



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



FACULTAD DE
INGENIERÍA

Ontologías computacionales para las metodologías de investigación cualitativas

Informe de Proyecto de Grado presentado por

Emilio Méndez , Juan Pablo Sandin y Facundo Fleitas

en cumplimiento parcial de los requerimientos para la graduación de la carrera
de Ingeniería en Computación de Facultad de Ingeniería de la Universidad de
la República

Supervisores

Regina Motz
Edelweis Rohrer

Montevideo, 17 de enero de 2025



Ontologías computacionales para las metodologías de investigación cualitativas por Emilio Méndez , Juan Pablo Sandin y Facundo Fleitas tiene licencia [CC Atribución 4.0](https://creativecommons.org/licenses/by/4.0/).

Agradecimientos

A nuestras familias, amigos y parejas, por su apoyo desde el principio hasta el final en este largo viaje. Gracias por compartir nuestras victorias, pero sobre todo, por estar ahí en cada una de nuestras derrotas, ayudándonos a levantar impulsándonos una y otra vez a intentarlo una vez más.

A cada docente que nos formó para estar hoy presentando este trabajo, pero en especial a Regina, Edelweis y Alén, que sin su guía y colaboración esto no habría sido posible.

Y por último, a nosotros mismos, porque a pesar de todas las dificultades en el largo camino, no bajamos los brazos.

Resumen

El proyecto desarrolla una ontología computacional que modela el proceso de investigación cualitativa, abordando desafíos relacionados con la integración, organización y análisis de datos. La solución propuesta permite estructurar información proveniente de diversos proyectos, facilitando consultas complejas, estructuración del conocimiento y colaboración entre investigadores sin interferir ni limitar el proceso de interpretación de los mismos. Para el desarrollo de la ontología, se utilizó el lenguaje OWL (Web Ontology Language), estándar del World Wide Web Consortium, que emplea lógica descriptiva para garantizar una representación formal y semántica precisa utilizando la plataforma protégé como entorno de desarrollo y ejecución.

El diseño de la ontología se basó en un enfoque bottom-up, comenzando con conceptos específicos del dominio definidos en colaboración con el usuario experto. A través de iteraciones y validaciones, se logró un modelo conceptual flexible que refleja las fases principales del proceso investigativo: formulación, campo y análisis. Además, este modelo destaca la relevancia del concepto de “interpretación” dentro del dominio, posicionándolo como un elemento central y principal factor de la flexibilidad mencionada previamente. Restringir el proceso de interpretación mediante la incorporación de axiomas o restricciones innecesarias en la ontología implicaría una limitación para el proceso investigativo y, en consecuencia, para la adopción de nuestra herramienta. Por esta razón, mantener dicha flexibilidad fue una de las directrices fundamentales durante el desarrollo del proyecto.

La evaluación de la ontología se llevó a cabo mediante preguntas de competencia y herramientas de análisis como OOPS! y OntoClean, validando su consistencia y capacidad para satisfacer las necesidades del dominio. El proyecto también incluye una herramienta visualizadora que actúa como prueba de concepto, integrando una interfaz gráfica para cargar instancias, ejecutar razonamientos, realizar consultas y explorar relaciones entre los datos. Esta herramienta demuestra la capacidad del modelo para responder a preguntas complejas y gestionar datos cualitativos de forma coherente y eficiente.

Palabras clave: Ontología, Investigación cualitativa, OWL, Protégé, Razonamiento computacional

Índice general

1. Introducción	1
2. Conceptos de base	7
2.1. Ontologías computacionales	7
2.1.1. Metodologías para el desarrollo de ontologías	8
2.2. Metodologías de investigación cualitativa	9
3. Desarrollo de la ontología	13
3.1. Objetivo de la ontología	13
3.2. Diseño y construcción	16
3.2.1. Diseño de la ontología	18
3.2.2. Construcción de la ontología	21
4. Validación de la Ontología	27
4.1. Validación de Buenas Prácticas	27
4.1.1. Ontology Pitfall Scanner (OOPS!)	28
4.1.2. Análisis basado en OntoClean	30
4.2. Validación mediante Preguntas de Competencia	34
4.2.1. Reflexiones finales de la validación	45
5. Carga y visualización de datos	47
5.1. Proceso de Carga de Datos	47
5.1.1. Preparación del Entorno	48
5.1.2. Creación de una Instancia	48
5.1.3. Asignación de Propiedades	48
5.1.4. Creación de Relaciones	51
5.1.5. Orden de Creación de Clases Relevantes	53
5.1.6. Validación de Carga	55
5.1.7. Casos Reales	56
5.1.8. Dificultades encontradas en la carga de datos	57
5.2. Herramienta de visualización	57
5.2.1. Objetivo y alcance	58
5.2.2. Arquitectura y diseño de la herramienta	59
5.2.3. Resultado final y trabajo futuro	61

6. Conclusiones y Trabajo Futuro	65
6.1. Evaluación de Resultados y principales aportes	65
6.2. Dificultades y Limitaciones Encontradas	67
6.3. Posibles Extensiones y Trabajo Futuro	68
Referencias	71

Capítulo 1

Introducción

Existen dos grandes grupos dentro de las investigaciones: investigaciones cuantitativas e investigaciones cualitativas. Si bien ambos definen procedimientos para la recopilación y el análisis de datos, se diferencian en el tipo de información que buscan. Mientras que las investigaciones cuantitativas se enfocan en datos numéricos y buscan identificar patrones y relaciones estadísticas, las investigaciones cualitativas se centran en interpretaciones más profundas de experiencias y conceptos, utilizando datos no numéricos como palabras y significados. Ambas metodologías son fundamentales, pero cada una responde a diferentes tipos de preguntas y necesidades de investigación. ([Academy, 2021](#))

Pongámonos en el lugar de un investigador cualitativo que busca categorizar todas las técnicas utilizadas en diferentes proyectos para entender cómo se recopilan y analizan los datos sobre un tema en particular. O, tal vez, quiere explorar todas las preguntas formuladas en entrevistas de proyectos que comparten un objetivo similar al suyo. O, por qué no, intentar obtener los esquemas de clasificación analíticos utilizados en diferentes proyectos para entender cómo se categoriza la información en estudios variados. Lo que parece ser una tarea sencilla en teoría, rápidamente se convierte en un desafío. Cada proyecto utiliza sus propias herramientas: unos guardan sus entrevistas en formularios de Google, otros almacenan transcripciones en archivos de Word dispersos en carpetas de Drive, e incluso algunos siguen optando por el lápiz y el papel. Aun si los datos estuvieran accesibles, su desorganización y la falta de una estructura común dificultan cruzar información entre investigaciones.

Este tipo de situaciones son una realidad cotidiana para los investigadores cualitativos. A pesar del avance de la tecnología, la forma en que los proyectos se organizan sigue siendo tan diversa como los enfoques metodológicos que utilizan. La falta de una estructura común para almacenar, consultar y compartir datos entre proyectos no solo hace que sea difícil extraer información relevante, sino que también limita la capacidad de los investigadores para construir sobre el trabajo de otros y fomentar una verdadera ciencia abierta y colaborativa.

En este contexto, proponemos una solución: una ontología computacional que modele el proceso de investigación cualitativa. Lejos de imponer restricciones sobre cómo los investigadores deben llevar a cabo su análisis, nuestra propuesta busca generar una infraestructura que permita integrar, relacionar y consultar datos de múltiples proyectos de manera eficiente y coherente. Al proporcionar una estructura estandarizada, esta ontología haría posible lo que hoy es difícil: la capacidad de realizar consultas complejas y detalladas sobre proyectos de investigación cualitativos, así como categorizar y vincular datos de diferentes etapas del proceso de investigación. Por ejemplo, gracias a esta ontología, se podría obtener con facilidad todas las investigaciones que mencionan a un autor específico en cualquier parte del proceso. Asimismo, se podrían recuperar las preguntas realizadas en entrevistas de proyectos que compartan un objetivo común, sin importar cómo y dónde fueron almacenadas originalmente. Hoy, una tarea así requiere una gran cantidad de tiempo y esfuerzo, debido a la naturaleza fragmentada y dispareja de las herramientas y formatos utilizados por los investigadores cualitativos.

La razón por la que creemos que una ontología es la herramienta indicada para enfrentar estos desafíos radica en la definición que tomaremos como referencia en nuestro proyecto: *“Una especificación formal y explícita de una conceptualización compartida.”* (Studer, Benjamins, y Fensel, 1998). En este contexto, la conceptualización compartida se refiere al proceso de investigación cualitativa, mientras que la especificación formal, basada en la lógica de descripciones, es precisamente lo que buscamos construir. En términos más sencillos, una ontología actúa como un marco de referencia común que permite estructurar y organizar la información de manera coherente y accesible. Esta estructura no solo facilita la integración y el intercambio de datos entre diferentes proyectos, sino que también permite resolver consultas complejas y derivar información valiosa a partir del modelo de la ontología, gracias al uso de diferentes mecanismos de razonamiento basados en la semántica formal de lógica descriptiva. De esta forma, los investigadores obtienen resultados en forma sencilla, en comparación con el esfuerzo que implica usar herramientas desarticuladas. En estas últimas, a veces resulta imposible llegar a los mismos resultados ya que no se registra la misma información.

Sin embargo, desarrollar una ontología que modele el proceso de investigación cualitativa presenta desafíos significativos. Crear un modelo conceptual que capture adecuadamente la diversidad de enfoques y metodologías utilizadas en este tipo de investigación es una tarea compleja. La dificultad radica en diseñar una estructura que no solo sea flexible para adaptarse a diferentes métodos de investigación, sino que también sea lo suficientemente detallada como para representar adecuadamente las múltiples dimensiones del proceso investigativo. Esta tarea requiere una profunda comprensión tanto del proceso investigativo en sí como de las técnicas de modelado ontológico, lo que hace que el desarrollo de esta ontología sea una herramienta valiosa para mejorar la integración y el

análisis de datos cualitativos.

Nuestro trabajo ha involucrado una serie de etapas fundamentales para abordar los desafíos en la modelización del proceso de investigación cualitativa. En primer lugar, realizamos una definición de conceptos base que nos sirven de guía a lo largo del desarrollo del proyecto. Tener estos conceptos definidos desde el inicio es importante ya que otorgan un marco conceptual para el desarrollo del proyecto y permiten establecer un lenguaje común a través del mismo. Luego, realizamos una revisión de antecedentes que nos permitió evaluar el estado actual de las herramientas de análisis cualitativo y comprender las limitaciones de las metodologías existentes. A partir de esta revisión, identificamos vacíos y oportunidades de mejora en la integración de datos y la colaboración entre proyectos, estableciendo una base sólida sobre la cual desarrollar nuestra propuesta ontológica.

Entre los estudios encontrados, uno de los más relevantes aborda los desafíos de desarrollar una ontología computacional a partir de una investigación de métodos mixtos, lo que resulta particularmente útil para la integración de datos cualitativos y cuantitativos en un modelo unificado. El estudio destaca cómo un equipo interdisciplinario, proveniente de paradigmas teóricos diferentes, pudo colaborar eficazmente en el desarrollo de herramientas digitales, como bases de datos y ontologías computacionales, *“sin sacrificar el rigor epistemológico del enfoque cualitativo.”* (Hanchard y Merrington, 2019), lo cual nos parece relevante ya que es un antecedente positivo de la integración de un proceso de investigación cualitativa con tecnología.

Por otro lado, encontramos relevante el estudio sobre QualiCO (Hocker, Bipat, McDonald, y Zachry, 2021), una ontología diseñada específicamente para esquemas de codificación cualitativa. Este estudio aborda la cuestión de cómo los métodos cualitativos han sido en gran medida omitidos en las discusiones sobre ciencia abierta. A pesar de que la ciencia abierta ha avanzado en la compartición de datos y métodos, los métodos cualitativos han sido menos abordados, a menudo debido a la diversidad de enfoques metodológicos dentro de esta categoría. La investigación revela que QualiCO facilita la reutilización de esquemas de codificación, mejorando la precisión y eficiencia en las tareas de codificación.

El estudio también señala que la naturaleza diversa de los métodos cualitativos, desde enfoques interpretativos hasta métodos más sistemáticos y positivistas, plantea desafíos únicos para la estandarización en la ciencia abierta. Las ontologías como QualiCO pueden ayudar a superar estos desafíos al proporcionar un marco estructurado para compartir esquemas de codificación, haciendo que la investigación cualitativa sea más transparente y facilitando la reutilización de esquemas. No obstante, el estudio advierte que para que tales sistemas faciliten la ciencia abierta de manera efectiva, los costos de participación deben ser mínimos para los investigadores, integrándose sin problemas en su flujo de trabajo.

Este enfoque es relevante para nuestro trabajo, ya que buscamos aplicar principios similares para optimizar la gestión y reutilización de datos cualitativos en nuestro proyecto. Sin embargo, aunque ambos trabajos comparten principios, nuestros objetivos y los de QualiCo son distintos dentro del proceso de investigación cualitativa. Mientras que nosotros consideramos y modelamos los esquemas de codificación como un elemento adicional en nuestro dominio, QualiCo se centra únicamente en estos. Esto plantea una pregunta importante sobre la posibilidad de integrar ambos enfoques en el futuro y modelar, dentro de nuestro dominio, los esquemas de codificación utilizando el desarrollo de QualiCo.

Por último, vale la pena mencionar los CAQDAS (Computer-Assisted Qualitative Data Analysis Software), herramientas informáticas diseñadas para apoyar la investigación cualitativa. Estos programas, como ATLAS.ti (Stewart, 2024), facilitan la organización, codificación y análisis de grandes volúmenes de datos textuales y no textuales, tales como imágenes y audios. Al permitir una gestión más eficiente de datos complejos, los CAQDAS ayudan a los investigadores a realizar análisis profundos y estructurados. La integración de estas herramientas en la investigación cualitativa representa un avance significativo, ofreciendo nuevas capacidades para explorar y entender los datos de manera más eficaz. Si bien esta herramienta no es una ontología, nos permite evidenciar cómo ya existen elementos de software en el área que asisten al proceso de investigación cualitativa.

Aunque no hemos encontrado proyectos específicos que generen una ontología computacional para el proceso de investigación cualitativa, sí hemos identificado trabajos relevantes en el área. Entre las conclusiones destacadas por los autores anteriormente mencionados, resalta la importancia de la integración interdisciplinaria y la necesidad de minimizar los costos de adopción de estas herramientas por parte de los investigadores. Además, se ha señalado que la adopción de un modelo estructurado para la investigación cualitativa podría percibirse como una restricción a la flexibilidad inherente a este tipo de investigación, lo cual podría generar resistencia entre los investigadores. Por lo tanto, cualquier iniciativa en este ámbito debe equilibrar la estandarización con la flexibilidad para ser aceptada y efectiva en la práctica.

Vemos entonces cómo nos encontramos con un proyecto innovador y a su vez desafiante. La creación de una ontología computacional para la investigación cualitativa ofrece una oportunidad para sistematizar la práctica investigativa, estableciendo una estructura que facilite la organización y el análisis de datos cualitativos de manera más ordenada y coherente. Sin embargo, este desafío implica enfrentar barreras prácticas y teóricas, tales como la integración de diversos enfoques metodológicos y la minimización de costos para los usuarios. El éxito de un proyecto en esta área dependerá de la capacidad para ofrecer una herramienta que respete la flexibilidad de la investigación cualitativa, mientras

proporciona un marco sistemático que mejore la eficiencia y claridad del proceso investigativo.

