muestra de sujetos sea heterogénea.

Validez interna:

Maduración. - Se refiere a la posibilidad de que el sujeto se comporte diferente a medida que pasa el tiempo ya sea por cansancio, aburrimiento u otro factor. Para minimizar esta amenaza, el experimento se diseñó con un problema de tres tareas de resolución corta y de fácil resolución considerando la experiencia de los participantes. Para evitar distracciones durante la prueba, solicitamos a los sujetos que asignen un tiempo específico para la prueba, así como también las condiciones adecuadas para llevarla a cabo.

Instrumentación. - Significa que los objetos experimentales (formas, cuestionarios, prototipos de UI) que no están diseñados adecuadamente pueden afectar negativamente el experimento. Para evitar esta amenaza, el cuestionario de percepción TAM y las tareas propuestas fueron previamente evaluadas por un investigador independiente del grupo de experimentadores. Además, previo a la prueba se explicó a los participantes las tareas y cuestionarios que deben ser realizados en el experimento con el fin de que no existan ambigüedades ni diversidad de interpretaciones.

Validez del constructo:

Sesgo mono-operatorio. - Esta amenaza significa que cuando se usa una única variable independiente, sujeto o tratamiento, el experimento puede representar de forma insuficiente o subestimar el constructo y, por lo tanto, no ofrecer una imagen completa de la teoría. Evitamos esta amenaza utilizando el valor neutral (3) de la escala de Likert (configurado en la herramienta SPSS) para evaluar las hipótesis como una alternativa válida en el caso de no contar con un tratamiento control (Abrahão, 2011).

Suposición de hipótesis. - Esta amenaza ocurre cuando los sujetos experimentales determinan el propósito y el resultado esperado del experimento. Esta amenaza se ha

minimizado realizando el experimento sin explicar a los sujetos experimentales los objetivos específicos del experimento ni las variables de respuesta que pretendíamos medir.

Temor a la evaluación. - Esta amenaza ocurre cuando los sujetos experimentales tienen miedo de ser evaluados y tratan de estar más atentos y enfocados en la evaluación afectando así el resultado del experimento. Para evitar esta amenaza, antes de comenzar el experimento, explicamos a los sujetos experimentales que la interfaz de usuario diseñada como parte del experimento no es juzgada como “buen diseño” o “mal diseño” sino considerada como un producto final del proceso.

Validez externa:

Interacción de selección y tratamiento. - Esta amenaza se refiere al efecto de no tener una muestra representativa de la población objetivo que queremos generalizar. Esta amenaza fue evitada asegurándonos que los sujetos experimentales sean desarrolladores con experiencia en el diseño de interfaces de usuario y provenientes de diferente localización geográfica (España y Ecuador).

7.3.3 Análisis de datos e Interpretación de Resultados

El análisis de los datos fue realizado utilizando el software estadístico IBM SPSS Statistics versión 2.0. El análisis está estructurado como sigue:

Paso 1: Confiabilidad de la encuesta. – Desde que cada variable de percepción es medida por las respuestas de sus ítems, es necesario evaluar la confiabilidad y consistencia interna de los ítems que conforman la encuesta. Para esto, aplicamos el coeficiente Alfa de Cronbach obteniendo todas las variables (UP= 0.86, FUP=0.83, IDU=0.82) un valor superior a 0.7, un puntaje considerado como un umbral común de confiabilidad (Nunnally et al., 1967). Por lo tanto, estos resultados nos indican que los ítems son medidas confiables y válidas de las variables basadas en la percepción.

Paso 2: Evaluación de hipótesis H_{01} , H_{02} y H_{03} .- La evaluación de estas hipótesis sigue la siguiente estructura:

- Estudio descriptivo. – Estudiamos las variables en función de sus estadísticos descriptivos tales como: media, mediana, mínimo, máximo, desviación estándar.
- Estudio analítico. – Determinamos la significancia de las respuestas obtenidas en el análisis descriptivo utilizando el método estadístico “Wilcoxon de los rangos con signo de una muestra”. Antes de aplicar los métodos estadísticos, evaluamos la normalidad de cada muestra.

Paso 3: Evaluación de hipótesis H_{04} , H_{05} y H_{06} .- Estas hipótesis plantean una relación de causalidad entre las variables: a) UP vs FUP, b) IDU vs FUP, c) IDU vs UP. Por lo tanto, el planteamiento requiere de:

- Evaluación de relación entre variables. – Aplicamos “Regresión Lineal” como método estadístico para verificar si existe relación entre los pares de variables evaluando el coeficiente de correlación de Pearson (**R**). De acuerdo al valor del coeficiente, podremos determinar el grado de correlación: *relación no significativa*, si R está entre 0 y 0.20; *correlación baja*, si R está entre 0.21 y 0.40; *correlación significativa*, si R está entre 0.41 y 0.70; *alto grado de correlación*, si R está entre 0.71 y 1. Además el

valor de R cuadrado nos indicará la medida en la que una variable es explicada por la otra.

- Modelo de regresión lineal. – Si se evidencia correlación entre las variables, se construye un modelo predictivo de la forma $y = ax + b$.

Estudio descriptivo

El estudio descriptivo (Imagen 7.10) indica que los valores medios de la UP (media=4.10), FUP (media:4.25) e IDU (media: 3.87) superan el puntaje neutral de 3, el punto medio de la escala de Likert. Estos resultados indican que los desarrolladores percibieron a GenomIUm como un método útil y fácil de usar y que tendrán la intención de usar el método en futuros proyectos de diseño de interfaces de usuario.

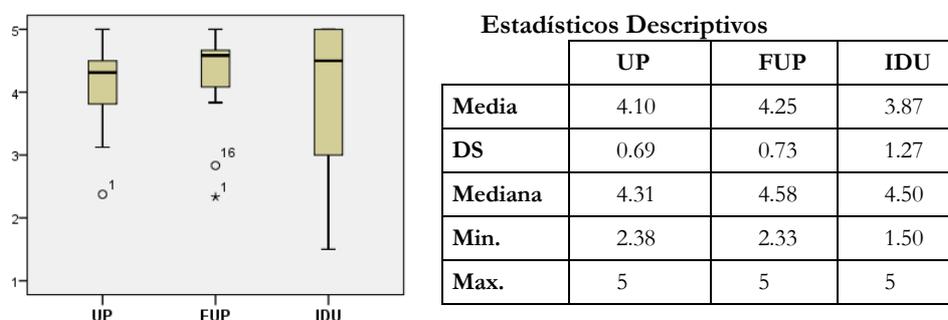


Imagen 7.10 Diagramas de Caja y Bigotes y estadísticos descriptivos para las variables de Utilidad Percibida (UP), Facilidad de Uso (FUP) e Intención de Uso (IDU)

Para corroborar las observaciones del estudio descriptivo, realizamos el estudio analítico a cada variable aplicando un método estadístico que nos permita verificar si las respuestas de los participantes son significativamente diferentes del valor neutral (3) de la escala de Likert.

Estudio analítico

El criterio para aplicar uno u otro método estadístico está sujeto a sus supuestos, especialmente a la prueba de normalidad. Los resultados de la prueba de normalidad de Shapiro Wilk indicaron que las muestras de UP, FUP e IDU no siguen una distribución normal (p -valor=0.05, p -valor=0.073 y p -valor=0.073, respectivamente). Por lo tanto, para confirmar analíticamente las observaciones del análisis descriptivo es necesario aplicar métodos estadísticos no paramétricos.

La Tabla 24 muestra el resultado de la prueba no paramétrica de “Wilcoxon de los rangos con signo de una muestra”, configurada con el valor hipotetizado igual a 3 (el valor neutral). La prueba evalúa tres hipótesis, una por cada variable de percepción, expresadas como “El valor de la variable es igual a 3”, siendo 3 el valor neutral y correspondiente al punto medio de la escala de Likert de 5 puntos utilizada.

Tabla 24 Resultados de la prueba Wilcoxon de los rangos con signo de una muestra, respecto al valor hipotetizado (3).

Hipótesis nula	p-valor	Decisión
La UP es igual a 3	0.001	Rechazar la hipótesis nula
La FUP es igual a 3	0.001	Rechazar la hipótesis nula
La IDU es igual a 3	0.001	Rechazar la hipótesis nula

A partir del estudio descriptivo y analítico, la decisión respecto a las hipótesis H_{01} , H_{02} y H_{03} es la siguiente:

La H_{01} es rechazada. Es decir que, el método GenomIUm es percibido por los desarrolladores como útil en un nivel superior al nivel medio. El p-valor=0.001 correspondiente a UP (Tabla 24) y su media=4.10 (Imagen 7.10) indican que el valor de utilidad percibida no solo es diferente al valor neutral de la escala de Likert, sino que es mayor a este.

La H_{02} es rechazada. Es decir que, el método GenomIUm es percibido como fácil de usar en un nivel superior al nivel medio. El p-valor=0.001 correspondiente a FUP (Tabla 24) y su media=4.25 (Imagen 7.10) indican que el valor de la facilidad de uso percibida no solo es diferente al valor neutral de la escala de Likert, sino que es mayor a este.

La H_{03} es rechazada. Esto quiere decir que, los desarrolladores tienen la intención de usar el método GenomIUm en futuros proyectos de desarrollo de interfaces de usuario y el nivel de esta intención es superior al nivel medio. El p-valor=0.001 correspondiente a IDU (Tabla 24) y su media=3.87 (Imagen 7.10) indican que el valor de la intención de uso no solo es diferente al valor neutral de la escala de Likert, sino que es mayor a este.

Evaluación de relación entre variables

Con respecto a las hipótesis H_{04} , H_{05} y H_{06} , la Tabla 25 muestra los resultados obtenidos con el método de regresión lineal que nos permite rechazar o aceptar las hipótesis.

Tabla 25 Resultados del modelo de regresión lineal simple

Relación	Modelo de regresión	T	p-valor	R	R cuadrado
UP vs FUP	Constante	1.21	0.246	0.829	0.688
	Coefficiente de correlación	5.52	<0.001		
IDU vs FUP	Constante	-0.485	0.635	0.624	0.390
	Coefficiente de correlación	2.991	<0.010		
IDU vs UP	Constante	-1.743	0.103	0.799	0.639
	Coefficiente de correlación	4.975	<0.000		

A partir de esta información podemos concluir que:

La H_{04} es aceptada. Para la relación UP vs FUP, el valor de R (0.829) indica un *alto grado de correlación* entre las variables y esto está corroborado por el valor de R cuadrado que indica que el 68% de la UP está explicada por la FUP.

La **H₀₅ es aceptada**. Para la relación IDU vs FUP, el valor de R (0.624) indica una *correlación significativa* entre las variables y esto está corroborado por el valor de R cuadrado que indica que el 63% de la IDU está explicada por la FUP.

La **H₀₆ es aceptada**. Para la relación IDU vs UP, el valor de R (0.799) indica un *alto grado de correlación* entre las variables y esto está corroborado por el valor de R cuadrado que indica que el 64% de la IDU está explicada por la UP.

Los valores de p -valor <0.05 de los coeficientes de correlación en cada relación de la Tabla 25 nos indican que es posible construir un modelo, de esta manera, utilizamos la constante y el coeficiente de correlación de cada relación para formar los modelos predictivos como se presentan a continuación:

Para UP vs FUP: $y = 5.52x + 1.21$.

Para IDU vs FUP: $y = 2.99x - 0.485$

Para IDU vs UP: $y = 4.975x - 1.743$.

7.3.4 Resumen

Este estudio empírico valoró cualitativamente la percepción de los desarrolladores/diseñadores de interfaces respecto al uso del proceso de diseño del método GenomIUm. La percepción de los desarrolladores fue valorada en tres variables principales definidas por el método TAM: utilidad percibida, facilidad de uso percibida e intención de uso percibida.

El estudio consistió en un experimento donde desarrolladores de software diseñaron interfaces de usuario siguiendo los pasos del proceso y luego evaluaron cualitativamente la utilidad, facilidad de uso del proceso y la intención de usar el proceso de diseño del método GenomIUm en futuros proyectos.

Los resultados de este estudio indicaron que:

- Los desarrolladores de software tienen intención de usar el proceso del método en futuros proyectos de diseño de interfaces de usuario.
- La intención descrita en el punto 1 está respaldada por el hecho de que los desarrolladores consideran que el proceso de diseño es útil y fácil de usar para el diseño de interfaces de usuario en el dominio genómico.

7.4 CASO DE ESTUDIO: GENOMA DE LOS CÍTRICOS

7.4.1 Objetivo

Para validar el método GenomIUm en un ambiente real, realizamos un caso de estudio dirigido al análisis del genoma de las frutas cítricas. El objetivo de este caso de estudio es primero aplicar el método para diseñar interfaces web de un proyecto de acceso a datos del genoma de los cítricos y entonces valorar los comentarios realizados por los usuarios finales respecto al proceso de diseño y el diseño de la interfaz de usuario final.

7.4.2 Dominio y participantes

El dominio de aplicación de este estudio es el campo de la investigación agrícola donde los investigadores (genetistas, biólogos, bioinformáticos) estudian el genoma de diferentes especies de cítricos (p. ej., mandarina, limón, pomelo) en sus diferentes variedades (p. ej., *clemensum* o *arrufatina*, son variedades de la mandarina) para priorizar las variaciones genéticas que determinan cambios en sus características biológicas (p. ej., color, sabor, aroma). Las variaciones genéticas resultantes del estudio se convierten en información valiosa para apoyar a programas de mejoramiento genético de las especies cultivadas. El caso de estudio fue llevado a cabo con investigadores del Centro de Genética del Instituto Valenciano de Investigaciones Agrarias (IVIA), quienes tienen amplia experiencia investigadora en las características agronómicas y estructura genómica de variedades de cítricos. En este contexto, aplicamos el método GenomUIm para diseñar interfaces de usuario web que faciliten a los investigadores del IVIA a priorizar las variaciones genéticas.

La idea general detrás del proceso de priorización de variaciones en el genoma de los cítricos es i) formar dos grupos de variedades de cítricos, a partir de los cuales se desea estudiar una determinada característica biológica (p. ej., acidez), ii) comparar los grupos para detectar qué variaciones genéticas están presentes en un grupo de variedades y no en el otro y, finalmente, iii) analizar detenidamente las variaciones genéticas resultantes con el fin de identificar un conjunto de variaciones relevantes y fuertemente relacionadas con la característica biológica bajo estudio.

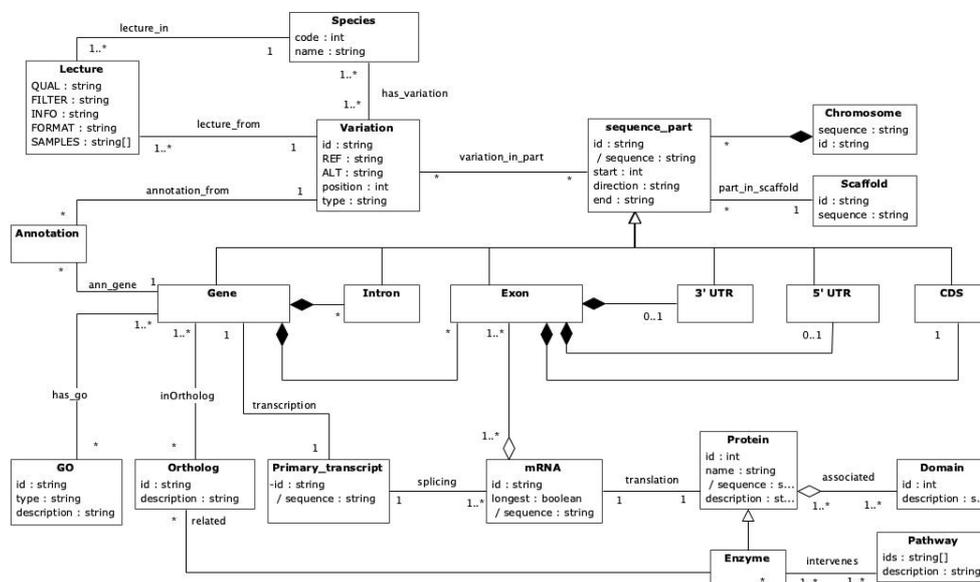


Imagen 7.11 Modelo conceptual del genoma de los cítricos. Tomado y adaptado de (García et al., 2019)

Para entender en detalle la idea general, es necesario especificar los conceptos principales inmersos en el dominio. El *modelo conceptual del genoma de cítricos* de la Imagen 7.11 reúne todos los conceptos y relaciones inmersos en el proceso de priorización de variaciones genéticas en el genoma de los cítricos. Este modelo fue realizado por García et al. (García, 2019) como una investigación del grupo de investigación PROS y paralela a esta Tesis. Nosotros tomamos este modelo como base para entender el dominio. Este modelo puede ser visto como una *instancia* del modelo conceptual del genoma mencionado en la Sección 3.1 ya que, además de incluir conceptos principales del modelo del genoma (clases Chromosome, Species, Variation, Gene, Protein), incluye otros conceptos necesarios para entender el dominio particular como, por ejemplo, scaffold (clase “Scaffold”), lectura realizada en la secuenciación (clase “Lecture”), información obtenida de Gene Ontology acerca de la función de los genes y sus productos (clase “GO”), grupos ortólogos (clase “Ortholog”). Estos y otros conceptos pueden ser revisados por el lector en (García, 2019).

Además de entender los conceptos en el dominio, especificamos las tareas inmersas en el proceso de priorización. La Imagen 7.12 muestra el *modelo de tareas*, expresado en notación CTT (Paternò, 2003), que detalla las tareas inmersas en el proceso de priorización y que fueron especificadas a partir de reuniones realizadas con los investigadores del IVIA. La primera imagen presenta el modelo de tareas completo; la segunda, por falta de espacio, detalla la tarea “Ver distribución variaciones” que está etiquetada con el signo “+” indicando que la tarea tiene tareas de menor nivel. El modelo de tareas resume la priorización de variaciones genéticas en dos tareas principales:

Seleccionar variedades (tarea “Seleccionar grupos de variedades”). – La priorización de variaciones inicia con la creación de dos grupos de variedades (tarea “Crear Grupo”) formados con las variaciones seleccionadas por los investigadores (tarea “Seleccionar Variedad” y tarea “Agregar a Grupo”).

Analizar (tarea “Análisis”). – Analizar implica comparar los grupos de variedades (tarea “Comparar variedades”) y examinar las variaciones resultantes hasta encontrar las variaciones de interés (tarea “Examinar Variaciones*”, el signo “+” después del nombre indica que la tarea es iterativa). Para comparar las variedades (tarea “Comparar grupos”) es necesario especificar las condiciones para ejecutar la operación de comparación (tarea “Configurar la Operación”) (p. ej., criterios de aceptabilidad, balance alélico, filtros). Las variaciones resultantes de la operación (tarea “Mostrar Resultados”) son examinadas por el investigador (tarea “Examinar Variaciones”), ya sea observando el comportamiento de la distribución de las variaciones respecto a diferentes elementos genómicos (tarea “Ver distribución variaciones”) o interactuando con los datos (tarea “Interactuar con datos”) para reducir el conjunto de variaciones a un conjunto de variaciones relevantes.

El modelo de tareas es el punto de entrada del método GenomIUm y es la base para conseguir el objetivo final del método: diseñar el concepto de la interfaz de usuario. En la siguiente sección explicamos las decisiones de diseño tomadas en cada paso del proceso de diseño del método GenomIUm. Aplicación del método

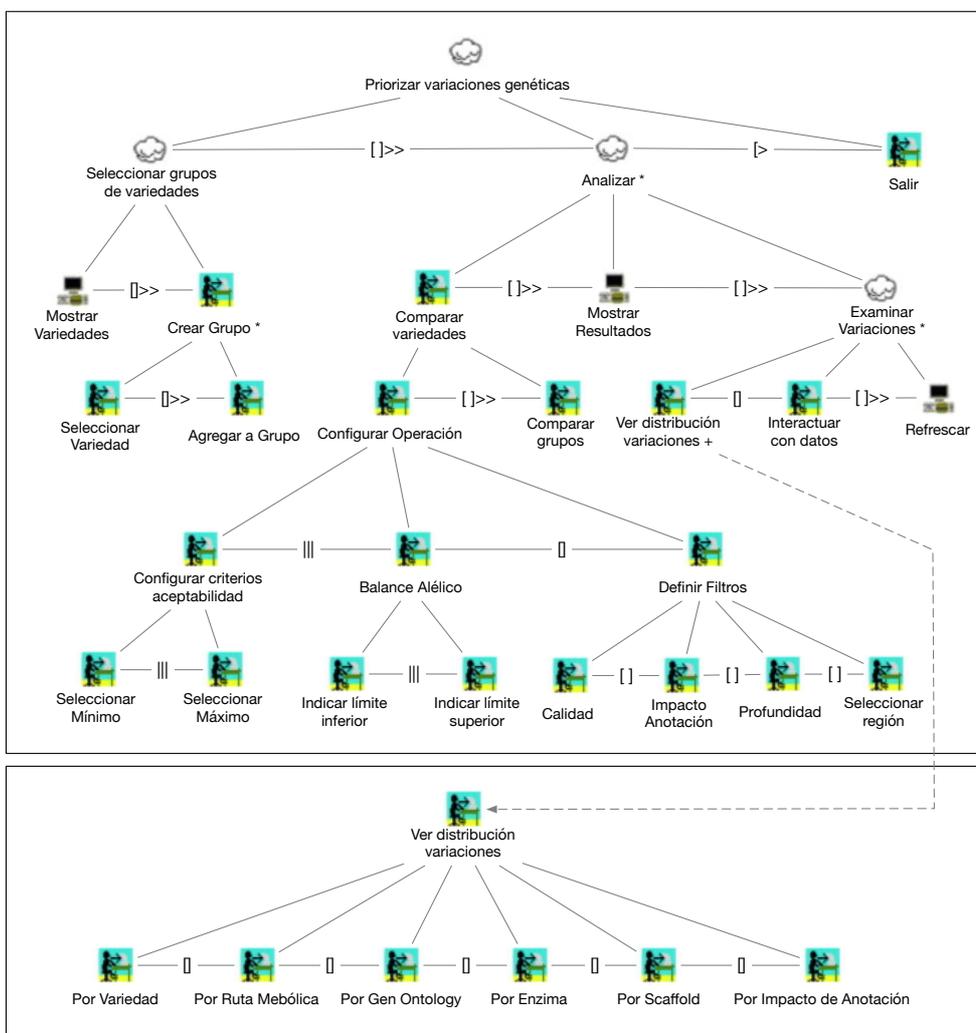


Imagen 7.12 Modelo de tareas de la priorización de variaciones genéticas

A lo largo del proceso de diseño y considerando el modelo de tareas, seleccionamos varios patrones del catálogo GenomIU y los combinamos para diseñar el concepto de la interfaz de usuario de acceso a datos genómicos, basados en las relaciones entre los patrones. La Imagen 7.13 resume los patrones (y sus relaciones) del catálogo GenomIU que son utilizados en cada paso del proceso de diseño.