Por otro lado, en la etapa de diseño y construcción, llevamos a cabo un análisis de los requerimientos específicos de los investigadores cualitativos. Este análisis resultó clave para la arquitectura de la ontología, en la que definimos conceptos, relaciones y restricciones, asegurando una estructura que permitiera organizar y consultar de manera eficiente los datos de investigación cualitativa. El proceso de diseño incluyó varias iteraciones, en las que trabajamos estrechamente con Alén Pérez (MSc. Sociólogo), docente responsable del curso "Metodología de la Investigación en Información y Comunicación" de la Facultad de Información y Comunicación de la Universidad de la República, lo cual permitió ajustar el modelo para captar la complejidad y diversidad de este tipo de investigaciones. En esta sección, se detalla cómo logramos una estructura flexible y reutilizable que se adapta a las diversas necesidades de los investigadores.

Para garantizar que nuestra ontología cumpliera con los objetivos planteados, formulamos una serie de preguntas de competencia que guiaron tanto el diseño como la evaluación de la solución. Estas preguntas se enfocaron en aspectos clave como la integración de técnicas de investigación, la gestión de datos y la capacidad de realizar consultas complejas. En la sección de validación, evaluamos la ontología en escenarios reales, probando su eficacia en responder a las preguntas de competencia y realizando ajustes según los resultados obtenidos. Este proceso de validación demostró la capacidad de la ontología para mejorar la organización y el acceso a la información, cumpliendo con los requisitos de los investigadores cualitativos.

Con el objetivo de ilustrar el proceso de carga de instancias en la ontología de forma manual y compartir las experiencias obtenidas al respecto, desarrollamos una sección dedicada exclusivamente a la carga de datos. Además, creamos una herramienta visualizadora como prueba de concepto, diseñada para integrar diferentes elementos a través de una interfaz gráfica. Esta herramienta permite cargar una ontología con instancias de proyectos, utilizar el razonador, visualizar datos de proyectos, realizar consultas y, de forma ordenada y orientada a nuestro trabajo, ingresar nuevas instancias. La herramienta visualizadora está disponible en https://eldkaii.github.io/ontology_viewer/index.html, acompañada de un manual de usuario y una descripción detallada de la prueba de concepto desarrollada.

Finalmente, en el capítulo de conclusiones y trabajo futuro, reflexionamos sobre los logros alcanzados, las lecciones aprendidas y las oportunidades de expansión para la ontología. Esta última sección explora posibilidades de extensión y mejoras, abriendo la puerta a desarrollos futuros que podrían incluir sistemas integrados de gestión y análisis de datos cualitativos, así como la incorporación de tecnologías de inteligencia artificial para mejorar la accesibilidad y aplicabilidad del modelo en investigaciones cualitativas.

Capítulo 2

Conceptos de base

En el desarrollo de la presente investigación, resulta fundamental situar el problema de estudio en un contexto más amplio, explorando y analizando investigaciones previas y teorías relevantes que han abordado temas similares. Esta sección tiene como objetivo proporcionar un marco teórico que no solo facilite la comprensión del tema en cuestión, sino que también evidencie las brechas existentes en el conocimiento actual, las cuales justifican y fundamentan la realización de este estudio.

2.1. Ontologías computacionales

Las ontologías computacionales son el concepto central del desarrollo de este trabajo. Es por ello que es importante entender a qué nos referimos cuando hablamos de una ontología, cuáles son sus capacidades y cómo se pueden aplicar en distintos contextos.

Para representar un dominio y almacenar información, existen diferentes enfoques, como bases de datos relacionales y no relacionales, que ofrecen estructuras organizativas eficaces, validadas por su amplio uso en la industria del desarrollo de software. Sin embargo, en este proyecto, elegimos utilizar ontologías, ya que permiten modelar y formalizar conocimientos complejos de manera flexible y adaptativa, adecuándose a las necesidades específicas de este tipo de investigación.

El término **ontología** tiene diversas interpretaciones dependiendo del campo de estudio. Por un lado, en el ámbito de la Filosofía, se define como la *“parte de la metafísica que trata del ser en general y de sus propiedades trascendentales”*. Por otro lado, en el contexto computacional, que es el que nos concierne en este estudio, se refiere a *“una red o sistema de datos que define las relaciones existentes entre los conceptos de un dominio o área del conocimiento.”* (RAE, 2024).

Este término ha sido explorado y redefinido por diversos autores; sin embargo, en este contexto, hemos decidido destacar la definición propuesta por Studer, Benjamins y Fensel (Studer y cols., 1998), quienes describen una ontología como:

“Una especificación formal y explícita de una conceptualización compartida.

Dado que esta definición es clara, concisa y precisa, la adoptaremos como referencia fundamental a lo largo de este trabajo al tratar el concepto de ontologías.

Para la construcción de la ontología en este trabajo, utilizamos el OWL (Web Ontology Language), lenguaje estándar de W3C (W3C, 2024) para las ontologías computacionales, que permite a los usuarios escribir conceptualizaciones formales y explícitas de modelos de dominio, cumpliendo con requisitos esenciales como una sintaxis bien definida, una semántica precisa y soporte para razonamiento eficiente.

La semántica formal de OWL basada en lógica descriptiva (Baader, Calvanese, McGuinness, Nardi, y Patel-Schneider, 2007) evita interpretaciones subjetivas y permite que tanto las máquinas como los humanos razonen sobre la información. Esto incluye la aplicación de los mecanismos que forman parte de la lógica descriptiva como la consistencia, inclusión o “subsumption” de clases e instanciación. Es importante destacar que OWL se construye a partir de RDF, por lo que este mantiene una sintaxis muy similar. (Staab y Studer, 2009)

Por tanto, vemos que las ontologías proporcionan una estructura explícita y compartida que resulta fundamental en un contexto donde los proyectos de investigación tienden a utilizar distintos formatos y metodologías. A diferencia de otros modelos, una ontología permite establecer un marco común en el cual se pueden organizar y relacionar conceptos heterogéneos. Esto no solo facilita la integración y el acceso a datos entre proyectos, sino que también abre la puerta a consultas complejas y detalladas, aprovechando capacidades de razonamiento lógico que no están disponibles en las bases de datos convencionales.

En conclusión, una ontología no solo ofrece un marco unificado para gestionar datos complejos, sino que también maximiza el valor de los datos al facilitar inferencias y la integración entre proyectos, superando las limitaciones de otros sistemas de modelado. Por ello, consideramos que las ontologías son la herramienta óptima para abordar los objetivos de este proyecto, proporcionando una base sólida y versátil para los procesos de investigación cualitativa.

2.1.1. Metodologías para el desarrollo de ontologías

Para desarrollar una ontología, existen dos enfoques principales: top-down (de arriba hacia abajo) y bottom-up (de abajo hacia arriba). En términos generales, el enfoque top-down comienza desde conceptos abstractos y de alto nivel, conocidos como ontologías fundacionales o de nivel superior, que proporcionan

principios y estructuras genéricas para el modelado de conocimiento. Este enfoque se basa en axiomas y patrones ontológicos predefinidos para evitar reinventar la rueda”, logrando una estructura general y coherente desde el inicio que puede luego personalizarse con detalles más específicos. Así, el top-down puede guiar en la definición y adición de axiomas en ontologías más especializadas, lo que resulta particularmente útil para proyectos que buscan una base sólida y formal desde el comienzo.

Por otro lado, el enfoque bottom-up parte del conocimiento específico del dominio y de materiales existentes, tales como datos, documentación y prácticas ya en uso, para construir la ontología de manera progresiva. Este método permite reutilizar fuentes de conocimiento que los expertos del dominio ya manejan, facilitando así la adquisición y organización de conocimiento sin empezar desde cero. Según estudios como los de Simperl et al. (Simperl, Mochol, Bürger, y Popov, 2010), un análisis exhaustivo del dominio puede tener un impacto significativo en el desarrollo de ontologías, aunque las herramientas para soportar este análisis suelen ser limitadas. En este sentido, el enfoque bottom-up ofrece una ventaja práctica, permitiendo aprovechar estos recursos ”legados” de manera efectiva y minimizando el esfuerzo de extracción de conocimiento de los expertos. (Keet, 2020)

2.2. Metodologías de investigación cualitativa

Anteriormente, mencionamos que las ontologías computacionales constituyen el concepto central de este trabajo. Sin embargo, para comprender su utilidad en este contexto, es importante definir el ámbito específico de aplicación: las metodologías de investigación cualitativa. En esta sección, ofreceremos una breve introducción a las metodologías de investigación cualitativa, presentando los conceptos fundamentales necesarios para el desarrollo de nuestro estudio.

De forma análoga que con las ontologías, comencemos por definir a qué nos referimos con el término ”Metodologías de investigación cualitativa”. La investigación cualitativa es un tipo de investigación científica que busca entender cierta problemática o caso de estudio asociado a la perspectiva de los sujetos involucrados. Esta es especialmente efectiva para obtener información cultural acerca de valores, opiniones, comportamientos y contextos sociales de poblaciones específicas. (Mack, Woodsong, M.MacQueen, Guest, y Namey, 2005)

Para recolectar la información deseada, estos enfoques se apoyan en una variedad de métodos propios de la investigación cualitativa, cada uno aportando una perspectiva única y valiosa. Entre los más relevantes se encuentran la observación participante, la entrevista en profundidad, el focus group y el análisis de documentos. A continuación, ofrecemos una breve descripción de cada uno de estos enfoques para ilustrar cómo contribuyen a la riqueza de los datos recopilados.

La observación participante se realiza mediante la inserción del investigador en cierta población con el objetivo de analizar las distintas perspectivas que existen sobre una temática en cuestión. De forma general, podríamos decir que el investigador intenta entender cómo es la vida para un “insider”¹ mientras permanece, inevitablemente, como un “outsider”.² (Mack y cols., 2005)

La entrevista en profundidad es una técnica que permite obtener la perspectiva de un individuo en concreto. Es un método efectivo para entender lo que el individuo siente, cuáles son sus opiniones y cuáles son sus experiencias. En adición, esta técnica nos permite atacar temáticas sensibles que un individuo, por diversas razones, no esté abierto a compartir mediante técnicas grupales. (Mack y cols., 2005)

El focus group (grupo focal) es un método cualitativo de recolección de datos donde uno o dos investigadores se reúnen con varios participantes para discutir un tema específico de investigación. Durante la sesión, un investigador actúa como moderador, guiando la discusión a través de preguntas abiertas que fomentan respuestas detalladas, mientras otro investigador toma notas detalladas de la conversación. Este método es especialmente útil para obtener una amplia gama de perspectivas en un corto período de tiempo, aunque no es el más adecuado para explorar temas personales o socialmente sensibles, para los cuales las entrevistas individuales suelen ser más efectivas. (Mack y cols., 2005)

El análisis de documentos es un método cualitativo que revisa y evalúa materiales tanto impresos como electrónicos para extraer significado, comprender un fenómeno y desarrollar conocimiento empírico. Este proceso sistemático implica encontrar, seleccionar, interpretar y sintetizar datos contenidos en documentos, organizándolos en temas y categorías relevantes. Los documentos que se analizan pueden incluir una amplia variedad de fuentes, como actas de reuniones, diarios, informes, recortes de periódicos y programas de eventos. Considerados “hechos sociales”, estos documentos reflejan la organización y el contexto en el que fueron producidos y utilizados. (Mack y cols., 2005)

Separación en fases del proceso de investigación cualitativa

Uno de los primeros desafíos que enfrentamos fue modelar las distintas fases del proceso de investigación cualitativa. Este tipo de investigación se caracteriza por su versatilidad y el énfasis en la interpretación de los investigadores, lo que genera un movimiento iterativo hacia adelante y atrás entre las distintas fases. La principal dificultad radica en equilibrar esta flexibilidad con la necesidad de estructurar el proceso en etapas definidas para su análisis y documentación.

¹Insider refiere a alguien que pertenece a cierto grupo

²Outsider refiere a alguien que no pertenece a cierto grupo

En este contexto, encontramos un marco de referencia en el artículo “Towards a Human-in-the-Loop Curation: A Qualitative Perspective” (Adorjan, Vargas-Solar, y Motz, 2022), donde se identifican cinco fases principales del proceso investigativo: Definición del problema, Recolección de datos, Gestión de los datos, Análisis de datos y Reporte. Estas fases se presentan en un esquema de circularidad bidireccional, enfatizando la naturaleza iterativa y cíclica del proceso (ver figura 2.1).

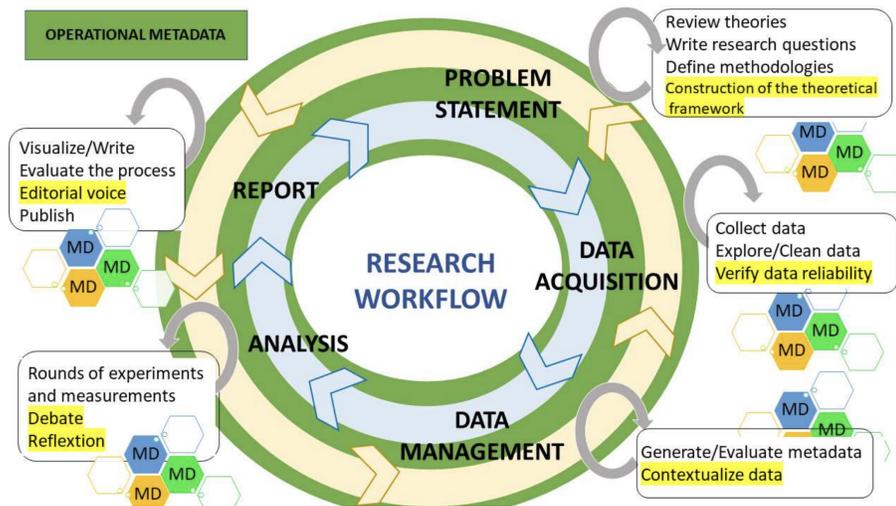


Figura 2.1: Flujo de fases del proceso investigativo, extraído de el artículo “Towards a Human-in-the-Loop Curation: A Qualitative Perspective”

El modelo propuesto organiza el proceso cualitativo en fases y permite integrar metodologías híbridas que combinan enfoques cualitativos y cuantitativos. Aunque el enfoque del artículo introduce un entorno de curación de datos orientado a registrar, mantener y enriquecer datos de investigación mediante metodologías mixtas, nuestro interés principal no radica en la curación de datos, sino en la estructura y representación de las fases del proceso investigativo.

Estas fases se apoyan en un view model que representa los datos generados en cada etapa, facilitando su manejo y exploración. Aunque se proponen herramientas y operadores para gestionar versiones de datos curados, nos enfocaremos más en las dinámicas y particularidades de las fases del proceso híbrido, con el objetivo de comprender cómo interactúan y se retroalimentan en proyectos de investigación cualitativa.

Capítulo 3

Desarrollo de la ontología

El capítulo central de este trabajo se enfoca en el desarrollo y aporte del proyecto de grado, abarcando tanto las decisiones tomadas como las justificaciones detrás de cada una de ellas. En este capítulo, se detallan los requerimientos identificados durante la etapa inicial del proyecto, así como el análisis que permitió estructurar una solución adecuada. Asimismo, se presenta el diseño conceptual y técnico de la solución, destacando las alternativas consideradas y la justificación de las elecciones finales. Este proceso incluye la definición de la ontología, la selección de tecnologías, y la forma en que la solución fue modelada para cumplir con los requerimientos.

3.1. Objetivo de la ontología

El propósito de esta sección es comprender la situación actual, el uso previsto para la ontología, los escenarios de aplicación y entender al usuario final involucrado. Para ello, realizamos un análisis del dominio en colaboración con Alén Pérez, nuestro usuario experto. A través de una serie de reuniones virtuales, logramos identificar y definir las preguntas de competencia, de manera que estas comprendiesen tanto las exigencias teóricas del modelo como los escenarios de uso práctico para los investigadores. Estas permiten estructurar de manera más precisa los elementos y relaciones que la ontología debe contemplar, asegurando así que cubra las necesidades reales de los investigadores. Cabe destacar que, además de las valiosas aportaciones de Alén, nos apoyamos en el libro *Metodología de la Investigación* (Sampieri, Collado, y del Pilar Baptista Lucio, 2014), el cual resultó fundamental para comprender ciertos conceptos clave relacionados con el dominio.

Aunque las preguntas de competencia son útiles para validar posteriormente la funcionalidad de la ontología, su objetivo principal es garantizar que los casos de uso más valiosos (desde el punto de vista del usuario experto) del proceso

de investigación cualitativa estén representados de forma clara y precisa en el modelo propuesto. A continuación, presentamos el conjunto de preguntas de competencia que nuestra ontología debe ser capaz de resolver. Estas preguntas cubren desde la aplicación de técnicas y la gestión de los datos, hasta la obtención de objetivos y hallazgos en el contexto de un proyecto de investigación cualitativa.

1. Dado un proyecto de investigación, nos interesa obtener el conjunto de técnicas aplicadas y qué información se obtiene a partir de su aplicación.
2. Dado un proyecto de investigación y una técnica aplicada, ¿sobre qué sujetos y/o objetos se aplican?
3. Dado un proyecto de investigación y una técnica aplicada, ¿en qué soporte se encuentran los datos?
4. Dado un proyecto de investigación, nos interesa obtener el conjunto de objetivos.
5. Dado un proyecto de investigación, obtener los esquemas de clasificación analíticos y qué información tienen asociados.
6. Dado un proyecto de investigación, obtener los esquemas de clasificación descriptiva y qué registros tienen asociados.
7. Dado un proyecto de investigación que aplica una o varias técnicas de tipo entrevista, obtener los núcleos temáticos, guía de entrevista y/o cuestionarios.
8. Dado un proyecto de investigación, obtener las reformulaciones de objetivos (si hubo) y/o supuestos iniciales.
9. Dado un proyecto de investigación, obtener los objetivos propuestos y/o alcanzados.
10. Dado un proyecto de investigación, obtener hallazgos principales o conclusiones.

En la actualidad, los investigadores enfrentan varios desafíos debido a la falta de herramientas específicas en su dominio para manejar el volumen y la complejidad de los datos cualitativos. Esto incluye dificultades para rastrear relaciones entre entrevistas, categorizar respuestas o seguir el análisis de los distintos datos obtenidos. El uso de herramientas genéricas, como hojas de cálculo o procesadores de texto, limita la capacidad de interconectar elementos, resultando en una mayor carga de trabajo y posibles errores. La falta de interoperabilidad semántica impide visualizar relaciones o realizar consultas eficientes sobre los datos. La ontología propuesta se presenta como una herramienta que permitirá a los investigadores gestionar mejor sus datos y acompañar el proceso durante

la investigación cualitativa. Adicionalmente, su diseño actuará como una herramienta didáctica para entender cómo sus investigaciones se alinean con el modelo.

En términos generales, buscamos centralizar la información para resolver problemas de gobernanza de datos, especialmente en investigaciones que manejan datos sensibles. Si bien este enfoque facilita la creación de ambientes de trabajo amigables para los investigadores, el alcance de nuestro trabajo se centra en la conceptualización y construcción de la ontología. Esto sienta una base estructurada para que, en etapas futuras, se desarrollen y utilicen estos ambientes de manera efectiva.

Además, el modelo debe ser reutilizable en diversas investigaciones, permitiendo a los investigadores evitar la necesidad de redefinir constantemente conceptos comunes, como el de “entrevista”. Se valora que los investigadores puedan acompañar el proyecto en sus distintas etapas, por lo que se busca modelar cada fase y sus elementos, pero sin la rigidez de asociar un estado a una etapa específica, ya que muchas veces estas fases se transitan de forma simultánea.

En conclusión, todas las preguntas de competencia parten del concepto de proyecto de investigación, que sirve como eje central para conectar los diferentes aspectos del proceso investigativo cualitativo. Estas preguntas se centran en el proyecto y abordan distintos elementos clave, como los objetivos, las técnicas aplicadas, los datos obtenidos y los resultados. Cada pregunta busca reflejar una dimensión particular del proyecto y cómo se interrelacionan los distintos componentes, como los sujetos, los objetos y las técnicas empleadas. De esta manera, se intenta garantizar que la ontología pueda representar correctamente las relaciones y procesos que ocurren durante una investigación cualitativa.

Además, se observa que estas preguntas parecen orientarse más hacia proyectos ya completados, en lugar de aquellos en desarrollo. Dado que todas asumen que el proyecto ya está finalizado, su enfoque tiende a recorrer y explorar sus distintas dimensiones, facilitando el análisis y consulta de la información registrada. Esto sugiere que la ontología podría ser especialmente útil como herramienta para acceder y revisar proyectos finalizados, permitiendo examinar sus diferentes aspectos en detalle.

No obstante, es importante destacar que, aunque la ontología no está diseñada específicamente para capturar el estado del proyecto, también puede servir como una base sólida para un marco de gestión de proyectos cualitativos. Esta capacidad de estructurar el proyecto a lo largo de sus fases y elementos resulta valiosa no solo para los proyectos completados, sino también para aquellos en desarrollo, ayudando a los investigadores a organizar y gestionar la información desde el inicio del proceso. La falta de una distinción clara entre proyectos en curso y proyectos finalizados puede ser una limitación, pero no invalida la capacidad de la ontología para ser una herramienta útil en ambos casos.

3.2. Diseño y construcción

El objetivo de esta sección es documentar el proceso de diseño y construcción de la ontología, proporcionando una guía comprensiva sobre cada etapa que contribuyó al diseño y construcción final, de manera de comprender las consideraciones conceptuales y metodológicas que fundamentan la ontología y aseguran su consistencia y utilidad dentro del contexto de estudio. Al exponer este proceso, buscamos demostrar cómo las distintas fases, desde la planificación y conceptualización inicial hasta la implementación y ajuste final, se alinearon con los requerimientos y objetivos propuestos y garantizan una representación estructurada y precisa del dominio de conocimiento.

En esta sección, presentamos el diseño final de nuestra ontología mediante un modelo conceptual, construido a partir de una serie de iteraciones que permitieron un refinamiento en diferentes aspectos para asegurar una ontología coherente y libre de redundancias. En primer lugar, se detalla el diseño de clases y relaciones, describiendo el proceso de identificación y organización de las principales entidades y conexiones, junto con las propiedades, criterios y restricciones aplicadas. Finalmente, se detallan aspectos clave de la construcción de la ontología, incluyendo la herramienta empleada para su desarrollo y la metodología aplicada durante las etapas de diseño y construcción.

3.2.1. Diseño de la ontología

En esta sección presentamos el modelo conceptual que captura los aspectos más relevantes del dominio, según el alcance definido. Aunque nos enfocamos en el modelo final, destacamos cómo algunas decisiones clave tomadas durante el proceso de diseño impactaron directamente en la estructura del modelo, sentando las bases para construir una ontología alineada con los objetivos del proyecto.

Como metodología de desarrollo adoptamos el enfoque bottom-up (ver sección 2.1.1), priorizando la adaptación de la ontología a partir del conocimiento específico del dominio ya disponible. Uno de los mayores impedimentos de este enfoque suele ser la dificultad para adquirir dicho conocimiento, lo cual fue uno de los principales desafíos del proyecto. Sin embargo, contamos con la colaboración cercana del docente Alén Pérez, nuestro usuario experto que, mediante iteraciones, nos ayudó a modelar el dominio de manera más precisa. Partimos de una definición preliminar del dominio realizada por las docentes Regina Motz y Edelweis Rohrer en conjunto con Alén Pérez. Si bien no entramos en detalle acerca de esta definición y su modelo conceptual, es importante destacar este trabajo ya que nos brindó una base sólida para comenzar nuestras iteraciones hasta llegar a la versión final.

Para asegurar un resultado de calidad y coherente con el enfoque bottom-up adoptado, nuestra metodología de trabajo se basó en una planificación y diseño antes de entrar en la implementación. Nos tomamos el tiempo necesario para modelar esquemas conceptuales detallados que capturaran las particularidades del dominio. Esta inversión inicial fue clave para establecer una base sólida, en la que cada elemento y relación estuviera bien fundamentado, evitando así la necesidad de ajustes o soluciones improvisadas durante el desarrollo. Gracias a este enfoque anticipado, logramos un modelo final simple y coherente, en el cual no fue necesario realizar grandes correcciones imprevistas o ajustes posteriores.

El esquema de la figura 3.1 representa el modelo final actualmente implementado, el cual fue utilizado durante el proceso de validación descrito en las secciones posteriores. A modo general, los cuadrados con nombres se corresponden con clases, mientras que las flechas representan relaciones entre ellas. Otros elementos, como la nube correspondiente al Proceso de interpretación, tienen un carácter más conceptual. Además, los cuadrados en color celeste indican clases intermedias que facilitan la representación de relaciones ternarias, mientras que las flechas azules representan dichas relaciones. Por último, las flechas de color rojo no indican relaciones sino una categorización en subclases.

Un elemento central en este esquema es el Proceso de interpretación, ubicado estratégicamente en el centro del diagrama. Este componente, que resultó ser el más complejo de capturar y representar durante las iteraciones, se convirtió en el eje articulador del modelo. Una vez identificado, su posición central permitió estructurar y conectar los demás elementos de manera coherente, sirviendo co-

mo punto de partida para la construcción del modelo conceptual final.

La posición central del Proceso de interpretación en el modelo no solo responde a razones estructurales, sino que también es coherente con la importancia fundamental que este concepto tiene en nuestro dominio, tal como lo expresó nuestro usuario experto Alen Pérez. La interpretación de los investigadores es el eje clave del proceso de investigación, ya que ellos analizan y otorgan significado a los datos obtenidos de la realidad desde su perspectiva. Este carácter interpretativo agrega un nivel de complejidad significativo a la representación, ya que limitar la interpretación mediante restricciones estrictas en el modelo podría comprometer tanto la efectividad del proceso investigativo como las capacidades de los investigadores que empleen la ontología. Por ello, al ser un elemento central, también orientó nuestro diseño hacia la búsqueda de la mínima cantidad de restricciones posibles en el resto de la implementación, garantizando así un balance entre estructura y flexibilidad.

Luego, en el diagrama, observamos tres grandes grupos a los que pertenecen la mayoría de los elementos del proceso investigativo: Formulación, Campo y Análisis. Estos grupos representan las tres fases principales del proceso de investigación cualitativa, cuya denominación y división surgen del trabajo realizado por Alejandro Adorjan, Genoveva Vargas-Solar y Regina Motz en el artículo *Towards a Human-in-the-Loop Curation: A Qualitative Perspective* (Adorjan y cols., 2022). Este trabajo, presentado con mayor detalle en la sección 2.2, realiza una división en las siguientes fases: Definición del problema, Recolección de datos, Gestión de los datos, Análisis de datos y Reporte.

Considerando el trabajo de Adorjan, Vargas-Solar y Motz, así como los intercambios realizados con nuestro usuario experto Alen Pérez, identificamos que la Definición del problema se transforma en la fase de Formulación; Recolección de datos y Gestión de los datos se unifican en la fase de Campo; y Análisis de datos, junto con Reporte, se agrupan en la nueva fase de Análisis. Estas categorías principales, representadas por las superclases Formulación, Campo y Análisis en la implementación, desempeñan un papel clave en la estructuración del modelo. Formulación incluye los elementos relacionados con la definición del problema y la planificación del proyecto; Campo abarca las actividades de recolección y gestión de datos realizadas en el contexto del trabajo de campo; y Análisis se enfoca en los procesos de interpretación de datos y generación de resultados.

Además, estas superclases están diseñadas con una visión a futuro. En caso de que el modelo evolucione hacia una red de ontologías más especializada por fase, estas categorías funcionarían como delimitadores del alcance de cada ontología, permitiendo una expansión modular y manteniendo la cohesión entre los diferentes elementos de la realidad. Vemos, además, que en este modelo no existe una transición ni linealidad fija entre las fases, pues entendimos la dificultad de ubicar el proyecto en una fase concreta en un momento dado. Esto permite un enfoque más fluido y menos fragmentado, logrando representar el proceso de

investigación de forma más integrada. Los nuevos grupos reflejan la progresión del trabajo de campo y su conexión con el análisis, donde la interpretación se convierte en el eje articulador, asegurando así una visión continua y cohesiva del desarrollo investigativo.

Por otro lado, se implementaron relaciones de superclase-subclase como solución para modelar las distintas técnicas, sujetos u objetos, soportes y registros. Esta elección permite que dichos elementos sean extensibles a futuro, una característica destacada como fundamental por nuestro usuario experto durante el proceso de diseño del problema. Esta estructura facilita, por ejemplo, la posibilidad de expandir el conjunto de técnicas soportadas por la ontología o de incluir nuevos tipos de registros en el futuro, garantizando así la adaptabilidad y la evolución del modelo conforme surjan nuevas necesidades o requisitos en el dominio.

En la misma línea, durante el proceso se consideró la opción de modelar a los investigadores y participantes utilizando alguna ontología especializada en el modelado de personas. Esto habría permitido una representación más detallada de dichos elementos, pero finalmente quedó fuera del alcance de esta versión inicial de la ontología en base a nuestras conversaciones con el usuario experto. No obstante, nuestra representación actual permite que estos elementos sean fácilmente reemplazables por representaciones provenientes de dichas ontologías especializadas, en caso de que se desee realizar esta ampliación en el futuro.

Otro elemento que creemos importante a destacar es el de Reformulación. Este concepto se diseña para permitir que cualquier elemento del Análisis pueda llevar a reformular alguna parte del proyecto, es decir, volver a la fase de formulación, logrando así una mayor flexibilidad en la representación. Esta capacidad se implementa a través de la estructura de subclases, facilitando una relación dinámica entre ambos procesos.

En conclusión, el proceso de construcción bottom-up fue fundamental para capturar el conocimiento específico del dominio con precisión. La colaboración estrecha con el usuario experto, Alén Pérez, resultó esencial para desarrollar un modelo detallado y fundamentado que respondiera a las necesidades concretas del dominio. Su experiencia guió decisiones clave y aportó claridad conceptual, lo que permitió un refinamiento continuo del modelo. Este enfoque nos permitió establecer una base sólida y facilitó un desarrollo progresivo y alineado con los requerimientos prácticos de la ontología. El enfoque en las preguntas de competencia desempeñó un papel crucial al proporcionar una guía clara y facilitar una evaluación continua de los avances, asegurando que la ontología fuera relevante y adecuada al contexto. Así, la colaboración con el experto y la orientación a través de las preguntas de competencia se consolidaron como pilares esenciales en el desarrollo, resultando en una construcción coherente y alineada con los objetivos planteados.

Finalmente, logramos generar una conceptualización del proceso de inves-

tigación cualitativa que es lo suficientemente detallada para facilitar su comprensión y aplicación, al tiempo que mantiene un enfoque no estructurado y adaptable que no restringe la libertad de interpretación de los investigadores que la utilicen.

3.2.2. Construcción de la ontología

Para desarrollar la ontología de este proyecto utilizamos Protégé ([Stanford Center for Biomedical Informatics Research, 2024](#)), una plataforma ampliamente reconocida en la comunidad de modelado ontológico. Protégé destaca por su arquitectura modular y su capacidad de ampliación mediante plug-ins, lo que permite construir aplicaciones ontológicas tanto simples como complejas. Un aspecto importante en la elección de Protégé es su total compatibilidad con los estándares del World Wide Web Consortium (W3C), como OWL (Web Ontology Language) y RDF (Resource Description Framework). Esta compatibilidad asegura que las ontologías creadas en Protégé cumplan con las especificaciones actuales y puedan interoperar con otros sistemas y aplicaciones basadas en ontologías. Así, los proyectos desarrollados en Protégé están alineados con las prácticas y protocolos estándar de la web semántica, ampliando su aplicabilidad y utilidad.