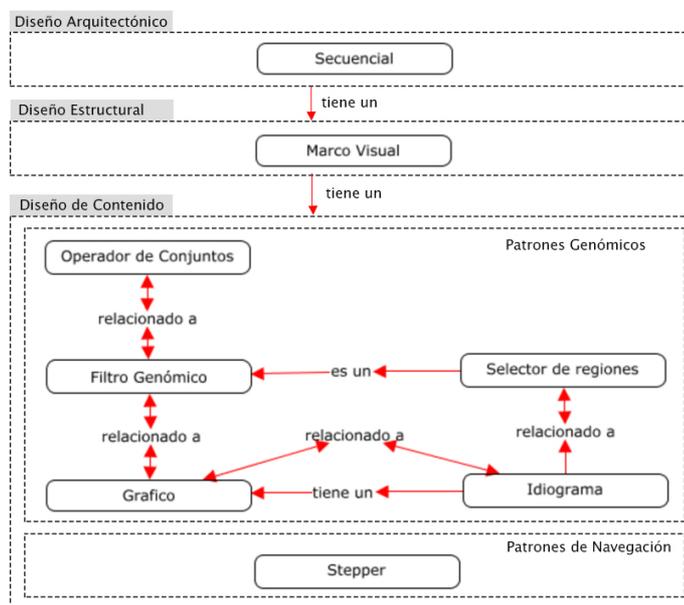


Imagen 7.13 Patrones inmersos en el caso de estudio

Considerando estos patrones y sus relaciones, describimos cómo diseñamos el concepto de la UI a través de los pasos del proceso de diseño:

Paso1: Diseño Arquitectónico

En el diseño arquitectónico, seleccionamos el patrón Secuencial porque es adecuado para guiar fácilmente al investigador por el proceso de priorización de variaciones. Es importante mencionar que en un principio planteamos tener dos UIs en el sitio, una por cada tarea principal del modelo de tareas, es decir, una para la tarea “Seleccionar grupos de variedades” a la que denominamos “UI Seleccionar Variedades” y otra para la tarea “Analizar” a la que denominamos “UI Análisis”. Sin embargo, consideramos que la tarea “Analizar” es muy amplia para ser abordada en una sola UI, puesto que la tarea implica dos subtareas: “Comparar variedades” y “Examinar Variaciones”.

Por lo tanto, decidimos dividir la UI Análisis en dos UIs. La primera UI trata la subtarea “Comparar variedades” permitiendo configurar los filtros para la comparación de variedades, a la cual hemos denominado *UI Filtro*. La segunda trata la subtarea “Examinar Variaciones” y que, por facilidad de entendimiento y preferencia de los investigadores, ha sido denominada *UI Visualizar*.

Con esta consideración, la arquitectura del sitio consistió en tres UIs conectadas con el patrón Secuencial (como lo muestra la Imagen 7.14): *UI Seleccionar Variedades*, *UI Filtrar* y *UI Visualizar*.

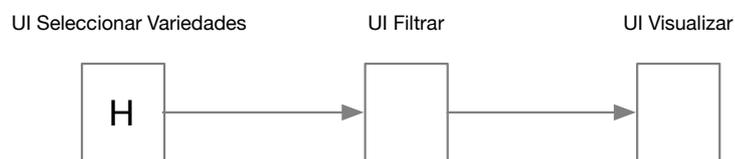


Imagen 7.14 Diseño arquitectónico del sitio web de cítricos

Paso 2: Diseño estructural

En el segundo paso del proceso de diseño, consideramos como decisión importante de diseño el mantener una estructura general y uniforme para todas las UIs del sitio con el fin de producir en el investigador la sensación de que está trabajando en un mismo ambiente, en lugar de UIs aisladas. Para esto, aplicamos el patrón “Marco Visual” de la categoría “Patrones de Página” del catálogo GenomIU.

El patrón Marco Visual es un patrón genérico y está documentado en la página 141 del libro *Designing User Interfaces* (Tidwell, 2010), bajo el nombre *Visual Framework*. Tidwell describe la solución de este patrón como: “Diseñe cada página para usar el mismo diseño básico, colores y elementos estilísticos, pero brinde al diseño suficiente flexibilidad para manejar el contenido variado de la página”. Es así como definimos un diseño básico para todas las UIs y que consiste en tres sectores (como lo muestra la Imagen 7.15): encabezado, contenido y pie de página.

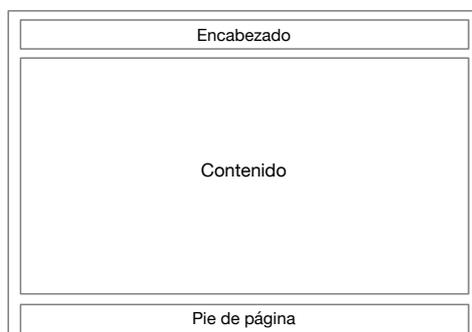


Imagen 7.15 Diseño estructural básico de las UIs del sitio web de cítricos

Paso 3: Diseño de Contenido

En este paso, enfocamos nuestra atención en diseñar el contenido de cada sector de la estructura interna de las UIs, en especial el sector “contenido” que es el sector con contenido flexible en todas las UIs. Para esto, nos apoyamos de los patrones de contenido, específicamente, en los patrones de la categoría “genómicos” y “navegación”. Cabe mencionar que los patrones de la categoría “contenedores de información”, que son parte también de los patrones de contenido, no fueron aplicados en el diseño del sitio, sin embargo, esto es una particularidad de este estudio y no se descarta el uso de este tipo de

patrones en otros proyectos de diseño. La Tabla 26 lista los patrones seleccionados en cada UI e indica el problema de diseño que resuelve cada patrón. Los patrones han sido codificados por cuestiones de explicación.

Tabla 26 Lista de patrones de contenido aplicados en el sitio web de cítricos

Código	Patrón	Problema de diseño
UI Seleccionar Variedades		
1	Operación Conjuntos	Diseñar la interacción para obtener la diferencia entre dos grupos de variedades
UI Filtrar		
2	Filtro genómico	Especificar el criterio de aceptabilidad (número mínimo y máximo de veces que una variación debe aparecer en un grupo de variedades para ser considerada en la operación de comparación)
3	Filtro genómico	Especificar el nivel de calidad de la lectura de una variación (aparición de dicha variación en una variedad durante la secuenciación).
4	Filtro genómico	Especificar el Impacto de Anotación que debe tener una variación genética para ser considerada en la operación de comparación.
5	Filtro genómico	Filtrar por Balance Alélico
6	Filtro genómico	Filtrar por Profundidad
7	Selector de regiones	Segregar las variaciones resultantes a una región específica de la secuencia
UI Visualizar		
8	Gráfico	Mostrar el número total de variaciones genéticas resultantes de la operación de diferencia entre los grupos de variedades
9	Gráfico	Mostrar el número total de SNPs resultantes
10	Gráfico	Mostrar el número de variaciones genéticas del tipo INDEL
11	Idiograma + Gráfico	Mostrar la distribución de frecuencias de variaciones genéticas sobre la secuencia genética completa
12	Idiograma + Gráfico	Mostrar la distribución de frecuencias de variaciones genéticas sobre la secuencia genética de un cromosoma
13	Gráfico	Mostrar resumen de variaciones agrupadas por variedades
14	Gráfico	Mostrar resumen de variaciones agrupadas por cromosoma (scaffold)
15	Gráfico	Mostrar variaciones agrupadas por impacto de anotación
16	Gráfico	Mostrar resumen de variaciones agrupadas por Gene Ontology
17	Gráfico	Mostrar resumen de variaciones agrupadas por enzima
18	Gráfico	Mostrar resumen de variaciones agrupadas por ruta metabólica
19	Stepper	Facilitar la navegación lineal por el proceso de priorización de variaciones.

La Imagen 7.16 muestra el diseño de contenido de las UIs con los patrones asignados a cada UI e indicados mediante un recuadro con el nombre y código identificador del patrón. La asignación de los patrones en cada UI se explica a continuación:

En la **UI Seleccionar Variedades** asignamos el patrón “Operación de Conjuntos” (recuadro 1) para resolver el problema de diseñar la interacción que permite al investigador seleccionar las variedades, definir dos grupos de variedades y ejecutar la operación diferencia de conjuntos entre los dos grupos de variedades.

En la **UI Filtrar**, asignamos cinco patrones Filtro Genómico (recuadros 2 al 6) y un patrón Selector de regiones (recuadro 7) para cubrir las tareas bajo la tarea “Configurar Operación”. Tanto el patrón Filtro Genómico como el Selector de regiones fueron identificados explotando las relaciones del patrón Operador de Conjuntos ya que el patrón Operador de Conjuntos está **relacionado al** patrón Filtro Genómico y el patrón Selector de regiones **es un** tipo de patrón Filtro Genómico, como lo muestra la Imagen 7.13.

En la **UI Visualizar** asignamos los patrones Idiograma y Gráfico, los cuales tienen dos relaciones del tipo **relacionado a** y **puede tener** que aplicaremos para solucionar los problemas de diseño derivados de las necesidades de los investigadores del IVIA: 1) los patrones Gráfico de los recuadros 8, 9 y 10 muestran el número total de resultados de la comparación en función de variaciones, SNPs y variaciones del tipo INDEL, respectivamente; 2) el patrón Idiograma en el recuadro 11 soluciona el problema de diseñar una representación completa del genoma, visualizada en forma de cromosomas o “scaffolds”, donde se pueda distinguir la localización de las variaciones dentro de la secuencia genética. Sin embargo, existe un problema: la escala de visualización de la secuencia genética es muy alta para mostrar detalladamente las variaciones, lo cual resultará en una representación sobrecargada difícil de entender. Para resolver este problema, aprovechamos que el patrón Idiograma **puede tener** un patrón Gráfico, lo cual nos permite agregar al idiograma un *histograma* que presente, a escala, la frecuencia de variaciones en rangos de la secuencia genética. Por otro lado, el patrón Idiograma del recuadro 12 está conectado interactivamente al patrón Idiograma en el recuadro 11 y presenta el idiograma y la frecuencia de variaciones de un cromosoma específico que ha sido seleccionado en el recuadro 11. Esto facilita la exploración y filtrado de las variaciones genéticas ya que el investigador puede enfocarse en las variaciones de un cromosoma específico; 3) Los patrones Gráfico de los recuadros 13 al 18 representan gráficos de resumen que muestran el porcentaje de variaciones genéticas agrupadas por varios elementos genómicos que fueron mencionados en la Tabla 26.

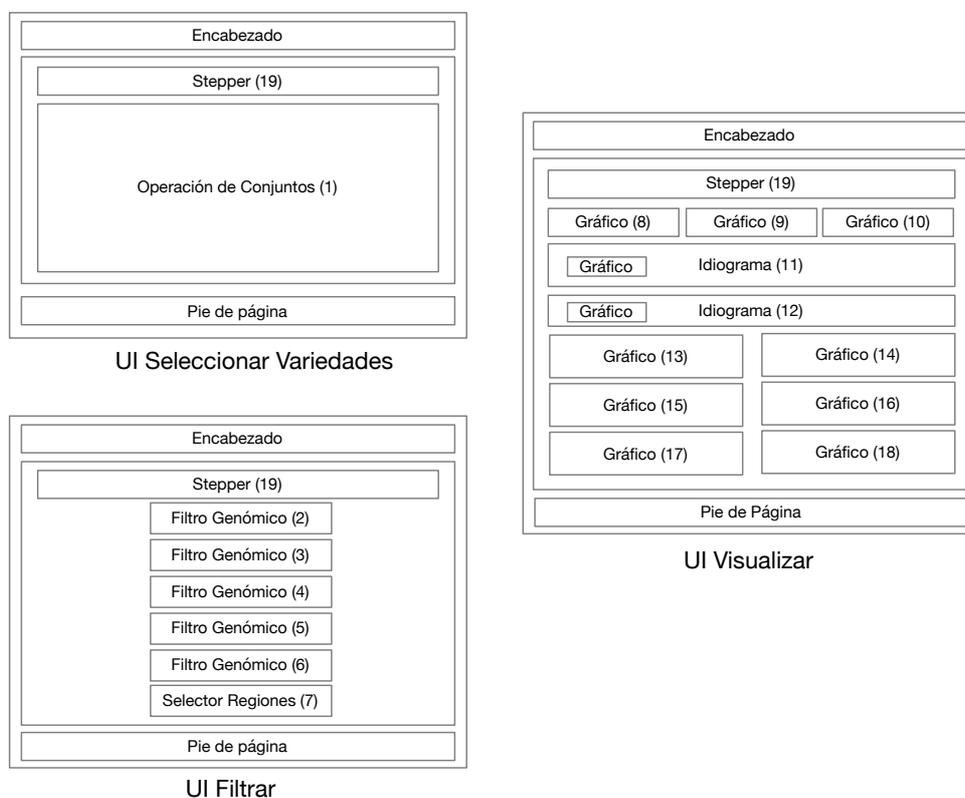


Imagen 7.16 Diseño de contenido de las UIs del sitio web de cítricos

Finalmente, el patrón Stepper, representado con el recuadro 19, ha sido asignado a todas las UIs para guiar la navegación a través de los pasos del proceso de priorización de variaciones genéticas.