Si bien anteriormente mencionamos que buscamos un modelo que sea lo menos restrictivo posible para no interferir en el proceso de interpretación, la ontología desarrollada contiene axiomas y restricciones que permiten estructurar los datos y habilitar la inferencia automática de conocimiento. A continuación, se describe la implementación de dominios, rangos, relaciones fundamentales y otros axiomas claves, con un enfoque en las capacidades de inferencia y validación.

Relaciones Fundamentales: Definición de Dominios y Rangos

En la ontología, las relaciones que consideramos fundamentales son aquellas que articulan su estructura principal y generan los caminos necesarios para conectar las distintas instancias dentro de un mismo proyecto. Estas relaciones son esenciales porque permiten vincular conceptos clave, facilitando la recuperación de información relevante para responder a las preguntas de competencia definidas en el proyecto. En otras palabras, estas relaciones no solo organizan el conocimiento dentro de la ontología, sino que también habilitan su capacidad para soportar consultas específicas que reflejan necesidades reales del dominio.

Con el objetivo de mantener la ontología clara y accesible, se decidió implementar relaciones simples, es decir, propiedades directas que tienen dominios y rangos claramente establecidos, lo que asegura que las conexiones entre las instancias sean coherentes con las clases a las que pertenecen.

Relaciones funcionales e inversas

En el diseño de esta ontología, se tomó una decisión consciente de no incluir restricciones estrictas asociadas a las propiedades, como relaciones inversas, funcionales, simétricas o transitivas. Esta decisión responde a la necesidad de priorizar la libertad de los usuarios al cargar sus investigaciones, permitiéndoles modelar la información de acuerdo con sus necesidades específicas. El dominio modelado es intrínsecamente variable y dinámico, y establecer axiomas restrictivos podría haber limitado su flexibilidad, dificultando la representación de casos particulares que no fueran considerados en la fase de diseño.

Adicionalmente, se consideró que muchos usuarios potenciales de esta ontología podrían no estar familiarizados con conceptos avanzados de lógica descriptiva o modelado ontológico. En muchos casos, estos usuarios interactuarán con la ontología únicamente a través de aplicaciones diseñadas para simplificar su uso, sin necesidad de comprender conceptos de lógica descriptiva o de construcción de ontologías. Por ejemplo, un investigador cualitativo podría usar una interfaz para registrar información sobre un proyecto, técnicas o resultados, sin tener conocimiento explícito sobre propiedades de una relación como "funcionalidad." "simetría". Introducir restricciones ontológicas que requieran conocimiento avanzado podría resultar en confusión o errores al cargar datos, ya que los usuarios no estarían preparados para resolver inconsistencias detectadas por el razonador.

Por ejemplo, si una propiedad se definiera como funcional y un usuario intentara asociar múltiples valores a un recurso, el razonador marcaría esto como inconsistente. Para alguien sin conocimientos sobre ontologías, esta restricción puede ser percibida como un error del sistema o una limitación innecesaria, generando frustración y desincentivando el uso de la herramienta. Este escenario subraya la importancia de mantener la ontología lo suficientemente flexible como para adaptarse a un público que no necesariamente comprende los fundamentos de la lógica descriptiva. Por este motivo, en lugar de imponer restricciones estrictas como relaciones inversas o funcionales, se optó por permitir que los usuarios carguen y conecten instancias con la mayor libertad posible. Esto asegura que la ontología sea accesible y útil incluso para aquellos sin experiencia técnica, evitando que el sistema presente barreras adicionales al modelado de información.

Por otro lado, los razonadores empleados (como Pellet o HermiT) aseguran la consistencia del modelo y generan relaciones adicionales a partir de los axiomas definidos. Por ejemplo, si un usuario asocia una técnica a un proyecto y esta técnica está vinculada a un investigador, el razonador puede deducir automáticamente que dicho investigador participa en el proyecto, incluso si esto no se declaró explícitamente.

En resumen, la decisión de no incluir relaciones inversas, funcionales, simétri-

cas u otras restricciones estrictas responde tanto a la necesidad de ofrecer flexibilidad a los usuarios como a la intención de hacer que la ontología sea accesible incluso para personas sin conocimientos técnicos avanzados. Al mismo tiempo, la incorporación de axiomas de composición asegura la completitud de los datos y maximiza la potencia de inferencias del modelo, permitiendo descubrir conocimiento implícito y simplificando el trabajo de los usuarios. Este enfoque garantiza que la ontología sea una herramienta adaptable, accesible y eficiente para la representación y análisis de conocimiento en un dominio complejo y dinámico.

Axiomas de composición de relaciones

Si bien hicimos énfasis en la flexibilidad de nuestro diseño, esta flexibilidad no implica una falta de estructura o de mecanismos para garantizar la completitud de los datos. Para abordar este aspecto, se incorporaron axiomas de composición, que son fundamentales para mantener la coherencia y maximizar el potencial de inferencia de la ontología. Estos axiomas permiten deducir relaciones implícitas entre los elementos del modelo, incluso si los usuarios no las declaran explícitamente.

A continuación, presentamos las relaciones compuestas más relevantes introducidas en la ontología. En primer lugar, tenemos la relación “tiene esquema de clasificación descriptiva” la cual se compone con la relación “categoría superior” (ver figura 3.2). Esta composición permite agrupar distintas instancias de interpretación de un registro (y, por lo tanto, las instancias de la clase Registro asociadas) bajo una jerarquía de esquemas de clasificación descriptiva, evitando la necesidad de crear relaciones explícitas para cada nivel de la jerarquía. Por ejemplo, si E1, E2 y E3 son tres instancias de esquemas en orden jerárquico (donde E3 tiene una categoría superior a E2, y E2 a E1), al asignar la relación “Tiene esquema de clasificación descriptiva.” a una instancia de interpretación de un registro con E3, no sería necesario repetir esta relación con E2 ni con E1, lo cual reduciría exponencialmente la carga de trabajo a medida que aumentan los niveles de los esquemas. La relación definida permite inferir automáticamente las relaciones necesarias con los niveles superiores.

Por otro lado, tenemos otra relación compuesta en la relación “responde a entrevista” (ver figura 3.3, esta composición de relaciones ilustra cómo simplificar la carga de información para los usuarios. Al tener una relación trilateral entre Entrevista, Pregunta y Respuesta a pregunta, sería necesario, para cada registro de respuesta, relacionarlo manualmente con su pregunta y su entrevista. La composición permite declarar solo una parte de esta información, y dejar que los procesos de inferencia completen el resto. En este caso, se definieron dos composiciones diferentes para la misma relación, garantizando que la información en el proyecto permanezca consistente y completa, sin importar el criterio de carga utilizado.

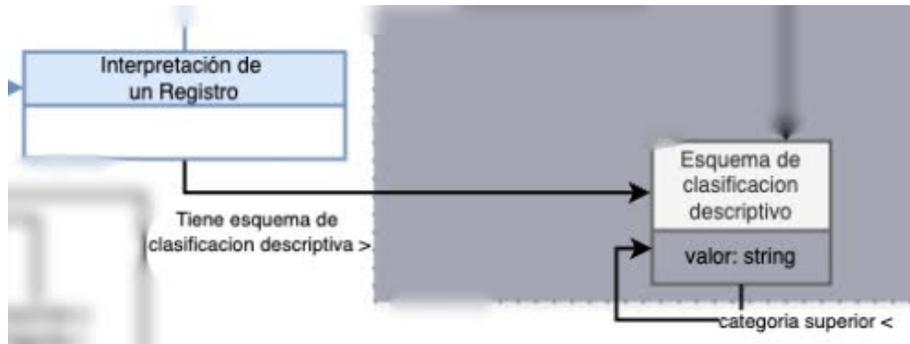


Figura 3.2: Relación compuesta “tiene esquema de clasificación descriptiva”

Impacto de la Composición en la recuperación de información

Además de las ventajas que ofrecen la validación y la inferencia, los axiomas de composición también simplifican significativamente las consultas realizadas sobre la ontología. Gracias a las reglas de composición, las relaciones entre las distintas instancias se consolidan en caminos bien definidos, lo que elimina la necesidad de que las consultas aniden múltiples patrones o recorran diversas clases de manera explícita.

Por ejemplo, una consulta compleja que originalmente requeriría unir manualmente varios pasos intermedios para conectar un proyecto con sus resultados puede ser simplificada drásticamente. Al aplicar las inferencias derivadas de las composiciones, el sistema ya ha recorrido y concatenado las relaciones necesarias. Esto permite que las consultas se reduzcan a una simple declaración de interés, confiando en que el razonador ya ha realizado las conexiones subyacentes. Este enfoque no solo reduce la complejidad técnica de las consultas, sino que también mejora la eficiencia y la legibilidad del proceso de recuperación de información.

En conclusión, podemos observar cómo, a través del uso de propiedades y relaciones especiales, como la composición, logramos aprovechar las capacidades que ofrece una ontología. Si bien hacemos uso de estas capacidades, lo hacemos desde un enfoque poco restrictivo, permitiendo flexibilidad en su aplicación pero manteniendo una estructura bien definida. Esta estructura está enmarcada por clases y propiedades con dominios y rangos cuidadosamente especificados, asegurando coherencia y claridad en el modelo.

La ontología desarrollada está disponible en TechnoPortal, un repositorio de acceso abierto de ontologías, accesible en la siguiente url: <https://technoportal.hevs.ch/ontologies/ONTINVCUALI>.

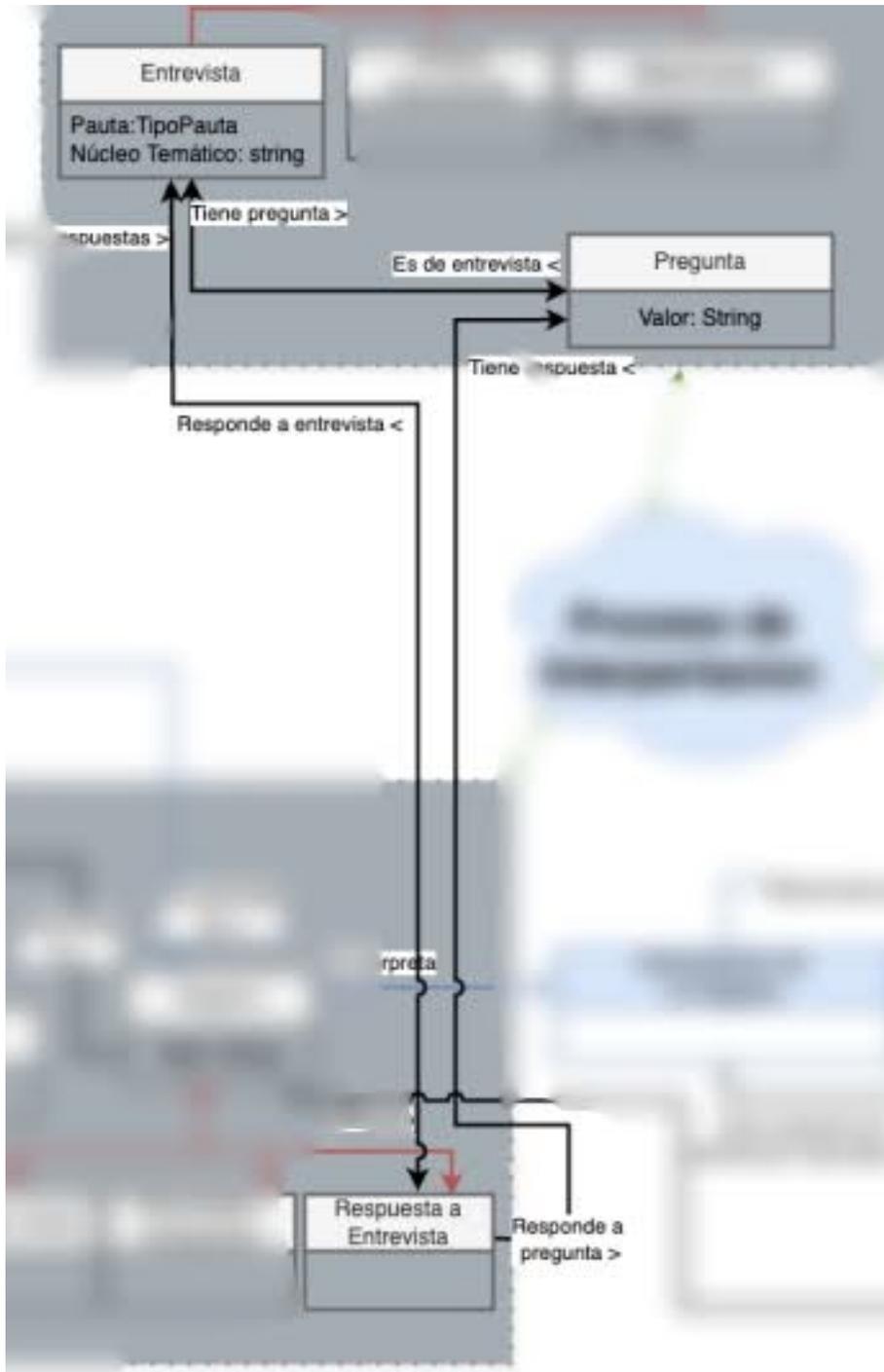


Figura 3.3: Relación compuesta “responde a entrevista”

Capítulo 4

Validación de la Ontología

Antes de realizar la validación principal, se llevó a cabo una verificación preliminar de la consistencia lógica de la ontología utilizando el razonador HermiT, provisto por Protégé. Este razonador implementa el algoritmo de Tableau, que determina si una ontología definida en lenguaje OWL es consistente. Esta verificación inicial asegura que la estructura lógica de la ontología es coherente y no presenta contradicciones fundamentales, sirviendo como un paso esencial en el proceso de validación. Esto se realiza simplemente ejecutando el razonador en Protégé y comprobando que este no arroje errores en su ejecución.

En este capítulo, nos centraremos en una validación más detallada de la ontología construida desde dos grandes aspectos. Primero, evaluaremos la ontología en términos de las buenas prácticas de diseño ontológico. Este análisis permite identificar posibles problemas de consistencia (lógica y ontológica), redundancia y calidad estructural, garantizando que la ontología cumple con lineamientos establecidos para su correcto desarrollo. Segundo, validaremos la ontología mediante las preguntas de competencia formuladas al inicio del proceso de desarrollo. Estas consultas verificarán si la ontología puede representar el conocimiento del dominio de manera adecuada y responder con precisión a las necesidades de los usuarios.

4.1. Validación de Buenas Prácticas

En esta sección, se aplican metodologías que permiten asegurar que la ontología sigue los principios fundamentales del diseño ontológico. Estas herramientas revisan aspectos clave como la taxonomía, la coherencia entre clases y la adecuación de las relaciones semánticas, para evitar errores comunes que pueden afectar la calidad del modelo.

Primero, se utilizará la herramienta OOPS (Ontology Pitfall Scanner), que realiza una revisión automatizada de la ontología para detectar anti-patronnes

previamente definidos. Estos incluyen problemas comunes como la creación de relaciones incorrectas, la ausencia de dominios o rangos en propiedades, y la inclusión de ciclos en la jerarquía. Aunque resulta útil para encontrar errores estructurales, OOPS no aborda las deducciones indeseables desde una perspectiva más profunda del modelado. Por ejemplo, una deducción indeseable podría ocurrir si, debido a una relación mal definida, el razonador concluye que dos conceptos son equivalentes cuando en realidad deberían ser distintos, o si genera jerarquías que contradicen el conocimiento del dominio.