Paso 4: Refinamiento

El prototipo de UIs logrado hasta aquí se convierte en un diseño conceptual temprano de la interfaz de usuario, que muestra los patrones que van a soportar las tareas de los usuarios. En el refinamiento, el último paso del proceso de diseño, nos concentramos en instanciar cada uno de los patrones adaptando la solución genérica de cada patrón al contexto del proyecto que estamos diseñando, lo que incluye definir los detalles visuales de cada patrón considerando los datos genómicos propios del proyecto.

Para ilustrar el refinamiento, nos enfocaremos en la UI Visualizar ya contiene la mayor cantidad de patrones. La Imagen 7.17 muestra el diseño refinado de la UI Visualizar donde se puede observar cómo los patrones Idiograma y Gráfico han sido instanciados para presentar el conjunto de variaciones genéticas resultantes de la comparación entre dos

grupos de variedades cítricas. Cada patrón ha sido etiquetado con el número correspondiente en la Tabla 26.

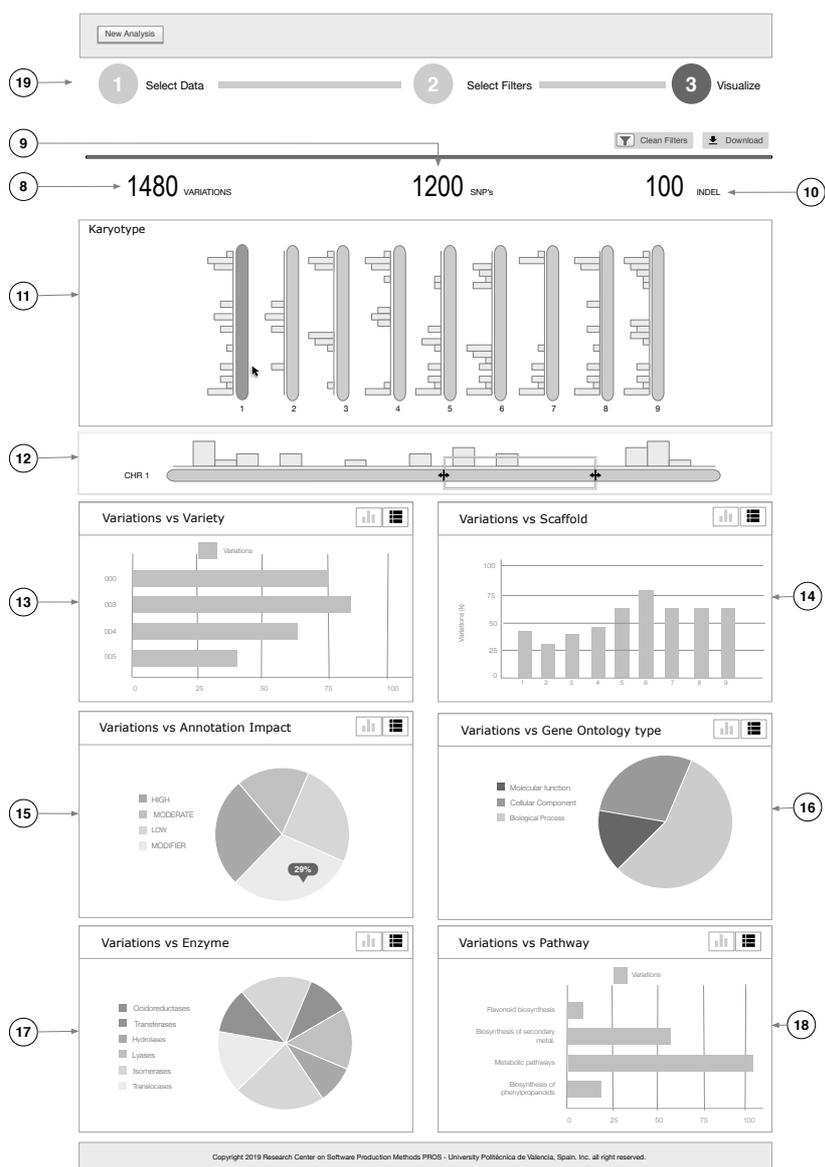


Imagen 7.17 Diseño refinado de la interfaz de “Análisis” del sistema de los cítricos

El prototipo de la UI Visualizar presentado, al igual que los prototipos de UI Seleccionar Variedades y UI Filtrar, recoge los comentarios y sugerencias obtenidas en continuas evaluaciones del prototipo con los investigadores del IVIA.

Los prototipos logrados son el resultado final del método GenomIUm y se convierten los modelos conceptuales terminados de las interfaces de usuario. Estos prototipos son el insumo para continuar con la etapa de implementación, la cual no tratamos en esta Tesis.

7.4.3 Resultados

Una vez terminado el diseño de la interfaz de usuario, analizamos la retroalimentación de los investigadores del IVIA como usuarios finales de la interfaz. Para esto, realizamos una discusión final con el investigador que participó activamente en el diseño y evaluación de la interfaz de usuario. Los comentarios en la sesión de discusión fueron grabados en formato audio para luego ser analizados detenidamente mediante un software de reproducción de audio. A continuación, describimos los resultados del análisis resumiendo los comentarios del investigador en dos afirmaciones:

1) La interfaz de usuario facilita el acceso a los datos y reduce la complejidad de manipular los datos. El siguiente comentario refuerza la afirmación: *“La operación de comparación de las variedades se torna más fácil puesto que anteriormente se utilizaba Python para manipular y operar los datos.”*.

2) La interfaz de usuario es útil y fácil de usar. El siguiente comentario refuerza la afirmación: *“...la interfaz es intuitiva y es fácil de usar...incluso esto abriría la posibilidad de que no solo investigadores del centro puedan hacer los análisis sino también los estudiantes de biología, agronomía, bioquímica, ingeniería agronómica que acuden al Centro para realizar sus prácticas y que con frecuencia no disponen de conocimientos profundos en el genoma de los cítricos e informática”*.

7.5 CONCLUSIONES

En este capítulo se ha presentado la validación del método GenomIUm a través de tres experiencias empíricas. Las dos primeras experiencias, basadas en un *experimento de comparación estadística*, validaron el método desde el punto de vista de sus componentes (patrones y proceso de diseño) y desde dos perspectivas diferentes: usuarios finales y diseñadores de las interfaces de usuario. La tercera experiencia, basada en un *caso de estudio*, validó el método en condiciones reales, aplicando el método en el contexto del diseño de interfaces de usuario para acceder a datos del genoma de las frutas cítricas.

Las conclusiones obtenidas en cada experiencia empírica se describen a continuación:

- La validación de los patrones. – Esta primera experiencia, descrita en la Sección 7.2, evaluó el impacto de los patrones en la usabilidad de las interfaces de usuario diseñadas con tales patrones desde el punto de vista de los usuarios que usan dichas interfaces. Para esto, realizamos dos réplicas de un experimento controlado que consistió en medir la eficiencia, efectividad y satisfacción de los usuarios que usaron interfaces de usuario diseñadas con y sin patrones. Un total de 30 sujetos participaron en el experimento, 24 en la primera réplica y ocho en la segunda réplica. Los resultados de este experimento indicaron que los usuarios que usaron interfaces diseñadas con los patrones experimentaron mejora en la efectividad,

eficiencia y satisfacción. Los resultados sugieren que los patrones impactan positivamente en la usabilidad de la interfaz y este impacto puede traducirse en usuarios más satisfechos que pueden completar más tareas en menos tiempo. Es importante destacar que nuestros resultados deben contextualizarse de acuerdo con las características del experimento (es decir, número de participantes, perfil de los participantes, problemas propuestos, número de réplicas). Por lo tanto, son deseables más réplicas con un número similar de participantes para poder confirmar los resultados obtenidos de los patrones.

- La validación del proceso. - Esta segunda experiencia, descrita en la Sección 7.3, evaluó cualitativamente la percepción de los diseñadores de interfaces respecto al uso del proceso de diseño propuesto por el método GenomIUm. La percepción de los diseñadores fue evaluada en términos de tres variables de percepción: facilidad de uso, utilidad percibida e intención de uso. El estudio consistió en un experimento donde 16 diseñadores de interfaces siguieron los pasos del proceso para construir el diseño conceptual de la interfaz y luego evaluaron cualitativamente la utilidad, facilidad de uso del proceso y la intención de usar dicho proceso en futuros proyectos. Los resultados obtenidos indicaron que los diseñadores estarían motivados en usar el proceso en futuros proyectos de diseño de interfaces de usuario y que esta motivación está influenciada por lo útil y fácil que representa para los diseñadores usar el proceso del método.
- Caso de estudio. - Esta tercera experiencia, descrita en la Sección 7.4, de enfoque más práctico y dirigido a la investigación agraria, consistió en aplicar GenomIUm para diseñar interfaces de usuario que permitan a investigadores comparar el genoma de varias frutas cítricas para identificar variaciones genéticas asociadas a determinadas características biológicas. Este caso de estudio se llevó a cabo con investigadores del Centro de Genética del Centro de Investigaciones IVIA de la ciudad de Valencia, España. Junto con los investigadores del Centro de Genética y en varias reuniones y talleres de trabajo, construimos el diseño conceptual de la interfaz de usuario ajustado a las necesidades de acceso a los datos. El diseño incorporó patrones genéricos y específicos del catálogo de patrones. Una vez terminado el diseño de la interfaz, analizamos la percepción de los investigadores del IVIA respecto al diseño conseguido. Los resultados indican que la interfaz conseguida fue percibida por los investigadores como útil y fácil de usar. Además, los investigadores consideran que la interfaz de usuario representa un cambio positivo y considerable en el acceso a los datos en comparación a la manera habitual de trabajo donde la línea de comandos es el mecanismo para interactuar con los datos, demandando conocimientos profundos en programación y bases de datos.

Las tres experiencias de validación representaron tres ciclos empíricos que cierran el estudio empírico de la metodología Design Science y que concluyen a su vez el ciclo de diseño con el que inició esta investigación. Los resultados obtenidos justificaron el cumplimiento de los objetivos para los cuales fue creado el método y se convierten en evidencia empírica que puede ser utilizada para futuras investigaciones.

Capítulo 8

Discusión Final

8.1 CONTRIBUCIONES

Las contribuciones de esta Tesis a la comunidad científica se derivan de los objetivos de la Tesis y pueden ser descritas como respuestas a las preguntas de investigación planteadas.

Pregunta 1 (¿Qué problemas existen en el acceso a datos genómicos?).

Respuesta: Revisamos la literatura, realizamos reuniones de discusión y entrevistas con investigadores que interactúan con datos genómicos, tanto en el entorno académico como en el industrial, e identificamos y listamos 16 problemas de interacción con los datos genómicos (descritos en la Sección 5.2.2). Esta lista de problemas se convierte en el espacio de problemas sobre el cual se construye el catálogo de patrones propuesto en esta Tesis.

Pregunta 2 (¿Cuáles son los enfoques o herramientas existentes para el diseño de interfaces que pueden ser aplicables al contexto del acceso a datos genómicos?).

Respuesta: Nuestros hallazgos en el Capítulo 4 (Estado del Arte) reportaron varios enfoques relevantes (DCU, PSA, GDD o POD) para el diseño de interfaces de usuario. Estos enfoques promueven el uso de patrones como una herramienta clave en el proceso de diseño de la interfaz y son de carácter general, es decir, pueden ser aplicados para diseñar interfaces en cualquier dominio. Respecto al uso de patrones, identificamos varias librerías de patrones dirigidas a varios dominios (p. ej., sistemas de recomendación, Big Data) y plataformas (p. ej., escritorio, web, móvil) y que pueden ser aplicadas a los enfoques mencionados. Sin embargo, a pesar de la gran cantidad de patrones contenidos en estas librerías, estos patrones no son suficientes para cubrir las particularidades del acceso a datos en el dominio genómico. Esto es lo que nos ha motivado a crear un conjunto de patrones que aborde problemas generales y específicos del diseño de interfaces en el dominio genómico.

Pregunta 3 (¿Cómo definir el método GenomIUm para diseñar interfaces de usuario de acceso a datos genómicos?).

Respuesta: Presentamos el método GenomIUm (Capítulo 6) compuesto de dos artefactos: un catálogo de patrones de diseño de interfaz, denominado GenomIU, y un proceso de diseño que es soportado por el catálogo. Los dos artefactos son abordados mediante las preguntas 3.1 y 3.2, respectivamente.

Pregunta 3.1 (¿Cómo definir un catálogo de patrones de diseño de interfaz específico para el acceso a datos genómicos?).

Respuesta: En la Sección 6.2 presentamos el proceso sistemático que hemos seguido para definir el catálogo de patrones GenomIU. El catálogo resultante contiene patrones agrupados en categorías que cubren varios problemas de diseño generales y específicos de la interfaz de un sistema genómico. Los patrones bajo la categoría “patrones genómicos” cubren problemas específicos en la interacción con los datos genómicos y se convierten en la contribución principal de esta Tesis. La identificación de los patrones genómicos que componen el catálogo de GenomIU sigue una estrategia de razonamiento inductiva: se estudian las soluciones de diseño particulares implementadas en interfaces de usuario de herramientas existentes en el dominio genómico y en el dominio del Big Data, y a partir de las soluciones particulares definimos soluciones generales aplicables al dominio genómico.

Pregunta 3.2 (¿Cómo guiar al desarrollador en el diseño conceptual de la interfaz de usuario?).

Respuesta: GenomIU está concebido bajo el enfoque de Diseño Orientado por Patrones (POD), es decir, el método propone un proceso de diseño que guía al diseñador en la definición del concepto de la interfaz de usuario, cubriendo desde la definición de las interfaces de alto nivel que estructuran el espacio de información de un sistema genómico hasta la definición detallada de la estructura y contenido de cada interfaz. Para esto, cada paso del proceso está soportado por una categoría particular de patrones del catálogo GenomIU.

Pregunta 4. (¿Qué impacto produce el método GenomIU en el diseño de interfaces de usuario de acceso a datos genómicos?).

Respuesta: Aportamos evidencia de tres experiencias empíricas que evalúan el método en el contexto del problema. Las dos primeras son experimentos controlados de comparación estadística que evalúan el método en función de sus componentes (patrones y proceso de diseño) y desde dos perspectivas diferentes: usuarios finales que usan las interfaces de usuario y diseñadores que diseñan las interfaces. La tercera experiencia es de carácter práctico y evalúa el método en un contexto real. Cada experiencia se corresponde con las respuestas a las preguntas 4.1, 4.2 y 4.3, derivadas de esta pregunta.

Pregunta 4.1 (¿Cuál es el impacto del uso de los patrones en la usabilidad de las interfaces de usuario?).

Respuesta: Aportamos evidencia empírica del impacto de los patrones en la usabilidad de interfaces de usuarios diseñadas con dichos. Esta experiencia empírica consistió en un experimento de comparación estadística que evaluó el impacto de un conjunto de cuatro patrones definidos en la primera iteración en la identificación de patrones. El experimento evaluó la usabilidad, en términos de efectividad, eficiencia y satisfacción, de interfaces de usuario diseñadas con y sin patrones. El análisis de los datos fue realizado aplicando el Modelo Lineal Mixto como método estadístico para inferir el impacto del uso de los patrones en condiciones reales. Los resultados obtenidos indican que el uso de las interfaces diseñadas con los patrones produce mejores resultados de efectividad, eficiencia y satisfacción respecto a las interfaces reales diseñadas sin patrones.

Pregunta 4.2 (¿Cuál es el impacto del uso del proceso de diseño del método GenomIU desde el punto de vista de los diseñadores de interfaces?).

Respuesta: Aportamos con evidencia empírica sobre la percepción de los diseñadores de interfaces en el uso del proceso de diseño propuesto por GenomIUm. El experimento llevado a cabo evaluó la percepción de 16 diseñadores de interfaces en términos de variables basadas en percepción: utilidad, facilidad de uso e intención de uso. Los diseñadores siguieron los pasos del proceso y usaron patrones para diseñar una interfaz de usuario solicitada. Los resultados obtenidos indican que los diseñadores perciben que el proceso de diseño es útil y fácil de aprender y, además, esta percepción es la motivación para intentar usar el proceso en futuros proyectos de diseño de interfaz de usuario.

Pregunta 4.3 (¿Cuáles es el impacto de la aplicación del método GenomIUm en un contexto del mundo real desde el punto de vista de las percepciones de las partes interesadas?).

Respuesta: Aportamos con los resultados de un caso de estudio aplicado al dominio de la investigación agraria, específicamente en el acceso a los datos del genoma de las frutas cítricas. El caso de estudio fue realizado con la colaboración de los investigadores del Centro de Genética del Instituto Valenciano de Investigaciones Agrarias (IVIA). El caso de estudio consistió en aplicar GenomIUm para diseñar interfaces de usuario que permitan comparar el genoma de diferentes frutas cítricas para identificar variaciones genéticas asociadas a diferentes propiedades biológicas. El caso de estudio provee retroalimentación relevante respecto a la aplicación GenomIUm. La aplicación del método ha generado impresiones positivas en el usuario investigador resaltando principalmente la “utilidad y “facilidad de uso” de la interfaz conseguida con GenomIUm.

8.2 RESUMEN DE PUBLICACIONES

Este trabajo de Tesis Doctoral ha generado un amplio conjunto de resultados de investigación variados, originales y relevantes, que se presentan a continuación estructurados en publicaciones, participación en proyectos de I+D competitiva y organización de congresos internacionales de primer nivel. Todo ello pone de manifiesto la intensa y exitosa actividad de diseminación realizada durante el desarrollo de la Tesis.

8.2.1 Publicaciones

La Tabla 27 muestra el resumen de un total de 11 contribuciones realizadas tanto como primer autor como en relación de coautoría y publicadas en revistas de alto impacto y en congresos nacionales e internacionales puntuados según el CORE Ranking Portal. La columna “Referencia” indica la referencia a la publicación detallada en la Tabla 28.

Tabla 27 Resumen de publicaciones

	Primer autor		Coautoría		TOTAL
	Cantidad	Referencia (Tabla 28)	Cantidad	Referencia (Tabla 28)	
Revistas					
Q1	1 (aceptado en MTA)	(A)			1
Congresos					
CORE A	2	(B, C)			2
CORE B	1	(D)	2	(E, F)	3
Sin CORE*	2	(G, H)			2
Capítulos de Libro	2	(I, J)	1	(K)	3
TOTAL	8		3		11

*Publicaciones no registradas en CORE

Tabla 28 Detalle de publicaciones

Líteral	Publicación	Año	Revista/Congreso
A	Improvement of Usability in User Interfaces for Massive Data Analysis: An Empirical Study	2019	Artículo aceptado en Multimedia Tools and Applications (MTA) (Q1)
B	Defining Interaction Design Patterns to Extract Knowledge from Big Data. (C. E. Iñiguez-Jarrín, 2018)	2018	CAiSE (A)
C	User Centered Design for Biomedical Literature Search User Interfaces (C. Iñiguez-Jarrín, Panach, et al., 2018)	2019	ISD (A)
D	GenDomus: Interactive and Collaboration Mechanisms for Diagnosing Genetic Diseases.(Iñiguez-Jarrín, 2017)	2017	ENASE (B)
E	GenesLove.Me: A Model-based Web-application for Direct-to-consumer Genetic Tests. (Reyes R. et al., 2017)	2017	ENASE (B)
F	GeIS based on Conceptual Models for the risk assessment of Neuroblastoma. (Pastor Lopez, 2016)	2017	RSIS (B)
G	Designing a Tool for Linking Genetic Variations to Diseases. (C. E. Iñiguez-Jarrín, 2018)	2018	*SISTEDES
H	A conceptual modelling-based approach to generate data value through the end-user interactions: A case study in the genomics domain. (C. Iñiguez-Jarrín, 2016)	2016	*POEM
Capítulo de Libro			
I	User Interface Design for Searching Biomedical Literature (C. Iñiguez-Jarrín, 2018)	2018	ISD
J	Guidelines for Designing User Interfaces to Analyze Genetic Data. Case of Study: GenDomus. (C. Iñiguez-Jarrín, García S., et al., 2018)	2018	ENASE
K	Genomic tools*: Web-applications based on conceptual models for the genomic diagnosis. (Reyes R. et al., 2018)	2018	Communications in Computer and Information Science

8.2.2 Participación en proyectos de investigación

- “DATAME - Desarrollo de aplicaciones Big Data” Financiado por la Universidad Politécnica de Valencia (UPV). Fecha de inicio: 01/09/2016. Duración 1 año. Ref: TIN 2016-80811-P. Importe: 18.800 €
- “IDEO - Innovative services for Digital Enterprises with ORCA (Servicios Innovadores para Empresas Digitales con ORCA)” (PROMETEO/2014/039) Financiado por la Generalitat Valenciana. Fecha de inicio: 01/06/2014. Duración 3 años y 6 meses. Importe: 143.180 €

8.2.3 Organización de congresos

- “The 11th ACM SIGCHI Symposium on Engeneering Interactive Computing Systems” del 18 al 21 de junio de 2019 en Valencia, España. Miembro del comité local organizador.
- “ER 2017 The 36th International Conference on Conceptual Modeling”, del 6 al 9 de noviembre del 2017, Valencia, España. Miembro del equipo organizador.
- “8th IFIP WG 8.1 working conference on the Practice of Enterprise Modelling (POEM 2015)” del 10 al 12 de noviembre de 2015 en Valencia, España. Miembro del equipo organizador.

8.3 TRABAJOS FUTUROS

El desarrollo de esta investigación fue guiado por la metodología DS, iniciando el ciclo investigativo con el estudio del problema, pasando por el diseño de la solución y completándolo con la validación de la solución. Durante este ciclo investigativo emergieron temas que, a pesar de ser interesantes, quedaron fuera del alcance del objetivo perseguido por la investigación. Estos temas son presentados en esta sección como futuros proyectos de investigación con el objetivo de dar continuidad a la presente investigación.

- El catálogo propuesto en esta Tesis debe ser visto como un primer esfuerzo en la definición de patrones. El siguiente paso en esta línea es enriquecer y actualizar el catálogo GenomIU con nuevos patrones observados en interfaces de usuario de aplicaciones emergentes en el dominio genómico.
- El caso de estudio presentado en esta Tesis es una experiencia que acerca al método GenomIU a su aplicación en el contexto industrial. Sin embargo, los resultados obtenidos del caso de estudio son difíciles de generalizar porque expresan situaciones particulares propias del escenario de estudio. Por lo tanto, un proyecto futuro es aplicar el método GenomIU a diversos casos de estudio con el fin de obtener resultados que permitan generalizar los efectos positivos del método reportados en esta Tesis.
- Uno de los propósitos de la Tesis fue evaluar el impacto producido por los patrones en la usabilidad de la interfaz de usuario, considerando que el impacto es producido por todo el conjunto de patrones más que por cada patrón. Un próximo proyecto es evaluar el impacto producido por cada patrón. Para esto, una potencial propuesta es evaluar los patrones por separado mediante un taller de expertos

organizado en el formato de “taller de escritores” comúnmente utilizado en las conferencias de lenguajes de patrones PLOP. Es importante también considerar criterios de evaluación diferentes al que hemos aplicado, como por ejemplo evaluar los patrones en el contexto de su usabilidad (¿Son los usuarios capaces de entender el patrón y su solución?) o la factibilidad (¿Es la información proporcionada por el patrón suficiente para implementarlo?).