Además, se realizará un análisis basado en OntoClean para evaluar la coherencia ontológica, centrándose en la correcta jerarquización de clases según principios filosóficos como la rigidez y la dependencia. OntoClean garantiza que las clases y sus propiedades respeten estos principios, aunque tiene limitaciones para resolver problemas relacionados con la estructura de propiedades o deducciones complejas. (Keet, 2020)

4.1.1. Ontology Pitfall Scanner (OOPS!)

OOPS (Ontology Pitfall Scanner) es una herramienta desarrollada por el Ontology Engineering Group, perteneciente a la Escuela Técnica Superior de Ingeniería Informática de la Universidad Politécnica de Madrid. (Ontology Engineering Group (OEG), 2024) Es una herramienta web independiente de cualquier entorno de desarrollo ontológico, creada para identificar posibles errores o "pitfalls" que podrían ocasionar problemas de modelado. Su objetivo es asistir a los desarrolladores de ontologías en el proceso de validación, que incluye tanto la fase de diagnóstico como la de reparación. Actualmente, OOPS! ofrece mecanismos automáticos para la detección de una serie de "pitfalls", contribuyendo a mejorar la precisión y calidad de las ontologías. (Poveda-Villalón, Gómez-Pérez, y Suárez-Figueroa, 2014) Los "pitfalls" o errores son categorizados en tres niveles de severidad: Críticos, Importantes y Menores. (ver figura 4.1) Estos niveles permiten priorizar los problemas detectados y enfocar los esfuerzos de corrección en los aspectos más relevantes para garantizar la calidad del modelo ontológico.

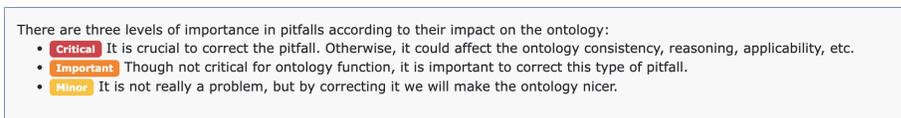


Figura 4.1: OOPS: Categorías de pitfalls

A continuación en la figura 4.2, presentamos los resultados obtenidos al aplicar OOPS! a nuestra ontología, destacando los tipos de errores identificados en cada categoría y analizando las posibles implicaciones de estos en el diseño y la estructura de nuestra ontología. Cabe mencionar que las "pitfalls" se identifican

según el prefijo P seguido de un número específico, que corresponde a un tipo particular de error o advertencia en la ontología. Este sistema de identificación facilita su clasificación y seguimiento durante el proceso de corrección y mejora.

Pitfalls detected:		
Results for P04: Creating unconnected ontology elements. 1 case	1 case	Minor
Results for P13: Inverse relationships not explicitly declared.	26 cases	Minor
Results for P22: Using different naming conventions in the ontology.	Ontology*	Minor

Figura 4.2: OOPS: Resultados del análisis de nuestra ontología

En primer lugar, se detectó **“P04: Creating unconnected ontology element”**, un problema menor que señala la existencia de elementos no conectados adecuadamente con el resto de la ontología. Este tipo de desconexión puede indicar la presencia de conceptos o clases que están aislados, lo que podría afectar la integración y el uso eficiente de la ontología. Si examinamos con más detalle a partir de los datos que nos provee la herramienta (ver figura 4.3), podemos observar que este elemento desconectado se trata de la clase campo:

Results for P04: Creating unconnected ontology elements. 1 case		1 case	Minor
Ontology elements (classes, object properties and datatype properties) are created isolated, with no relation to the rest of the ontology.			
• This pitfall appears in the following elements:			
• http://www.semanticweb.org/tesis_inv_cualitativa#campo			

Figura 4.3: OOPS: Detalle de pitfall P04

Como se menciona en la sección 3.2.1, la clase Campo actúa como una superclase que representa una fase específica del proceso de investigación, agrupando diversas subclases relacionadas. A diferencia de las fases de Formulación y Análisis, que están vinculadas entre sí mediante la reformulación, Campo no establece una relación directa con otros elementos del modelo. Dado que esta clase tiene un propósito más conceptual que práctico en nuestro contexto, consideramos que esta advertencia puede ser desestimada.

En segundo lugar, se identificaron 25 casos de **“P13: Inverse relationships not explicitly declared”**, un problema clasificado como menor que señala la ausencia de relaciones inversas explícitas. Las relaciones inversas pueden facilitar ciertas consultas al proporcionar rutas más directas para recuperar información,

como lo permite SPARQL a través de los property paths. En nuestra ontología, priorizamos la legibilidad y simplicidad, por lo que incluimos únicamente aquellas relaciones inversas que aportan claridad al modelo o simplifican consultas específicas, en función de las necesidades del dominio y las preguntas de competencia planteadas. Este enfoque asegura que la ontología sea manejable y comprensible sin comprometer su funcionalidad.

Otro aspecto detectado es **“P22: Using different naming conventions in the ontology”**, clasificado también como menor. La consistencia en las convenciones de nomenclatura es importante para asegurar la claridad y la mantenibilidad de la ontología, especialmente en entornos colaborativos o cuando esta crece en complejidad. En nuestro caso, dado que todas las entidades de nuestra ontología tienen etiquetas (labels), consideramos que este problema no añade demasiada complejidad en cuanto a la lectura y manejo de la ontología. Sin embargo, mantenemos presente la importancia de la uniformidad en futuras expansiones.

El análisis realizado con OOPS! nos ha permitido identificar y evaluar una serie de errores o “pitfalls” presentes en nuestra ontología. Como hemos visto, algunos de estos problemas, como la falta de relaciones inversas o elementos desconectados, fueron desestimados en el contexto de nuestra aplicación actual debido a diversas razones expuestas anteriormente. No obstante, el hecho de haber identificado estos problemas resalta la importancia de seguir las mejores prácticas de desarrollo ontológico, ya que en aplicaciones de mayor envergadura, como en entornos de producción o proyectos colaborativos, estos detalles pueden llegar a comprometer la eficiencia, la claridad y la capacidad de expansión del modelo. Las decisiones tomadas en nuestra ontología fueron orientadas por la necesidad de mantener la legibilidad y el propósito educativo del proyecto, asegurándonos, al mismo tiempo, de no perder la capacidad de responder las preguntas de competencia definidas al inicio.

En conclusión, el uso de OOPS! ha sido valioso para evaluar la calidad y consistencia de nuestra ontología, permitiéndonos tomar decisiones informadas sobre qué problemas abordar o ignorar según el contexto de aplicación. La corrección de estos “pitfalls” en el futuro contribuirá a asegurar que nuestra ontología no solo sea funcional para su propósito actual, sino que también pueda adaptarse y crecer en nuevos escenarios, manteniendo su calidad y robustez.

4.1.2. Análisis basado en OntoClean

OntoClean es una metodología utilizada para validar la corrección y consistencia de la taxonomía de una ontología desde un punto de vista filosófico. Aunque no aplicaremos el método completo, utilizaremos sus conceptos para evaluar nuestra ontología. A continuación, se describen los conceptos fundamentales de Rigidez, Identidad, Unidad y Dependencia.

La rigidez se refiere a las propiedades de una entidad que son esenciales para esa entidad, es decir, que deben ser verdaderas en cualquier situación posible. Por ejemplo, la propiedad de tener paredes es esencial para una casa. Toda casa debe tener paredes en cualquier situación posible. Si las paredes se demuelen, ya no tendríamos una casa. La rigidez ayuda a identificar las características que una entidad no puede perder sin dejar de ser ella misma. Por lo tanto, si existe una relación de subsunción entre dos propiedades, donde una propiedad z subsume a otra propiedad x , y z es anti-rígida, entonces x también debe ser anti-rígida. Esto implica que las propiedades no esenciales para una entidad en un nivel superior deben conservarse como no esenciales en sus niveles más específicos.

La identidad se refiere a la capacidad de identificar entidades individuales en el mundo como iguales o diferentes. Esto implica poder distinguir cuándo estamos tratando con la misma entidad o con entidades distintas. Por ejemplo, los criterios que usamos para determinar si dos instancias son la misma persona, objeto o concepto son aspectos de la identidad. Este concepto es importante para mantener la consistencia al definir y relacionar entidades dentro de una ontología. Si una propiedad superior, z , define un criterio de identidad, entonces cualquier propiedad x que sea subsumida por z debe conservar el mismo criterio de identidad.

La unidad se refiere a la capacidad de describir las partes y límites de los objetos, y por lo tanto, saber qué partes constituyen un objeto, cuáles no, y bajo qué circunstancias el objeto es un todo. Esto ayuda a definir qué es y qué no es parte de un objeto específico. La unidad es importante para entender cómo se estructuran y organizan las entidades complejas, permitiendo una representación más precisa de su composición interna y de sus límites. Si una propiedad z define un criterio de unidad, cualquier propiedad específica x también debe poseer ese mismo criterio de unidad. De igual manera, si z tiene anti-unidad, x también debe tener anti-unidad.

La dependencia es la relación entre entidades donde una existe únicamente por la existencia de la otra. Por ejemplo, un agujero depende de la existencia de una superficie para existir. Este concepto destaca las relaciones de dependencia existencial entre entidades, subrayando cómo la existencia de una entidad puede estar condicionada por la presencia de otra. Es decir que, si una propiedad z depende existencialmente de otra propiedad y , entonces cualquier propiedad x subsumida por z también debe depender de y .

Para asegurar la consistencia ontológica o conceptual de las relaciones de subclase en una ontología, OntoClean establece ciertas reglas conocidas como restricciones, que regulan la asignación de meta-propiedades a las entidades. Estas restricciones permiten garantizar que las relaciones entre clases y subclases se mantengan coherentes a lo largo de la ontología, evitando conflictos conceptuales y asegurando que el modelo represente adecuadamente el dominio. Es importante señalar que esta verificación es diferente de la que realizamos con

los algoritmos de razonamiento, como el razonador Hermit, que nos garantiza la consistencia lógica de la ontología.

Estas restricciones fortalecen la ontología al asegurar que la taxonomía conserve una estructura conceptual coherente. Aplicar estas reglas en conjunto con los conceptos de Rigidez, Identidad, Unidad y Dependencia permite una evaluación precisa de la ontología, y facilita que las propiedades y relaciones esenciales de las entidades estén bien definidas en toda la estructura. (Mahlaza y Keet, s.f.)

A continuación exponemos en la figura 4.4 la estructura de clases de la ontología desarrollada con el fin de demostrar el cumplimiento de las propiedades y reglas expuestas en los párrafos anteriores.

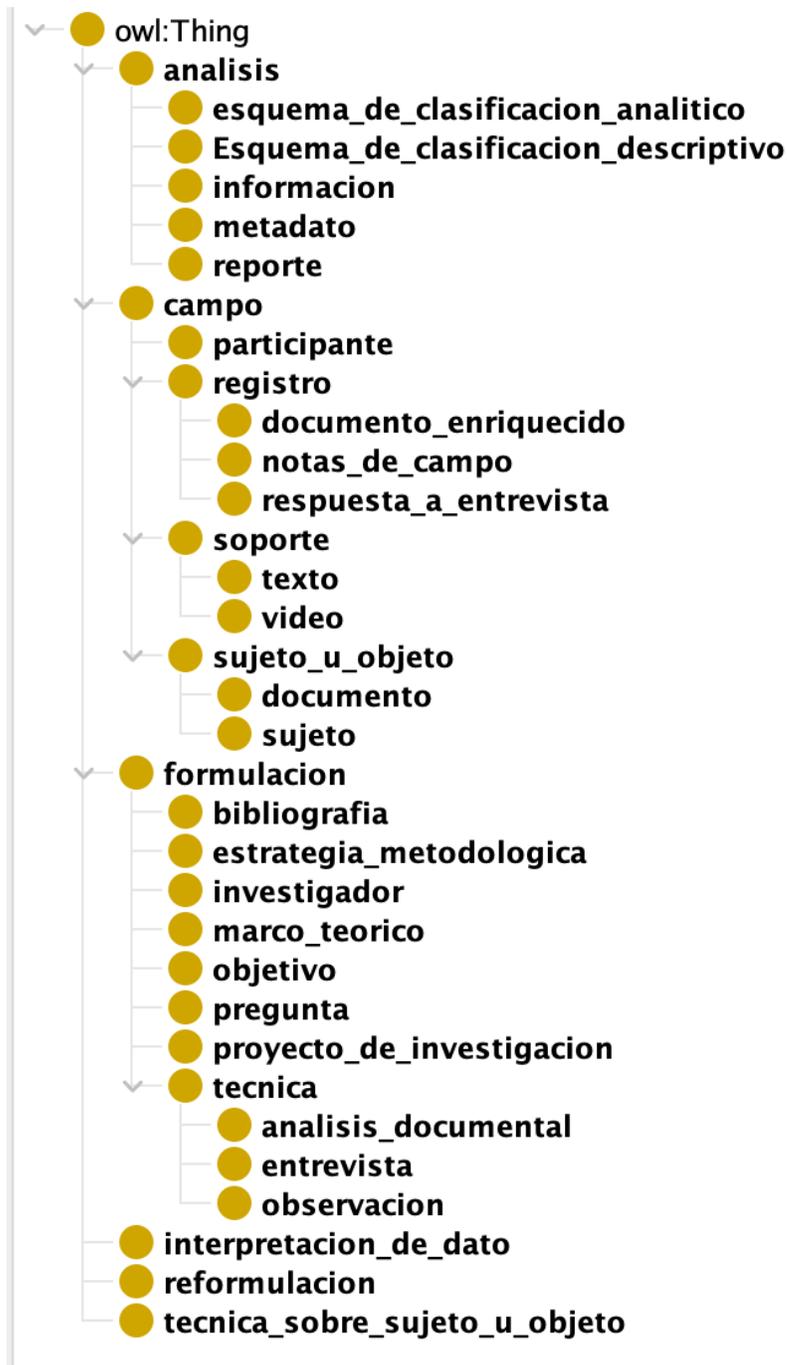


Figura 4.4: Estructura de clases de la ontología

Primero, consideremos las tres clases principales que representan las distintas fases del proceso investigativo: Análisis, Campo y Formulación. Estas clases son rígidas, ya que, aunque el proceso implica una transición entre fases, una instancia que pertenece a Análisis, Campo o Formulación siempre mantendrá su identidad dentro de esa fase. En otras palabras, una entidad que ha sido clasificada en una de estas fases no cambia de clase aunque el proceso investigativo avance; siempre será una instancia de Análisis, Campo o Formulación.

Siguiendo esta misma lógica, observemos el grupo de clases con subclases, como Registro, Soporte, Sujeto u Objeto y Técnica. Estas también pueden considerarse clases rígidas, ya que representan conceptos e instrumentos definidos de forma precisa en el contexto del proceso de investigación y no son susceptibles de cambio en cuanto a su esencia. Si bien podríamos pensar en Registro como una clase anti-rígida porque el registro de datos podría “convertirse” en información a través de la interpretación, es precisamente esta interpretación la que actúa como intermediaria y define una línea divisoria entre el registro y la información resultante. Esto permite que el dato, aunque sea interpretado de diversas maneras, mantenga su identidad como registro de datos, pues sigue siendo el mismo dato base aunque sus interpretaciones sean múltiples y variables.