- Para evaluar los efectos de los patrones en la usabilidad de la interfaz de usuario, hemos presentado dos réplicas de un experimento que produjo resultados alentadores. Con la intención de incrementar los buenos resultados obtenidos, es necesario llevar a cabo nuevas réplicas del experimento cuyos resultados agreguen evidencia empírica sobre la validez de los patrones propuestos.
- El alcance de esta investigación fue producir el diseño conceptual de la interfaz de usuario usando los patrones GenomIU. La implementación de la interfaz de usuario basada en el diseño conceptual está fuera de los límites de la investigación, sin embargo, las interfaces de usuario del caso de estudio presentado están en proceso de implementación por colaboradores del Centro de Investigación PROS. Esta implementación es manual y utilizando lenguajes estándar de la Web (p.ej., HTML, CSS, JavaScript) y entornos de desarrollo modernos (Angular, Node, etc). En este contexto nos planteamos dos proyectos:
 - La descripción en lenguaje natural de los patrones GenomIU ha facilitado la comunicación entre los implicados en el diseño de la interfaz de usuario. Sin embargo, esta forma de expresión “natural” no es suficiente para implementar eficientemente los patrones. El reto a futuro es transformar los patrones descriptivos presentados en esta Tesis a patrones generativos de forma que puedan ser entendidos por herramientas de desarrollo y lograr que su implementación sea automática o semiautomática.
 - Crear una herramienta que automatice el proceso completo de generación de interfaces de usuario, desde la caracterización del contexto de uso, pasando por el diseño conceptual de la interfaz, hasta la generación automática (código) de la interfaz de usuario. Confiamos en que el enfoque de Desarrollo Dirigido por Modelos (MDD por sus siglas en inglés) puede ser una alternativa adecuada para abordar este desafío.

Referencias

- Abrahão, S., Insfran, E., Carsí, J. A., & Genero, M. (2011). Evaluating requirements modeling methods based on user perceptions: A family of experiments. *Information Sciences*, *181*(16), 3356–3378. <https://doi.org/10.1016/j.ins.2011.04.005>
- Alemán, A., Garcia-Garcia, F., Salavert, F., Medina, I., & Dopazo, J. (2014). A web-based interactive framework to assist in the prioritization of disease candidate genes in whole-exome sequencing studies. *Nucleic Acids Research*, *42*(W1), 1–6. <https://doi.org/10.1093/nar/gku407>
- Alexander, C., Ishikawa, S., & Silverstein, M. (1977). A Pattern Language. *Ch. Alexander*. <https://doi.org/10.2307/1574526>
- Alonso, R., Salavert, F., Garcia-Garcia, F., Carbonell-Caballero, J., Bleda, M., Garcia-Alonso, L., ... Dopazo, J. (2015). Babelomics 5.0: Functional interpretation for new generations of genomic data. *Nucleic Acids Research*, *43*(W1), W117–W121. <https://doi.org/10.1093/nar/gkv384>
- An, J., Lai, J., Sajjanhar, A., Batra, J., Wang, C., & Nelson, C. C. (2015). J-Circos: An interactive Circos plotter. *Bioinformatics*, *31*(9), 1463–1465. <https://doi.org/10.1093/bioinformatics/btu842>
- Appleton, B. (2000). Patterns and Software: Essential Concepts and Terminology. *Object Magazine Online*, *2010*(07/21), 1. <https://doi.org/10.1.1.47.986>
- Bernasconi, A., Ceri, S., Campi, A., & Masseroli Marco. (2017). Conceptual Modeling for Genomics: Building an Integrated Repository of Open Data. *ER 2017*, 325–339. https://doi.org/10.1007/978-3-319-58544-4_3
- Borchers, J. O. (2000). Interaction Design Patterns : Twelve Theses. *Position Paper for the CHI Workshop "Pattern Languages for INteractoin Design: Building Momentum."*
- Bornberg-Bauer, E., & Paton, N. W. (2002). Conceptual data modelling for bioinformatics. *Briefings in Bioinformatics*, *3*(2), 166–180. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12139436>
- Broad Institute. (2019). Genome Analysis Toolkit. Retrieved February 22, 2019, from <https://software.broadinstitute.org/gatk/>
- Burriel, V., Román, J. F. R., Casanoves, A. H., Iñiguez-Jarrín, C., & León, A. (2017). GeIS based on Conceptual Models for the risk assessment of Neuroblastoma. In *11th International Conference on Research Challenges in Information Science, {RCIS} 2017, Brighton, United Kingdom, May 10-12, 2017* (pp. 451–452). <https://doi.org/10.1109/RCIS.2017.7956581>

-
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* 2nd edn. Erlbaum Associates, Hillsdale.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Boston: L-Ioughton-Mifflin.
- Cooper, A., & others. (2004). *The inmates are running the asylum: [Why high-tech products drive us crazy and how to restore the sanity]*. Sams Indianapolis.
- Cooper, A., Reinmann, R., & Cronin, D. (2007). *About Face 3.0: The essentials of interaction design. Information Visualization* (Vol. 3). <https://doi.org/10.1057/palgrave.ivs.9500066>
- Coram, T., & Lee, J. (1996). Experiences - A Pattern Language for User Interface Design. *Proc. of Joint Pattern Languages of Programs Conferences PLOP, 96*, 1–16.
- Cremonesi, P., Elahi, M., & Garzotto, F. (2015). Interaction Design Patterns in Recommender Systems. In *Proceedings of the Biannual Conference on Italian SIGCHI Chapter* (pp. 66–73). New York, NY, USA: ACM. <https://doi.org/10.1145/2808435.2808442>
- Cremonesi, P., Elahi, M., & Garzotto, F. (2017). User interface patterns in recommendation-empowered content intensive multimedia applications. *Multimedia Tools and Applications*, 76(4), 5275–5309. <https://doi.org/10.1007/s11042-016-3946-5>
- Crumlish, C., & Malone, E. (2009). *Designing social interfaces: Principles, patterns, and practices for improving the user experience*. “O’Reilly Media, Inc.”
- Dashti, M. J. S., & Gamielien, J. (2017). A practical guide to filtering and prioritizing genetic variants. *BioTechniques*, 62(1), 18–30. <https://doi.org/10.2144/000114492>
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, 35(8), 982–1003.
- Douglas, S. M., Montelione, G. T., & Gerstein, M. (2005). PubNet: a flexible system for visualizing literature derived networks. *Genome Biology*, 6(9), R80. <https://doi.org/10.1186/gb-2005-6-9-r80>
- Elliott, A. C., & Woodward, W. A. (2006). *Statistical Analysis Quick Reference Guidebook: With SPSS Examples*. Sage Publications Pvt. Ltd.
- Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511761676>
- expero. (2015). *UI Design Patterns for Navigating Complex Data Sets (Online Seminar)*. Retrieved from <https://www.experoinc.com/post/ui-design-patterns-for-navigating-complex-data-sets->

online-seminar

- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage Publications Ltd.
- Filion, G. (2015). If cars were made by bioinformaticians... | The Grand Locus. Retrieved July 2, 2019, from <http://blog.thegrandlocus.com/2015/01/if-cars-were-made-by-bioinformaticians>
- Folmer, E. (2006). Usability patterns in games. *Future Play*, 6.
- Fritz, M. S., & Arthur, A. M. (2017). Moderator Variables. In *Oxford Research Encyclopedia of Psychology*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190236557.013.86>
- García, A., Román, J. F. R., Casamayor, J. C., & Pastor, O. (2019). Towards an Effective and Efficient Management of Genome Data: An Information Systems Engineering Perspective. In *International Conference on Advanced Information Systems Engineering* (pp. 99–110).
- Graham, I., & Ian. (2002). *A pattern language for Web usability*. Addison-Wesley.
- Granlund, Å., Lafrenière, D., & Carr, D. A. (2001). A Pattern-Supported Approach to the User Interface Design Process.
- Henderson-Sellers, B., Ralyté, J., Ågerfalk, P. J., & Rossi, M. (2014). *Situational Method Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-41467-1>
- Hu, P. J., Chau, P. Y. K., Sheng, O. R. L., & Tam, K. Y. (1999). Examining the Technology Acceptance Model Using Physician Acceptance of Telemedicine Technology. *Journal of Management Information Systems*, 16(2), 91–112. <https://doi.org/10.1080/07421222.1999.11518247>
- Hübscher, C., Pauwels, S. L., Roth, S. P., Bargas-Avila, J. A., & Opwis, K. (2011). The organization of interaction design pattern languages alongside the design process. *Interacting with Computers*, 23(3), 189–201. <https://doi.org/10.1016/j.intcom.2011.02.009>
- Iñiguez-Jarrín, C. (2016). A conceptual modelling-based approach to generate data value through the end-user interactions: A case study in the genomics domain. *CEUR Workshop Proceedings*, 1765(21/12/2016), 14–21. Retrieved from <http://ceur-ws.org/Vol-1765/>
- Iñiguez-Jarrín, C. E., Panach, J. I., & Pastor, Ó. (2018). Defining Interaction Design Patterns to Extract Knowledge from Big Data. In J. Krogstie & H. A. Reijers (Eds.), *Advanced Information Systems Engineering* (Vol. 10816, pp. 539–553). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-91563-0>

-
- Íñiguez-Jarrín, C., García, A., Reyes, J. F., & Pastor, O. (2017). GenDomus: Interactive and Collaboration Mechanisms for Diagnosing Genetic Diseases. In *{ENASE} 2017 - Proceedings of the 12th International Conference on Evaluation of Novel Approaches to Software Engineering, Porto, Portugal, April 28-29, 2017.* (pp. 91–102). <https://doi.org/10.5220/0006324000910102>
- Íñiguez-Jarrín, C., García S., A., Reyes Román, J. F., & Pastor López, Ó. (2018). Guidelines for Designing User Interfaces to Analyze Genetic Data. Case of Study: GenDomus. In E. Damiani, G. Spanoudakis, & L. Maciaszek (Eds.), *Evaluation of Novel Approaches to Software Engineering* (pp. 3–22). Cham: Springer International Publishing.
- Íñiguez-Jarrín, C., Panach, J. I., & López, O. P. (2018). User Interface Design for Searching Biomedical Literature. *International Conference on Information Systems Development (ISD)*. Retrieved from <https://aisel.aisnet.org/isd2014/proceedings2018/eHealth/8>
- Jacquette, D. (2002). Ontological commitment (on Quine). In *Ontology* (pp. 156–181). Acumen Publishing. <https://doi.org/10.1017/UPO9781844653386.008>
- Jamieson, S. (2004, December). Likert scales: How to (ab)use them. *Medical Education*. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Javahery, H., & Seffah, A. (2002). A Model for Usability Pattern-Oriented Design. In *Proceedings of the First International Workshop on Task Models and Diagrams for User Interface Design* (pp. 104–110). INFOREC Publishing House Bucharest. Retrieved from <http://dl.acm.org/citation.cfm?id=646617.697237>
- Javahery, H., Seffah, A., & Radhakrishnan, T. (2004). Beyond Power: Making Bioinformatics Tools User-centered. *Commun. ACM*, 47(11), 58–63. <https://doi.org/10.1145/1029496.1029527>
- Javahery, H., Sinnig, D., Seffah, A., Forbrig, P., & Radhakrishnan, T. (2007). Pattern-Based UI Design: Adding Rigor with User and Context Variables. In K. Coninx, K. Luyten, & K. A. Schneider (Eds.), *Task Models and Diagrams for Users Interface Design* (pp. 97–108). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70816-2_8
- Kruschitz, C., & Hitz, M. (2010). Analyzing the HCI design pattern variety. *Proceedings of the 1st Asian Conference on Pattern Languages of Programs - AsianPLoP '10*, 1. <https://doi.org/10.1145/2371736.2371745>
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., ... Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, 19(9), 1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Kuehl, R. O. (2001). *Diseño de experimentos: principios estadísticos de diseño y análisis de investigación*. (Thomson Learning, Ed.) (2 ed.). México.