4.2. Validación mediante Preguntas de Competencia

```

SELECT DISTINCT ?proyecto ?tecnica ?sujeto
WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
    ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Relacionado_a_tecnica_sobre_sujeto_u_objeto ?tssuo
    .
  ?tssuo ont:Relacionado_a_sujeto_u_objeto ?sujeto_u_objeto.
}
ORDER BY (?proyecto) (?tecnica) (?sujeto)

```

Listing 4.1: SPARQL query para la pregunta de competencia 3

```

SELECT DISTINCT ?proyecto ?tecnica ?sujeto
WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
    ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Se_aplica_en ?sujeto_u_objeto.
}
ORDER BY (?proyecto) (?tecnica) (?sujeto)

```

Listing 4.2: SPARQL query para la pregunta de competencia 3

En esta sección, la ontología será evaluada en función de su capacidad para responder preguntas específicas, conocidas como preguntas de competencia. Estas preguntas reflejan las consultas que los usuarios o investigadores podrían realizar sobre el dominio modelado y son clave para verificar si la ontología representa correctamente el conocimiento, proporcionando respuestas adecuadas de manera eficiente.

Para la validación de nuestro modelo ontológico cargamos instancias de las distintas clases basándonos en una investigación real de la Facultad de Información y Comunicación (FIC) sobre estrategias de estudio (Estudio exploratorio sobre las estrategias de estudio y modos de lectura en jóvenes universitarios de comunicación, realizado por Rosalía Winocur, Magela Cabrera, Soledad Morales y Camila Rojas). Trabajar con una investigación concreta nos permite comprobar que la ontología es capaz de representar el conocimiento generado en dicho contexto y que, mediante las preguntas de competencia, somos capaces de recuperar esa información. Este enfoque asegura que el modelo ontológico puede responder a consultas reales y relevantes para el dominio.

Cada pregunta de competencia está diseñada y validada en conjunto con nuestro usuario experto para simular escenarios de uso en los que los investigadores necesitan acceder a información. A través de consultas SPARQL, evaluaremos si la ontología puede ofrecer las respuestas deseadas para dichas preguntas.

Para facilitar la comprensión y simplificación del informe, a continuación se presentan los prefijos que serán utilizados en las consultas SPARQL posteriores. Estos prefijos permiten definir abreviaciones para las URI comunes que forman parte de la ontología y de las tecnologías relacionadas con RDF, OWL y XML Schema. Aunque estos prefijos son necesarios para ejecutar las consultas, y para evitar la repetición en cada ejemplo, los mencionamos al inicio de esta sección y asumimos que estarán implícitamente incluidos en todas las consultas mostradas más adelante.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX ont: <http://www.semanticweb.org/tesis_inv_cualitativa#>
```

Listing 4.3: SPARQL query prefixes

A continuación, se presentarán las preguntas de competencia junto con las consultas que las responden, los resultados obtenidos y un breve análisis de los resultados:

1. Dado un proyecto de investigación, nos interesa obtener el conjunto de técnicas aplicadas y qué información se obtiene a partir de su aplicación.

```
SELECT DISTINCT ?proyecto ?tecnica ?informacion
WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
    ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Relacionado_a_tecnica_sobre_sujeto_u_objeto ?
    tssuo.
  ?tssuo ont:Genera ?registro.
  OPTIONAL {
    ?registro ont:Se_interpreta ?interpretacion.
    ?interpretacion ont:Relacionado_a_informacion ?
      informacion.
  }
}
```

ORDER BY (?proyecto) (?tecnica) (?informacion)

Listing 4.4: SPARQL query para la pregunta de competencia 1

proyecto	ont:InvestigacionEstrategiaEstudio				
tecnica	ont:entrevista1	ont:entrevista2			
informacion	null	ont:informacion1	ont:informacion2	ont:informacion3	ont:informacion4 ont:informacion5

Figura 4.5: Resultados para la pregunta de competencia 1

Observamos que la consulta responde de manera satisfactoria, mostrando el proyecto, las técnicas aplicadas y, en los casos correspondientes, la información generada por dichas técnicas. En el caso de “entrevista1”, se muestra el valor “null”, lo que indica que esta técnica no generó información. Es importante recordar que las técnicas aplicadas sobre sujetos u objetos producen datos, y para que estos se “conviertan” en información, deben ser interpretados. Por lo tanto, la ausencia de información para “entrevista1” refleja que los datos generados no fueron interpretados, lo que explica el resultado.

2. Dado un proyecto de investigación y una técnica aplicada, ¿sobre qué sujetos y/o objetos se aplican?

```

SELECT DISTINCT ?proyecto ?tecnica ?sujeto ?descripcion
WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
    ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Se_aplica_en ?sujeto.
  ?sujeto ont:Descripcion ?descripcion.
}
ORDER BY (?proyecto) (?entrevista) (?sujeto)

```

Listing 4.5: SPARQL query para la pregunta de competencia 2

proyecto	ont:InvestigacionEstrategiaEstudio	
tecnica	ont:entrevista1	ont:entrevista2
sujeto	ont:sujeto1	ont:sujeto1
descripcion	Estudiantes de la asignatura Sociedad, Cultura y TIC, de la Carrera de Comunicación de la FIC.	Estudiantes de la asignatura Sociedad, Cultura y TIC, de la Carrera de Comunicación de la FIC.

Figura 4.6: Resultados para la pregunta de competencia 2

En este caso, observamos cómo se obtienen los sujetos u objetos sobre los cuales se aplican ambas técnicas (en este caso, los mismos para am-

bas). Además, se ha agregado la descripción de cada uno, ya que esta corresponde a una propiedad de datos de cada instancia. Esto nos permite enriquecer la consulta al incluir información adicional útil para el investigador que proporciona un contexto más detallado sobre los sujetos u objetos involucrados en la aplicación de las técnicas.

3. Dado un proyecto de investigación y una técnica aplicada, ¿en qué soporte se encuentran los datos que esta genera?

```
SELECT DISTINCT ?proyecto ?tecnica ?registro ?soporte
WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
    ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Relacionado_a_tecnica_sobre_sujeto_u_objeto ?
    tssuo.
  ?tssuo ont:Genera ?registro.
  ?registro ont:Soportado_en ?soporte.
}
ORDER BY (?proyecto) (?tecnica) (?registro) (?soporte)
```

Listing 4.6: SPARQL query para la pregunta de competencia 3

proyecto	ont:InvestigacionEstrategiaEstudio						
tecnica	ont:entrevista1			ont:entrevista2			
registro	ont:dato40	ont:dato41	ont:dato42	ont:dato44	ont:dato45	ont:dato46	ont:dato47
soporte	ont:soporte1	ont:soporte1	ont:soporte1	ont:soporte1	ont:soporte2	ont:soporte2	ont:soporte2

Figura 4.7: Resultados para la pregunta de competencia 3

A través de esta consulta, obtenemos los diferentes registros en los que se almacenan los datos generados por la aplicación de las técnicas sobre los sujetos u objetos correspondientes. Una vez identificados los soportes, es posible realizar consultas adicionales para obtener más detalles sobre ellos, lo que permite encadenar los resultados y acceder a la información necesaria de manera más completa. Cabe destacar que también se podría haber incluido esta información (nombre y ubicación de los soportes) directamente en la consulta anterior, pero se ha separado para ser fiel a la especificación de la pregunta de competencia. A continuación se busca el nombre y la ubicación de los soportes previamente encontrados:

```
SELECT DISTINCT ?soporte ?nombre ?ubicacion WHERE {
```

```

?soporte rdf:type ont:soporte.
?soporte ont:Nombre ?nombre.
?soporte ont:Ubicacion ?ubicacion.
  FILTER (?soporte IN (ont:soporte1, ont:soporte2))
}

```

Listing 4.7: SPARQL query para obtener nombre y ubicación de los soportes

soporte	nombre	ubicacion
ont:soporte1	"REFFDF~1"	"https://docs.google.com/document/d/1SLPe5xiSw0OnjHtq2sSkKesxWw0HZws0/edit"
ont:soporte2	"RE343F~1"	"https://docs.google.com/document/d/1FixEPvTI00QPCcpwJ_pUApwFctCvrheG/edit"

Figura 4.8: Resultados para la consulta acerca de los soportes

4. Dado un proyecto de investigación, nos interesa obtener el conjunto de objetivos.

```

SELECT DISTINCT ?proyecto ?objetivo WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
  ont:Tiene_objetivo ?objetivo.
}
ORDER BY (?proyecto) (?objetivo)

```

Listing 4.8: SPARQL query para la pregunta de competencia 4

proyecto	ont:InvestigacionEstrategiaEstudio		
objetivo	ont:objetivo1	ont:objetivo2	ont:objetivo3

Figura 4.9: Resultados para la pregunta de competencia 4

A partir de la consulta se obtienen los objetivos por proyecto de investigación. De forma análoga a la pregunta anterior, se podría realizar otra consulta para obtener mas información acerca de cada objetivo.

5. Dado un proyecto de investigación, obtener los esquemas de clasificación analíticos y qué información tienen asociados.

```

SELECT DISTINCT ?proyecto ?eca ?informacion WHERE {

```

```

?proyecto rdf:type ont:proyecto_de_investigacion;
  ont:Define ?estrategia.
?estrategia ont:Aplica_una_o_varias ?tecnica.
?tecnica ont:Relacionado_a_tecnica_sobre_sujeto_u_objeto ?
  tssuo.
?tssuo ont:Genera ?registro.
?registro ont:Se_interpreta ?interpretacion.
?interpretacion ont:Relacionado_a_informacion ?informacion
.
?informacion ont:Tiene_esquema_de_clasificacion_analitica ?
  eca.
}
ORDER BY (?proyecto) (?eca) (?informacion)

```

Listing 4.9: SPARQL query para la pregunta de competencia 5

proyecto	ont:InvestigacionEstrategiaEstudio				
eca	ont:eca1	ont:eca1.1	ont:eca2	ont:eca3	ont:eca4
informacion	ont:informacion1	ont:informacion1	ont:informacion3	ont:informacion4	ont:informacion5

Figura 4.10: Resultados para la pregunta de competencia 5

Esta consulta presenta una estructura muy similar a la de nuestra primera pregunta de competencia, ya que sigue el mismo recorrido hacia la obtención de la información generada. Si deseamos acceder a la información relacionada con los esquemas de clasificación analítica (eca) obtenidos, podemos ejecutar la siguiente consulta:

```

SELECT DISTINCT ?eca ?valor WHERE {
  ?eca rdf:type ont:esquema_de_clasificacion_analitico.
  ?eca ont:Valor ?valor.
  FILTER (?eca IN (ont:eca1, ont:eca1.1, ont:eca2, ont:eca3,
    ont:eca4))
}

```

Listing 4.10: SPARQL query para obtener informacion relacionada con los eca

soporte	nombre	ubicacion
ont:soporte1	"REFFDF~1"	"https://docs.google.com/document/d/1SLPe5xiSw00njHtq2sSkKesxWw0HZws0/edit"
ont:soporte2	"RE343F~1"	"https://docs.google.com/document/d/1FixEPvTI00QPcCpwJ_pUApwFctCvrheG/edit"

Figura 4.11: Resultados para la consulta acerca de los soportes

6. Dado un proyecto de investigación, obtener los esquemas de clasificación descriptiva y qué registros tienen asociados.

```
SELECT DISTINCT ?proyecto ?ecd ?registro WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
  ont:Define ?estrategia.
  ?estrategia ont:Aplica_una_o_varias ?tecnica.
  ?tecnica ont:Relacionado_a_tecnica_sobre_sujeto_u_objeto ?
    tssuo.
  ?tssuo ont:Genera ?registro.
  ?registro ont:Se_interpreta ?interpretacion.
  ?interpretacion ont:
    Tiene_esquema_de_clasificacion_descriptiva ?ecd.
}
ORDER BY (?proyecto) (?ecd) (?registro)
```

Listing 4.11: SPARQL query para la pregunta de competencia 6

ont:InvestigacionEstrategiaEstudio							
ont:ecd6				ont:ecd6.1			
ont:dato44	ont:dato45	ont:dato46	ont:dato47	ont:dato44	ont:dato45	ont:dato46	ont:dato47
1	1	1	1	1	1	1	1

Figura 4.12: Recorte de resultados para la pregunta de competencia 6

En la figura 4.12 se muestra un extracto de la información generada, ya que en este caso obtuvimos un total de 64 triplas como resultado de nuestra consulta. Podemos observar cómo se asocian correctamente los esquemas de clasificación descriptiva, los cuales se derivan de la interpretación de los registros. Además, todos ellos cuentan con una propiedad de datos asociada, "Valor", que si bien no la incluimos en esta consulta puede utilizarse posteriormente para obtener más información sobre estos esquemas de clasificación descriptiva.

7. Dado un proyecto de investigación que aplica una o varias técnicas de tipo entrevista, obtener los núcleos temáticos, guía de entrevista y/o cuestionarios.

```

SELECT DISTINCT ?proyecto ?tecnica ?nucleo_tematico ?ent_obj
?pregunta WHERE {
?proyecto rdf:type ont:proyecto_de_investigacion;
ont:Define ?estrategia.
?estrategia ont:Aplica_una_o_varias ?tecnica.
?tecnica rdf:type ont:entrevista;
ont:Nucleo_tematico ?nucleo_tematico;
ont:Objetivo ?ent_obj;
ont:Tiene_pregunta ?pregunta.
}
ORDER BY (?proyecto) (?tecnica) (?pregunta)

```

Listing 4.12: SPARQL query para la pregunta de competencia 7

proyecto					
tecnica	ont:entrevista1				
ent_obj	Reconstruir la percepción que los y las estudiantes tienen acerca de su condición de ser jóvenes				
nucleo_tematico	Autopercepcion, Estudiantes, Juventud				
pregunta	ont:pregunta1	ont:pregunta2	ont:pregunta3	ont:pregunta4	ont:pregunta5

Figura 4.13: Recorte de resultados para la pregunta de competencia 7

ont:InvestigacionEstrategiaEstudio									
ont:entrevista2									
Indagar las prácticas de estudio y de lectura de los y las estudiantes de comunicación									
Practicas de estudio, practicas de lectura									
ont:pregunta6	ont:pregunta7	ont:pregunta8	ont:pregunta9	ont:pregunta10	ont:pregunta11	ont:pregunta12	ont:pregunta13	ont:pregunta14	ont:pregunta15

Figura 4.14: Recorte de resultados para la pregunta de competencia 7

En este caso, obtenemos la información deseada para las técnicas de tipo entrevista. Es importante destacar que, si fuera necesario, se podría modificar la consulta o realizar una nueva para obtener más detalles sobre cada una de las preguntas asociadas a cada entrevista.

8. Dado un proyecto de investigación, obtener las reformulaciones de objetivos (si hubo) y/o supuestos iniciales.

La reformulación es un concepto complejo en nuestro esquema, por lo que es importante analizar en detalle las posibles respuestas a esta pregunta de competencia. Si observamos el diagrama de nuestra realidad (ver figura 3.1), podemos ver que la reformulación se modela como una clase asociada mediante dos propiedades de objeto: una hacia el análisis y otra hacia la formulación. La ventaja de este diseño es que permite que cualquier elemento de la fase de análisis pueda reformular cualquier elemento de la fase de formulación. Atendiendo a la especificación de la pregunta, solo obtendremos aquellos casos en los que el destino de la reformulación sea la clase “Objetivo”.

En la investigación que hemos ingresado a la ontología, no contamos con instancias de reformulación, por lo que crearemos una con el fin de ilustrar los resultados para esta consulta.

```
SELECT DISTINCT ?proyecto ?reformulacion ?origen ?destino ?
  descripcion WHERE {
  ?origen rdf:type ont:analisis;
  ont:Origen_de_reformulacion ?reformulacion.
  ?reformulacion ont:Descripcion ?descripcion;
  ont:Destino_de_reformulacion ?destino.
  ?destino rdf:type ont:objetivo.
  ?proyecto rdf:type ont:proyecto_de_investigacion;
  ont:Tiene_objetivo ?destino.
}
ORDER BY (?proyecto) (?reformulacion)
```

Listing 4.13: SPARQL query para la pregunta de competencia 8

proyecto	reformulacion	origen	destino	descripcion
ont:InvestigacionEstrategiaEstudio	ont:reformulacion1	ont:informacion4	ont:objetivo3	"xx/xx/xx: A partir de la informacion se reformula objetivo para hacerlo mas preciso."

Figura 4.15: Resultados para la pregunta de competencia 8

A partir de la consulta, obtenemos el resultado mostrado en la figura 4.15, el cual detalla, por proyecto, las reformulaciones realizadas. Este resultado incluye el origen de la reformulación, su destino, y una descripción adicional generada en el momento de la reformulación. En el caso analizado, observamos cómo el elemento “informacion4” originó una reformulación del “objetivo3”, con una descripción que proporciona una visión del motivo o

contenido de dicha reformulación, facilitando una mayor comprensión del proceso.

9. Dado un proyecto de investigación, obtener los objetivos propuestos y/o alcanzados.

En esta pregunta, al igual que en la anterior, en el proyecto instanciado no contamos con una instancia de reporte previamente cargada que esté vinculada a los objetivos mediante la propiedad de objeto “Alcanza_objetivo”. Por ello, para poder demostrar la respuesta a esta consulta, agregaremos dicha instancia con el fin de ilustrar cómo se relacionan estos elementos.

```
SELECT DISTINCT ?proyecto ?objetivo_alcanzado WHERE {
  ?proyecto rdf:type ont:proyecto_de_investigacion;
             ont:Tiene_objetivo ?objetivo_alcanzado.
  ?reporte rdf:type ont:reporte;
            ont:Reporta_proyecto ?proyecto;
            ont:Alcanza_objetivo ?objetivo_alcanzado.
}
ORDER BY (?proyecto)
```

Listing 4.14: SPARQL query para la pregunta de competencia 8



proyecto	objetivo_alcanzado
ont:InvestigacionEstrategiaEstudio	ont:objetivo1

Figura 4.16: Resultados para la pregunta de competencia 9

Si comparamos con los resultados de la pregunta de competencia 4 (ver 4.9) vemos como de todos los objetivos propuestos solo uno de ellos es alcanzado.

10. Dado un proyecto de investigación, obtener hallazgos principales o conclusiones.