-
- Laakso, S. (2003). User Interface Design Patterns. Retrieved February 15, 2019, from <https://www.cs.helsinki.fi/u/salaakso/patterns/index.html>
- Laurel, B., & Mountford, S. J. (1990). *The art of human-computer interface design*. Addison-Wesley Longman Publishing Co., Inc.
- Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, 12(1), 50.
- León Palacio, A. (2019). *SILE: A Method for the Efficient Management of Smart Genomic Information*. Universitat Politècnica de València.
- Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57–78.
- Lynch, P. J., & Horton, S. (2016). *Web style guide: Foundations of user experience design*. Yale University Press.
- Martín-Rodilla, P. (2016). *Software-Assisted Knowledge Generation in the Cultural Heritage Domain: A Conceptual Framework*. Universitat Politècnica de València, Valencia (Spain). <https://doi.org/10.4995/Thesis/10251/68496>
- Mayhew, D. J. (1999). *The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design* (1st ed.). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nature Reviews. Genetics*, 11(1), 31–46. <https://doi.org/10.1038/nrg2626>
- Molina Moreno, P. J. (2003). *Especificación de interfaz de usuario: de los requisitos a la generación automática*. Universitat Politècnica de València.
- Molina Moreno, P. J., Meliá, S., & Pastor, O. (2002). Just-ui: A user interface specification model. In *Computer-Aided Design of User Interfaces III* (pp. 63–74). Springer.
- Navarrete-Hidalgo, M., Reyes Román, J. F., & Pastor López, Ó. (2018). Design and Implementation of a Geis for the Genomic Diagnosis using the SILE Methodology. Case Study: Congenital Cataract (pp. 267–274). <https://doi.org/10.5220/0006705802670274>
- Nielsen, J. (2000). End of Web Design. Retrieved April 24, 2019, from <https://www.nngroup.com/articles/end-of-web-design/>
- Nilsson, E. G. (2009). Design patterns for user interface for mobile applications. *Advances in Engineering Software*, 40(12), 1318–1328.
- Norman, D. A. (2013). *The Design of Everyday Things. Human Factors and Ergonomics in Manufacturing*

(Vol. 16). <https://doi.org/10.1002/hfm.20127>

- Nunnally, J. C., Bernstein, I. H., & Berge, J. M. F. ten. (1967). *Psychometric theory* (Vol. 226). McGraw-Hill New York.
- Oliva, R., Oriola, J., Ballesta, F., Clària, J., & Mengual, L. (2013). *Genética mèdica* (2nd ed., Vol. 39). Edicions Universitat Barcelona.
- Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., ... Trajanoski, Z. (2014). A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings in Bioinformatics*, 15(2), 256–278. <https://doi.org/10.1093/bib/bbs086>
- Palacio, A., & Pastor López, Ó. (2018). Smart Data for Genomic Information Systems: the SILE Method. *Complex Systems Informatics and Modeling Quarterly*, 0(17), 1–23. <https://doi.org/10.7250/csimq.2018-17.01>
- Pastor Lopez, O., Román, J. F. R., & G, F. V. (2016). *Conceptual Schema of the Human Genome (CSHG)*. Retrieved from <https://riunet.upv.es:443/handle/10251/67297>
- Paternò, F. (2003). ConcurTaskTrees: An Engineered Notation for Task Models. *The Handbook of Task Analysis for Human-Computer Interaction*, 483–503.
- Peissner, M., Schuller, A., & Spath, D. (2011). A Design Patterns Approach to Adaptive User Interfaces for Users with Special Needs. In J. A. Jacko (Ed.), *Human-Computer Interaction. Design and Development Approaches* (pp. 268–277). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21602-2_30
- Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. Routledge.
- Preece, J., Rogers, Y., & Sharp, H. (2015). *Interaction design: beyond human-computer interaction*. (John Wiley & Sons Inc, Ed.) (4th ed.).
- Radeke, F., Forbrig, P., Seffah, A., & Sinnig, D. (2007). PIM Tool: Support for Pattern-Driven and Model-Based UI Development. In K. Coninx, K. Luyten, & K. A. Schneider (Eds.), *Task Models and Diagrams for Users Interface Design: 5th International Workshop, TAMODIA 2006, Hasselt, Belgium, October 23-24, 2006. Revised Papers* (pp. 82–96). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70816-2_7
- Ram, S., & Wei, W. (2004). Modeling the Semantics of 3D Protein Structures (pp. 696–708). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30464-7_52
- Reyes R., J. F. (2014). *Integración de haplotipos al modelo conceptual del genoma humano utilizando la metodología síle*. Universidad Politècnica de Valencia. Retrieved from <http://hdl.handle.net/10251/43776>

-
- Reyes R., J. F. (2018). *Diseño y Desarrollo de un sistema de Información Genómica Basado en un Modelo Conceptual Holístico del Genoma Humano*. Universidad Politécnica de Valencia, Valencia (Spain). <https://doi.org/10.4995/Thesis/10251/99565>
- Reyes R., J. F., Iñiguez-Jarrín, C., & López, O. P. (2017). GenesLove.Me: A Model-based Web-application for Direct-to-consumer Genetic Tests. In *ENASE 2017 - Proceedings of the 12th International Conference on Evaluation of Novel Approaches to Software Engineering, Porto, Portugal, April 28-29, 2017*. (pp. 133–143). <https://doi.org/10.5220/0006340201330143>
- Reyes R., J. F., Iñiguez-Jarrín, C., & Pastor, Ó. (2018). Genomic tools*: Web-applications based on conceptual models for the genomic diagnosis. *Communications in Computer and Information Science*, 866(Cm), 48–69. https://doi.org/10.1007/978-3-319-94135-6_3
- Ribbecca, S. (2015). The Data Visualisation Catalogue. Retrieved February 5, 2018, from <https://datavizcatalogue.com/>
- Richardson, J., Sallam, R. L., Howson, C., Tapadinhas, J., Oestreich, T. W., & Idoine, C. (2017). Magic Quadrant for Business Intelligence and Analytics Platforms Market Definition / Description. Retrieved from <https://www.gartner.com/en/documents/3611117/magic-quadrant-for-business-intelligence-and-analytics-p>
- Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*, 340, c221. <https://doi.org/10.1136/BMJ.C221>
- Roca, I., Fernández-Marmiesse, A., Gouveia, S., Segovia, M., & Couce, M. L. (2018, May 27). Prioritization of variants detected by next generation sequencing according to the mutation tolerance and mutational architecture of the corresponding genes. *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/ijms19061584>
- Román, J. F. R., & López, Ó. P. (2016). Use of GeIS for Early Diagnosis of Alcohol Sensitivity. In *Proceedings of the BIOSTEC2016* (pp. 284–289). SCITEPRESS - Science and Technology Publications. <https://doi.org/10.5220/0005822902840289>
- Rossi, G., Schwabe, D., & Garrido, A. (1997). Design reuse in hypermedia applications development. In *Hypertext* (Vol. 97, pp. 57–66).
- Rossi, G., Schwabe, D., & Lyardet, F. (2000). User Interface Patterns for Hypermedia Applications. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (pp. 136–142). New York, NY, USA: ACM. <https://doi.org/10.1145/345513.345288>
- Rowland, F. (2015). UX design, bioinformatics and cars – EBI Interfaces. Retrieved July 2, 2019, from <https://ebiinterfaces.wordpress.com/2015/02/13/ux-design-bioinformatics-and-cars/>
- Samtools. (2016). The Variant Call Format (VCF) Version 4.2 Specification. Retrieved June 5,

-
- 2017, from <https://samtools.github.io/hts-specs/VCFv4.2.pdf>
- Schmettow, M. (2006). User interaction design patterns for information retrieval. *EuroPLoP 2006*, 489–512.
- Schümmer, T. (2003). GAMA: A Pattern Language for Computer Supported Dynamic Collaboration. In *EuroPLoP* (pp. 53–114).
- Schümmer, T. (2005). *A pattern approach for end user centered groupware development*. FernUniversität in Hagen.
- Schummer, T., & Lukosch, S. (2013). *Patterns for computer-mediated interaction*. John Wiley & Sons.
- Scott, B., & Neil, T. (2009). *Designing web interfaces: Principles and patterns for rich interactions*. “O’Reilly Media, Inc.”
- Seffah, A. (2010). The Evolution of Design Patterns in HCI: From Pattern Languages to Pattern-oriented Design. In *Proceedings of the Workshop on PEICS’10* (pp. 4–9). New York, NY, USA: ACM. <https://doi.org/10.1145/1824749.1824751>
- Seffah, A. (2015). From HCI Patterns Languages to Pattern- Oriented Design. In *Patterns of HCI Design and HCI Design of Patterns: Bridging HCI Design and Model-Driven Software Engineering* (pp. 15–33). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-15687-3_2
- Seffah, A., Gulliksen, J., & Desmarais, M. C. (2005). *Human-Centered Software Engineering — Integrating Usability in the Software Development Lifecycle*.
- Seffah, A., & Taleb, M. (2012). Tracing the evolution of HCI patterns as an interaction design tool. *Innovations in Systems and Software Engineering*, 8(2), 93–109. <https://doi.org/10.1007/s11334-011-0178-8>
- Seltman, H. J. (2012). *Experimental design and analysis*. Retrieved from <http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf>
- Sencha. (2016). Ext Js. Retrieved February 1, 2016, from <https://www.sencha.com/products/extjs/>
- Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., ... Robinson, G. E. (2015). Big data: Astronomical or genetical? *PLoS Biology*, 13(7), e1002195. <https://doi.org/10.1371/journal.pbio.1002195>
- Templeton, G. F. (2011). A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research. *Communications of the Association for Information*, 28.

-
- Tidwell, J. (1999). *Common ground: A pattern language for human-computer interface design*. O'Reilly Media. Retrieved from https://www.mit.edu/~jtidwell/common_ground.html
- Tidwell, J. (2010). *Designing interfaces: Patterns for effective interaction design*. O'Reilly Media, Inc.
- Toxboe, A. (2007). User interface design pattern library. Retrieved February 5, 2018, from <http://ui-patterns.com/>
- Van Duyne, D. K., Landay, J. A., & Hong, J. I. (2003). *The design of sites: patterns, principles, and processes for crafting a customer-centered Web experience*. Addison-Wesley. Retrieved from <http://dl.acm.org/citation.cfm?id=548998>
- Van Solingen, R., Basili, V., Caldiera, G., & Rombach, H. D. (2002). Goal question metric (GQM) approach. *Encyclopedia of Software Engineering*.
- Van Welie, M. (2008). Patterns in Interaction Design. Retrieved March 1, 2018, from <http://www.welie.com/patterns/>
- Van Welie, M., & Van der Veer, G. C. (2003). Pattern Languages in Interaction Design: Structure and Organization. In *Proceedings of interact* (pp. 1–5).
- Vanderdonckt, J., & Simarro, F. M. (2010). Generative Pattern-based Design of User Interfaces. In *Proceedings of the Workshop on PEICS'10* (pp. 12–19). New York, NY, USA: ACM. <https://doi.org/10.1145/1824749.1824753>
- Vegas, S., Apa, C., & Juristo, N. (2016). Crossover Designs in Software Engineering Experiments: Benefits and Perils. *IEEE Transactions on Software Engineering*, 42(2), 120–135. <https://doi.org/10.1109/TSE.2015.2467378>
- Vulliard, L., Chen, X., & Cui, Y. (2019). BioCircos: Interactive Circular Visualization of Genomic Data using “htmlwidgets” and “BioCircos.js” version 0.3.4 from CRAN. Retrieved September 18, 2019, from <https://rdr.io/cran/BioCircos/>
- Wallach, D., & Scholz, S. C. (2012). User-Centered Design: Why and How to Put Users First in Software Development. In A. Maedche, A. Botzenhardt, & L. Neer (Eds.), *Software for People: Fundamentals, Trends and Best Practices* (pp. 11–38). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31371-4_2
- Wieringa, R. (2014). *Design science methodology. the 32nd ACM/IEEE International Conference*. <https://doi.org/10.1145/1810295.1810446>
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering. Experimentation in Software Engineering* (Vol. 9783642290). United States: Springer. <https://doi.org/10.1007/978-3-642-29044-2>

Yahoo. (2006). Yahoo Design Pattern Library. Retrieved April 3, 2017, from <https://developer.yahoo.com/ypatterns/everything.html>

Zhang, H., Meltzer, P., & Davis, S. (2013). RCircos: An R package for Circos 2D track plots. *BMC Bioinformatics*, *14*(1), 244. <https://doi.org/10.1186/1471-2105-14-244>

Zimmer, W., & others. (1995). Relationships between design patterns. *Pattern Languages of Program Design*, *57*, 345–364.

ZURB. (2017). PatternTap. Retrieved September 9, 2019, from <https://patterntap.com/patterntap>

Lista de Imágenes

Imagen 2.1 El objeto de estudio del proyecto de investigación: el artefacto interactuando con el contexto para mejorar un problema en ese contexto.....	10
Imagen 2.2 Marco de trabajo de la metodología DS aplicado a la Tesis	12
Imagen 2.3 Ciclo de diseño de la metodología DS. Adaptado de (Wieringa, 2014).....	14
Imagen 2.4 Ciclos Empírico de la metodología DS. Adaptado de (Wieringa, 2014)	15
Imagen 2.5 Ciclos aplicados a la Tesis. Fuente: El Autor.....	17
Imagen 3.1 Vista estructural del MCGH versión 2 (MCHG v2). Fuente: (Reyes R., 2018)	23
Imagen 3.2 Flujo de trabajo básico para proyectos de secuenciación del genoma, adaptado de Pabinger et. al. (Pabinger, 2014) y GATK (Broad Institute, 2019).....	24
Imagen 3.3 Ejemplo de diseño conceptual básico	28
Imagen 3.4 Diseño conceptual de la Interfaz refinado	28
Imagen 3.5 Descripción gráfica del patrón “Entrance transition”. (Alexander, 1977) .	30
Imagen 3.6 Ejemplo donde a) no se cumple el patrón “Entrance Transition” y b) donde se cumple el patrón.	30
Imagen 4.1 Proceso de Diseño Centrado en el Usuario (ISO 9241-210)	36
Imagen 4.2 Proceso de diseño de GDD, tomado y adaptado de (Cooper et al., 2007)	37
Imagen 4.3 Marco de trabajo del PSA. Fuente: (Seffah, 2015)	38
Imagen 4.4 Resumen del lenguaje de patrones UPADE. Fuente: (Seffah, 2015)	41
Imagen 6.1 Proceso de diseño de interfaz de usuario del método GenomIU	60
Imagen 6.2 Resumen del catálogo de patrones GenomIU	61
Imagen 6.3 Modelo del catálogo de patrones	63
Imagen 6.4 Proceso para identificar patrones genómicos	66
Imagen 6.5 Formulario de carga de datos de la aplicación Bier.app	74
Imagen 6.6 Esquema del patrón Cargar Datos	75
Imagen 6.7 Formulario de carga de datos. Fuente: ICGC Data Portal	76
Imagen 6.8 Operación de conjuntos entre cromosomas. Fuente: TumorMap	76
Imagen 6.9 Esquema del patrón Operación de Conjuntos	77
Imagen 6.10 Operación de conjuntos de muestras. Fuente: ICGC Portal	78
Imagen 6.11 Secuencia de acciones de Galaxy. Fuente: Tutorial Galaxy	78
Imagen 6.12 Esquema del patrón Secuencia de Acciones	79
Imagen 6.13 Porción de interfaz de la aplicación Genesis 2.0	80
Imagen 6.14 Patrón buscar. Fuente: aplicación BioGPS	81
Imagen 6.15 Esquema del patrón Buscar	81
Imagen 6.16 Pantalla principal de Mastermind	82
Imagen 6.17 Filtros genómico. Fuente: Sequence Viewer	83
Imagen 6.18 Esquema del patrón Filtro Genómico	84
Imagen 6.19 Filtros de la aplicación GENESIS 2.0	84
Imagen 6.20 Selección de una región del cromosoma 5. Fuente: Gramene	85
Imagen 6.21 Esquema del patrón Selector de Regiones	86
Imagen 6.22 Selector de regiones cromosómicas. Fuente: GENESIS 2.0.....	86

Imagen 6.23 Patrón gráfico. Fuente: GENESIS 2.0.	87
Imagen 6.24 Esquema del patrón Gráfico	87
Imagen 6.25 Gráficos de distribución de mutaciones. Fuente: ICGC Data Portal.	88
Imagen 6.26 Idiograma de la especie humana (GRCh38.p). Fuente: elEnsembl.....	89
Imagen 6.27 a) Ejemplo de nomenclatura ISCN y b) representación de la posición 7q31.2 obtenida con Genome Decoration Page	90
Imagen 6.28 Esquema para el patrón Idiograma	91
Imagen 6.29 a) Idiograma total del genoma e b) idiograma particular de un cromosoma. Fuente: Genome Maps	92
Imagen 6.30 Pistas con anotaciones genómicas (UCSC Genes, RefSeq Genes). Fuente: ChromoZoom.....	92
Imagen 6.31 Esquema del patrón Pista Comparativa	94
Imagen 6.32 Pistas de secuencias de datos de GBrowse	95
Imagen 6.33 Configuración de columnas visibles/ocultas. Fuente: Gramene.....	95
Imagen 6.34 Esquema del patrón Columnas Ocultas	96
Imagen 6.35 Idiogramas del ratón y la rata, mapas de calor y líneas de enlace. Fuente: RCircos (Zhang, 2013)	97
Imagen 6.36 Estructura del Círculo Cromosómico.....	98
Imagen 6.37 Esquema para el patrón Círculo Cromosómico.....	99
Imagen 6.38 Interfaz de Genome Data Viewer.....	100
Imagen 6.39 Esquema del patrón Navegador Genómico	101
Imagen 6.40 Solución de navegación continua. Fuente: Librería Ext JS (Sencha, 2016)	102
Imagen 6.41 Esquema del patrón Entrega Continua	103
Imagen 6.42 Relaciones entre patrones genómicos	104
Imagen 6.43 Ejemplo de asociación entre patrones (tomado y adaptado de ICGC Data Portal).....	105
Imagen 6.44 Ejemplo de composición entre patrones. Tomado y adaptado de Genome Maps	106
Imagen 6.45 Patrón secuencial. Tomado y modificado de (Lynch, 2016)	107
Imagen 6.46 Modelo de diseño estructural de la UI aplicando el patrón Pestaña de módulo	107
Imagen 6.47 Patrón Pestaña de módulo. Tomado y adaptado de ICGC Data Portal (https://dcc.icgc.org/analysis/sets).	108
Imagen 6.48 Diseño de contenido ilustrado con el patrón Idiograma y Filtro Genómico	109
Imagen 7.1 Interfaz de usuario de la plataforma de evaluación mostrando el prototipo de PubMed sin patrones (imagen superior) y con patrones (imagen inferior).	121
Imagen 7.2 Procedimiento del experimento	123
Imagen 7.3 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del Modelo lineal mixto (MLM) y el índice de tamaño del efecto (TE) para la variable efectividad en la Réplica 1 (R1)	129

Imagen 7.4 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable efectividad en la Réplica 2 (R2)	130
Imagen 7.5 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable eficiencia en la Réplica 1 (R1)	131
Imagen 7.6 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable eficiencia en la Réplica 2 (R2)	132
Imagen 7.7 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del modelo lineal mixto (MLM) y el índice de tamaño del efecto para la variable satisfacción en la Réplica 1 (R1)	133
Imagen 7.8 Diagramas de caja y bigotes junto con sus estadísticos descriptivos, el p-valor resultante del método del modelo lineal mixto (MLM) y el tamaño del efecto para la variable satisfacción en la Réplica 2 (R2)	134
Imagen 7.9 Comparación de las variables (a) efectividad, (b) eficiencia y (c) satisfacción entre los tratamientos aplicados a la combinación de las réplicas considerando al “perfil” como variable moderadora. P-valores resultantes del MLM y tamaño de efecto para cada variable (d).	136
Imagen 7.10 Diagramas de Caja y Bigotes y estadísticos descriptivos para las variables de Utilidad Percibida (UP), Facilidad de Uso (FUP) e Intención de Uso (IDU)	145
Imagen 7.11 Modelo conceptual del genoma de los cítricos. Tomado y adaptado de (García et al., 2019)	148
Imagen 7.12 Modelo de tareas de la priorización de variaciones genéticas	150
Imagen 7.13 Patrones inmersos en el caso de estudio	151
Imagen 7.14 Diseño arquitectónico del sitio web de cítricos	152
Imagen 7.15 Diseño estructural básico de las UIs del sitio web de cítricos	152
Imagen 7.16 Diseño de contenido de las UIs del sitio web de cítricos	155
Imagen 7.17 Diseño refinado de la interfaz de “Análisis” del sistema de los cítricos	156

Lista de Tablas

Tabla 1 Ejemplos de bases de datos biológicas	22
Tabla 2 Librerías de patrones encontradas en la literatura	43
Tabla 3 Tipos de librerías de patrones	44
Tabla 4 Dominios de aplicación de las librerías de patrones.....	44
Tabla 5 Librerías de patrones en línea	45
Tabla 6 Grupos de patrones definidos por Duyne. Traducido de (Van Duyne, 2003)	45
Tabla 7 Evaluación de enfoques respecto a los requisitos de la solución	55
Tabla 8 Razones de usar patrones mapeadas a los requisitos de GenomIUm	56
Tabla 9 Problemas cubiertos por los patrones	68
Tabla 10 Herramientas y librerías de visualización genómica	69
Tabla 11 Herramientas de bibliografía biomédica	71
Tabla 12 Herramientas de análisis de variaciones genéticas	71
Tabla 13 Herramientas de inteligencia de negocio y analítica de datos	72
Tabla 14 Formato para documentar patrones GenomIU.....	73
Tabla 15 Resumen de las Preguntas de Investigación, hipótesis, variables de respuesta y métricas.	116
Tabla 16 Perfil académico de los sujetos experimentales por réplica.	117
Tabla 17 Frecuencia de uso de sitios web de búsqueda utilizados por los sujetos de la réplica 1 (R1) y replica 2 (R2).	117
Tabla 18 Preguntas para el problema PubMed.....	118
Tabla 19 Preguntas para el problema Scopus.	119
Tabla 20 Diseño del experimento.....	122
Tabla 21 P-valores resultantes de la prueba de normalidad de Shapiro-Wilk aplicado al valor de residuo del Modelo Lineal Mixto (MLM) para la Replica 1(R1) y Replica 2 (R2)	128
Tabla 22 Ítems para medir las variables de percepción: utilidad percibida (UP), facilidad de uso percibida (FUP) e intención de uso (IDU)	141
Tabla 23 Resumen de variables dependientes, sus métricas y su escala de medida	141
Tabla 24 Resultados de la prueba Wilcoxon de los rangos con signo de una muestra, respecto al valor hipotetizado (3).....	146
Tabla 25 Resultados del modelo de regresión lineal simple	146
Tabla 26 Lista de patrones de contenido aplicados en el sitio web de cítricos.....	153
Tabla 27 Resumen de publicaciones	162
Tabla 28 Detalle de publicaciones	162