```
SELECT DISTINCT ?proyecto ?conclusiones WHERE {
  ?reporte rdf:type ont:reporte;
            ont:Reporta_proyecto ?proyecto;
            ont:Conclusiones ?conclusiones.
}
ORDER BY (?proyecto)
```

Listing 4.15: SPARQL query para la pregunta de competencia 8

proyecto	conclusiones
ont:InvestigacionEstrategiaEstudio	*Se concluye que..*

Figura 4.17: Resultados para la pregunta de competencia 10

Se puede observar como gracias a la relación entre el reporte y el proyecto de investigación accedemos fácilmente a las conclusiones y hallazgos principales del mismo.

4.2.1. Reflexiones finales de la validación

Finalmente, luego de presentar y ejecutar las consultas SPARQL para cada una de las preguntas de competencia, hemos comprobado cómo el modelo ontológico construido permite obtener los datos deseados de forma precisa y organizada. Cada consulta ha respondido de manera efectiva a las necesidades planteadas, demostrando que el diseño del modelo permite acceder a información clave relacionada con los proyectos de investigación, sus técnicas, resultados, y el conocimiento derivado de los datos generados e interpretados.

Esto valida que el modelo cumple con los objetivos establecidos y que es adecuado para representar la realidad que se desea modelar en la ontología. La capacidad del modelo para responder a estas preguntas fundamentales confirma su solidez y utilidad como herramienta para la exploración y el análisis de información en el contexto planteado. En conclusión, se ha logrado un modelo ontológico robusto y eficiente, capaz de satisfacer los requerimientos específicos y responder las consultas críticas que motivaron su construcción.

Capítulo 5

Carga y visualización de datos

En este capítulo, comenzamos describiendo el proceso de carga de datos en la ontología, detallando los distintos proyectos que se han incorporado como instancias en esta etapa inicial. Reflexionamos sobre la experiencia de realizar una carga manual de datos, analizando los desafíos y aprendizajes obtenidos. Este proceso nos ha permitido evaluar la flexibilidad de la ontología para representar proyectos cualitativos e identificar oportunidades de mejora en futuras implementaciones.

Luego, presentamos la herramienta desarrollada para visualizar los datos integrados en la ontología, con un enfoque centrado en los proyectos de investigación. Esta herramienta permite explorar las distintas instancias y relaciones de los datos cargados en la ontología por proyecto, además de facilitar consultas basadas en las preguntas de competencia, otorgando a los investigadores un acceso directo a la información de cada proyecto.

5.1. Proceso de Carga de Datos

En esta sección, detallamos el proceso de carga de datos de una investigación utilizando la plataforma Protégé. Nuestro enfoque se centra en cargar únicamente las instancias más relevantes necesarias para responder a las preguntas de competencia formuladas, demostrando así el potencial de la ontología en investigaciones reales.

Existen diversas formas de cargar datos o editar una ontología; sin embargo, hemos optado por utilizar Protégé. Esta herramienta no solo nos proporcionó la posibilidad de modelar nuestra ontología, sino que también nos permite crear instancias específicas de las clases definidas y asignarles las propiedades correspondientes. Dividiremos el proceso de carga en varias etapas: **preparación del entorno, creación de instancias, asignación de propiedades, creación**

de relaciones y validación de carga. A lo largo de esta sección, realizaremos un recorrido paso a paso en cada una de estas etapas, con el objetivo final de que esta estrategia de carga pueda ser replicada de forma exitosa para otros proyectos de investigación.

5.1.1. Preparación del Entorno

Antes de comenzar con la carga de datos, se debe preparar adecuadamente el entorno de trabajo. Esto implica descargar e instalar la última versión de Protégé desde su sitio oficial. ([Stanford Center for Biomedical Informatics Research, 2024](#)) Una vez instalado, es recomendable iniciarlo para familiarizarse con su interfaz y cargar el archivo `.rdf` de la ontología que se utilizará. Es importante comprender la estructura de la ontología y disponer de todos los datos e información necesarios antes de iniciar la creación de las instancias.

Si bien lo mencionamos anteriormente, el archivo `.rdf` de la ontología se puede obtener en el siguiente repositorio de acceso público: <https://technportal.hevs.ch/ontologies/ONTINVCUALI>

5.1.2. Creación de una Instancia

Para cargar una instancia, primero debemos identificar la clase a la cual pertenece. Esto implica seleccionar una de las clases definidas en nuestra ontología. Al ubicarnos en la pestaña *Individuals by Class*, podemos visualizar todas las clases disponibles en el panel izquierdo, como se muestra en la Figura 5.1. Si no visualizamos esta lista, es posible que haya ocurrido un error al cargar la ontología. En este caso recomendamos cerrar y abrir Protégé y volverlo a intentar.

Una vez seleccionada la clase en el panel izquierdo, nos dirigimos a la pestaña *Individuals* en el panel inferior y hacemos clic en el botón `+`. Veremos una ventana similar a la mostrada en la Figura 5.2. En el campo *Name*, asignamos un nombre a la instancia. Este nombre servirá para identificar la instancia; podemos utilizar números y símbolos, por ejemplo: `proyecto_nuevo_1`. Al confirmar, la instancia habrá sido creada exitosamente.

Es importante establecer convenciones de nomenclatura claras y consistentes para las instancias, ya que esto facilitará su identificación y manejo en etapas posteriores del proyecto. Recomendamos utilizar nombres descriptivos que reflejen el contenido o propósito de la instancia.

5.1.3. Asignación de Propiedades

Una vez creada la instancia, es necesario asignarle sus propiedades. Estas propiedades funcionan como campos que describen las características de la instancia y dependen de la clase a la que pertenece. Al seleccionar la instancia



Figura 5.1: Visualización de clases en Protégé

creada, en el panel derecho veremos la pestaña *Data Property Assertions*; allí, hacemos clic en el botón +.

Se abrirá una ventana que muestra todas las propiedades disponibles en el

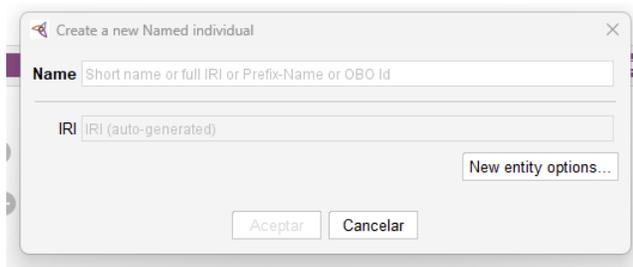


Figura 5.2: Creación de una nueva instancia

modelo. Debemos seleccionar las propiedades apropiadas, ya que no es posible asignar una propiedad no válida para la clase correspondiente. La lista de propiedades se muestra en la Figura 5.3.

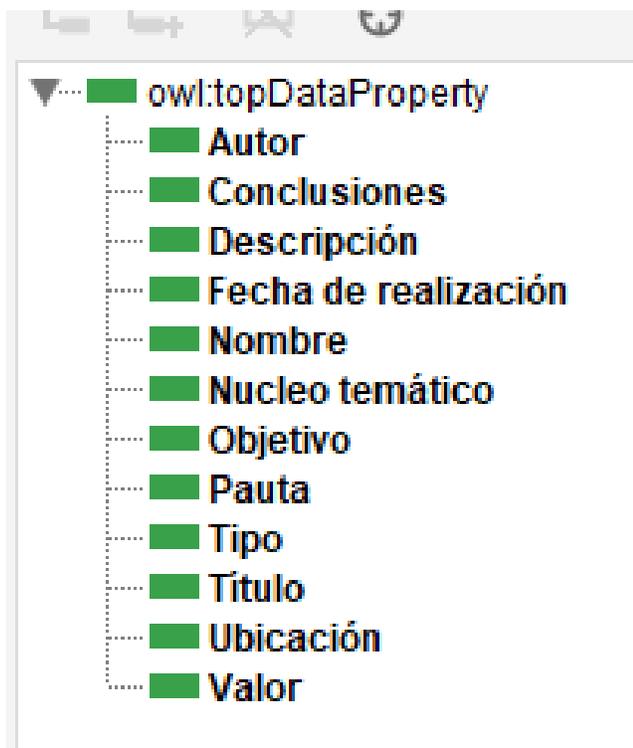


Figura 5.3: Selección de propiedades de datos

Seleccionamos la propiedad que deseamos asignar y, en el campo de texto de la derecha, ingresamos un valor adecuado. Por ejemplo, podemos asignar a la propiedad *Nombre* el valor *Proyecto de Investigación 2024*. Al hacer clic en

Aceptar, la propiedad quedará asignada a la instancia.

La asignación precisa de propiedades es fundamental para garantizar que la instancia represente correctamente los datos de la investigación. Esto facilitará futuras consultas y análisis dentro de la ontología.

5.1.4. Creación de Relaciones

Después de crear dos o más instancias, debemos establecer relaciones entre ellas para reflejar adecuadamente las interconexiones dentro de la ontología. En el mismo panel donde aparecen las propiedades, encontramos el panel *Object Property Assertions* en la parte superior. Hacemos clic en el botón +, y se abrirá una ventana similar a la mostrada en la Figura 5.4. En el campo de la izquierda, debemos ingresar el nombre de la relación que deseamos crear.

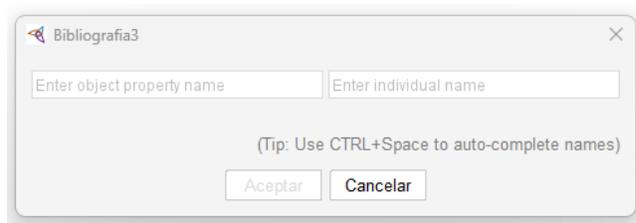


Figura 5.4: Creación de relaciones entre instancias

Todas las relaciones conectan dos clases y tienen una dirección, que va de la clase A hacia la clase B. Denominamos a la clase A como el **Rango** de la relación y a la clase B como el **Dominio**. Es importante revisar el diseño de la ontología para ingresar correctamente el nombre de la relación. La lista de relaciones se muestra en la Figura 5.5.

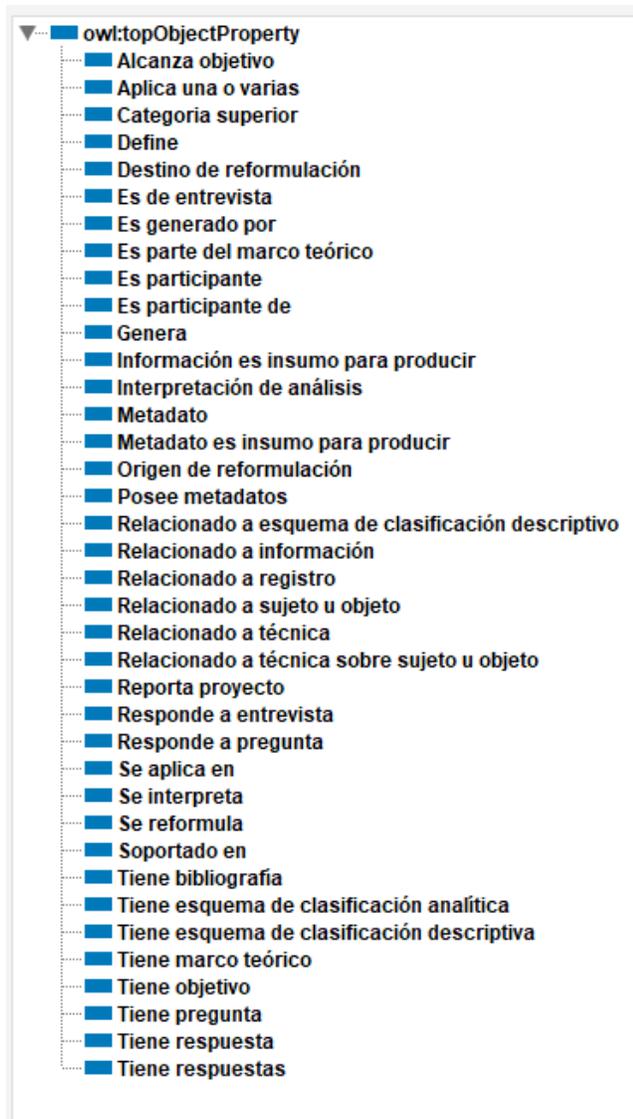


Figura 5.5: Lista de relaciones disponibles

En el campo de la derecha, ingresamos el nombre de la instancia perteneciente a la clase del Dominio de la relación que queremos establecer.

Ejemplo

Tenemos la instancia *Proyecto_1*, que representa el proyecto, y la instancia *Objetivo_1*, que representa el objetivo del proyecto. Necesitamos relacionarlas para que el modelo entienda que ese objetivo corresponde específicamente a

ese proyecto. Para ello, utilizamos la relación *Tiene_objetivo*, que tiene como Rango la clase *Proyecto* y como Dominio la clase *Objetivo*. Al seleccionar la instancia *Proyecto_1*, procedemos a crear una nueva relación. En el campo de la izquierda, ingresamos el nombre de la relación: *Tiene_objetivo*, y en el campo de la derecha, el nombre de la instancia que representa el objetivo: *Objetivo_1*.

5.1.5. Orden de Creación de Clases Relevantes

A continuación, proponemos un orden en el que las clases más relevantes deberían ser creadas, siguiendo el flujo natural de una investigación. Seguir un orden lógico en la creación de instancias y relaciones facilita el proceso de carga y asegura que la ontología represente adecuadamente la estructura y desarrollo de la investigación.

Clase *Proyecto*

Comenzamos creando una instancia de la clase *Proyecto de Investigación*. Esta clase es fundamental, ya que representa el proyecto principal del cual se derivan todas las demás instancias. A esta clase se le puede asignar la propiedad *Nombre*, donde especificaremos el título o identificación del proyecto.

Clase *Objetivo*

Cada objetivo del proyecto se representará con una instancia en la clase *Objetivo*. A esta instancia se le asignará la propiedad *Descripción*, donde se detallará el objetivo específico. Relacionaremos el proyecto con sus objetivos mediante la relación *Tiene_objetivo*, estableciendo así la conexión entre el proyecto y sus metas.

Clase *Estrategia Metodológica*

Cada proyecto se sustenta en una estrategia metodológica, la cual representaremos mediante una instancia en la clase *Estrategia Metodológica*. A esta instancia podemos asignarle la propiedad *Objetivo*, indicando el tipo de enfoque metodológico empleado, por ejemplo: *Exploratorio*. Al revisar el diseño de la ontología, observamos que el *Proyecto* está relacionado con la *Estrategia Metodológica* a través de la relación *Define*, por lo que debemos establecer esta relación.

Clase *Técnica*

Procedemos a representar las técnicas empleadas en el proyecto. Nuestra ontología permite crear instancias de tres técnicas posibles: *Análisis Documental*, *Entrevista en Profundidad* y *Observación Participante*. En el caso de la *Entrevista en Profundidad*, debemos asignar las siguientes propiedades:

- *Objetivo*: propósito de aplicar la técnica.

- *Núcleo Temático*: temas abordados durante la entrevista.
- *Pauta*: tipo de estructura de la entrevista (estructurada, semiestructurada o no estructurada).

Clase *Sujeto u Objeto*

Las técnicas se aplican sobre sujetos u objetos de estudio, y esta clase cuenta con dos subclases: *Documento* y *Sujeto*, que representan estas entidades en las investigaciones. Después de crear la instancia en la subclase apropiada, debemos asignar la propiedad *Descripción*, donde detallaremos características adicionales de esta instancia. Es necesario relacionar la instancia de la técnica con el sujeto u objeto al que se aplicó mediante la relación *Se aplica en*, estableciendo así la conexión entre la técnica y su objeto de estudio.

Clase *Registro*

La aplicación de técnicas sobre sujetos u objetos genera registros, que pueden clasificarse en una de las tres subclases: *Documento Enriquecido*, *Notas de Campo* o *Respuestas a Entrevista*. Cada registro posee la propiedad *Valor*, donde se almacena el contenido específico del registro. Por ejemplo, en el caso de una respuesta de entrevista, este campo contendrá la respuesta textual, permitiendo un seguimiento detallado de cada interacción.

Clase *Soporte*

El *Soporte* se refiere al medio en el cual se almacenan los registros generados. Puede ser de subtipo *Texto* o *Video*. Esta clase tiene las propiedades *Ubicación*, que indica dónde se encuentra el soporte, y *Nombre*, que permite identificarlo. Es necesario relacionar la instancia de *Registro* con la instancia de su *Soporte* mediante la relación *Soportado en*, estableciendo así el vínculo entre el contenido y su medio de almacenamiento.

Clase *Técnica sobre Sujeto u Objeto*

Esta clase nos permite representar una relación ternaria entre la *Técnica*, el *Sujeto u Objeto* y los *Registros* generados. Aunque no posee propiedades, es fundamental para establecer las relaciones correspondientes. Debemos relacionar esta instancia con los registros generados mediante la relación *Genera*. Asimismo, es necesario relacionar la *Técnica* con esta clase a través de la relación *Relacionado a Técnica sobre Sujeto u Objeto*, integrando así todos los elementos involucrados en el proceso de recolección de datos.

Clase *Información*

Nuestra ontología permite modelar la interpretación de los registros, generando así *Información*. Debemos crear las instancias necesarias de esta clase

para representar cada interpretación derivada de los registros. La clase *Información* posee la propiedad *Valor*, donde se documenta la información obtenida a partir del análisis de los registros.

Clase *Interpretación de un Dato*

Esta clase facilita la representación de que múltiples registros pueden ser interpretados conjuntamente para generar información. Se crea una instancia por cada conjunto de registros interpretados. Aunque no posee propiedades, es esencial para modelar las conexiones entre registros y la información resultante. Los registros interpretados se relacionan con esta clase mediante la relación *Se interpreta*. A su vez, esta clase se conecta con la *Información* generada mediante la relación *Relacionado a Información*.

Clase *Reporte*

La clase *Reporte* representa el informe final del proyecto. Posee la propiedad *Conclusiones*, donde se documentan los hallazgos y resultados de la investigación. Se relaciona con el *Proyecto* mediante la relación *Reporta Proyecto*. Además, para los objetivos que se han cumplido, se establece una relación con la clase *Objetivo* a través de *Alcanza Objetivo*, indicando así cuáles metas fueron logradas. El reporte final es la culminación de la investigación, sintetizando todos los hallazgos y conclusiones.

Con la carga de las clases mencionadas y el establecimiento de las relaciones correspondientes, obtenemos una versión inicial de la ontología que responde a las preguntas de competencia planteadas.

5.1.6. Validación de Carga

Después de haber realizado la carga de datos en nuestra ontología, debemos verificar que todo esté correctamente integrado y que no existan inconsistencias lógicas o errores semánticos. Para ello, utilizamos el razonador **HermiT** ([Oxford University Computing Laboratory, 2024](#)), que es una herramienta potente que nos va a permitir comprobar la consistencia ontológica de la ontología en Protégé.

Al iniciar el razonador, Protégé comenzará a procesar la ontología para detectar cualquier inconsistencia o error. Este proceso puede tardar desde unos pocos segundos hasta varios minutos, dependiendo del tamaño y complejidad de la ontología.

Durante el razonamiento, HermiT realizará las siguientes verificaciones:

- **Coherencia de la Ontología:** Comprueba que no existan contradicciones entre las definiciones de clases, propiedades e individuos.

- **Clasificación de Clases:** Organiza las clases en una jerarquía según sus relaciones de subclase y equivalencia.
- **Verificación de Restricciones:** Asegura que todas las restricciones y axiomas establecidos en la ontología se cumplan correctamente.

Si el razonador detecta inconsistencias, Protégé mostrará mensajes de error o advertencia en la pestaña *Inferences* o en la sección de *Explanation*. Es importante revisar estos mensajes para identificar y corregir los problemas encontrados.

Ejemplo de Verificación

Supongamos que, al cargar los datos, hemos definido una instancia que pertenece a dos clases disjuntas. Por ejemplo, una instancia que es tanto un *Sujeto* como un *Documento*, cuando estas clases se han definido como mutuamente excluyentes, es decir, un *Documento* no puede ser también un *Sujeto*. Al ejecutar HermiT, el razonador detectará esta inconsistencia y nos alertará sobre la incompatibilidad. Para corregir este error, debemos revisar las instancias y asegurarnos de que cada una pertenezca únicamente a las clases apropiadas, respetando las restricciones de la ontología.

Es recomendable llevar a cabo la **verificación regular** de la ontología, especialmente después de realizar cambios significativos. Ejecutar el razonador tras estas modificaciones nos permite identificar y corregir rápidamente cualquier inconsistencia que pueda haberse introducido. Es importante **revisar periódicamente las definiciones** de clases, propiedades y restricciones, asegurándonos de que estén bien establecidas y reflejen correctamente el dominio de conocimiento.

En resumen, la activación del razonador HermiT es una práctica esencial para validar la carga de datos en nuestra ontología, garantizando que todo esté correcto y que el modelo pueda utilizarse eficazmente en investigaciones reales.

5.1.7. Casos Reales

Para validar y ejemplificar el uso de nuestra ontología, hemos cargado algunos proyectos finalizados que emplean metodología cualitativa y son idóneos para ser representados en nuestra estructura. Estas investigaciones fueron encontradas en el portal *Colibrí* ([Universidad de la República, 2024](#)), el repositorio institucional de la Universidad de la República, que es una colección digital de acceso abierto.

Uno de estos trabajos es *La construcción de un docente inclusivo. Competencias y desafíos. Una mirada hacia la discapacidad. Aportes desde una investigación cualitativa en el departamento de Cerro Largo, Uruguay* ([Berní y Yesica, 2023](#)), una investigación que explora la construcción del perfil de un docente inclusivo, analizando las percepciones de docentes de apoyo en Cerro Largo.

Otro estudio que hemos cargado es la ontología *El encuentro inicial: Una oportunidad de facilitar la simbolización* (Conde, 2014), la cual analiza la práctica *El encuentro inicial* aplicada por maestras de Educación Inicial en jardines de infantes y escuelas públicas.

La incorporación de casos reales en la ontología no solo demuestra su aplicabilidad práctica, sino que también permite evaluar su eficacia y realizar ajustes necesarios para mejorar su capacidad de representar investigaciones cualitativas.

5.1.8. Dificultades encontradas en la carga de datos

Aunque la ontología está diseñada para abarcar una amplia gama de casos de investigación, es posible que no siempre encontremos suficiente información para crear instancias de cada elemento. Algunas relaciones son opcionales y dependen del alcance y detalle de cada estudio. En los casos de las instancias que hemos cargado, la mayoría no contaba con esquemas de clasificación ni con reformulaciones de objetivos, aspectos que nuestro modelo puede representar si están disponibles. Además, la ontología está concebida para acompañar la investigación durante sus fases intermedias. Al cargar instancias únicamente desde un informe final, como lo hemos hecho, existe el riesgo de perder información de etapas tempranas o detalles que no se incluyen en el documento final.

El uso de Protégé para la carga de datos puede incrementar la dificultad para quienes no están familiarizados con esta plataforma, haciendo que el proceso no sea tan sencillo. Sin embargo, esta complejidad puede reducirse mediante la implementación de un software específico para la carga de investigaciones, que esté conectado con nuestra ontología y facilite la interacción con el usuario.

A pesar de las dificultades mencionadas, la utilización de una ontología como herramienta para modelar investigaciones cualitativas ofrece múltiples beneficios, incluyendo una mejor organización de los datos y la posibilidad de realizar análisis más profundos. Mediante las investigaciones ingresadas pudimos comprobar que nuestra ontología responde adecuadamente al dominio para el cual fue diseñada, siendo nuestra mayor dificultad en la carga el acceso a datos completos de investigaciones cualitativas.

5.2. Herramienta de visualización

La herramienta de visualización desarrollada surge como un complemento práctico para la ontología, facilitando la comprensión y exploración de su estructura y funcionalidades. Su diseño se centra en ofrecer una experiencia intuitiva que permita interactuar con los datos y conceptos representados, destacando la aplicabilidad y el valor de la ontología en contextos reales.

A través de esta herramienta, es posible traducir los aspectos técnicos de la ontología en una representación visual accesible, haciendo que las relaciones y las inferencias resultantes sean más claras y comprensibles para los usuarios. Este enfoque busca no solo demostrar las capacidades del modelo, sino también fomentar su adopción en escenarios donde la representación semántica puede marcar la diferencia.

La herramienta está disponible en <https://eldkaii.github.io/ontology-viewer/index.html>, acompañada de un manual de usuario y una descripción detallada de la prueba de concepto desarrollada.

5.2.1. Objetivo y alcance

El principal objetivo de esta herramienta de visualización es servir como una prueba de concepto que permita demostrar la utilidad práctica de la ontología desarrollada. Lejos de aspirar a ser un sistema integral de gestión, la herramienta se centra en ilustrar cómo la ontología puede facilitar el acceso a información relevante en investigaciones, ofreciendo una interfaz más amigable y accesible para los usuarios.

Con este enfoque, se priorizó la implementación de funcionalidades clave por encima de aspectos relacionados con el diseño o la experiencia de usuario más sofisticada. El propósito fue garantizar que las capacidades centrales de la ontología, como la exploración de relaciones semánticas y el acceso de forma ordenada a los datos, sean claramente visibles y comprensibles.

En términos generales, el desarrollo de esta herramienta se enfocó en validar y ajustar sus capacidades para cumplir con su propósito principal. Entre los objetivos clave se destacan la representación clara de conceptos y relaciones, la interacción fluida con los datos semánticos y la integración de funcionalidades esenciales que faciliten la visualización de la información en diferentes contextos. Aunque no abarca todas las posibles aplicaciones de la ontología, la herramienta representa un paso inicial hacia su implementación práctica, destacando el potencial de este enfoque para resolver problemas específicos de acceso y gestión de información.

El alcance de la herramienta abarca diversas funcionalidades diseñadas para demostrar la utilidad de la ontología desarrollada y facilitar su interacción con los usuarios. Entre estas capacidades, se incluye la posibilidad de cargar y validar una ontología, asegurando su consistencia al considerar tanto su estructura como las instancias definidas. Además, la herramienta permite acceder a la información a dos niveles: uno global, para realizar consultas sobre múltiples proyectos, y otro específico, dentro de un proyecto particular, ofreciendo flexibilidad en el análisis de datos.

Para potenciar esta interacción, se ha implementado la capacidad de realizar y guardar consultas en lenguaje SPARQL, incluyendo todas las preguntas de competencia previamente definidas, las cuales han demostrado ser de relevancia práctica. Asimismo, la herramienta facilita la interacción con todas las instancias de un proyecto de investigación, proporcionando diversas vistas y la posibilidad de exportar datos en formato CSV. Esta última característica permite un análisis externo de los datos e interoperabilidad con otras herramientas, ampliando así las opciones de uso y aplicación.

Por otro lado, la herramienta incluye una funcionalidad inicial para agregar nuevas instancias a un proyecto de investigación, ya sea manualmente o mediante archivos Excel. Sin embargo, aunque esta característica forma parte del alcance, se presenta como una prueba de concepto y no como un objetivo principal de la herramienta. El enfoque central sigue siendo la visualización, buscando destacar la utilidad de la ontología para el acceso intuitivo y estructurado a la información.

5.2.2. Arquitectura y diseño de la herramienta

Aunque mencionamos anteriormente que la herramienta fue diseñada como un visualizador y una prueba de concepto, consideramos importante detallar los principales componentes que conforman la arquitectura de esta solución. Esto puede facilitar que, en el futuro, algunos de estos componentes puedan ser reutilizados con mínimas modificaciones, mientras que otros podrían ser descartados y sustituidos por alternativas más adecuadas. A continuación, presentamos un diagrama de la arquitectura y realizamos una descripción de sus principales componentes.

Como se puede observar en la figura 5.6, contamos con una arquitectura dividida en 5 grandes módulos: Interfaz de usuario, Consultas, Razonador, Backend y Loader. Estos 5 módulos nos permiten tener una división clara de responsabilidades y de realizar cambios de forma independiente en cada uno de ellos.

El módulo de Interfaz de Usuario fue desarrollado utilizando PyQt6 como herramienta principal, una biblioteca de Python que facilita la construcción de interfaces gráficas mediante diversos elementos dinámicos, como pestañas, ventanas, listas desplegables, tablas y gráficos interactivos. Estos componentes permiten ofrecer al usuario una experiencia de interacción intuitiva y funcional, posibilitando alcanzar los objetivos definidos para la herramienta. (Limited, 2024)

Por otro lado, tenemos el módulo del razonador, que es el que permite validar la consistencia de las ontologías, inferir nuevas relaciones y propiedades que no estaban explícitas en los datos originales y generar un archivo RDF enriquecido con las inferencias realizadas. Este módulo supuso un desafío técnico ya que

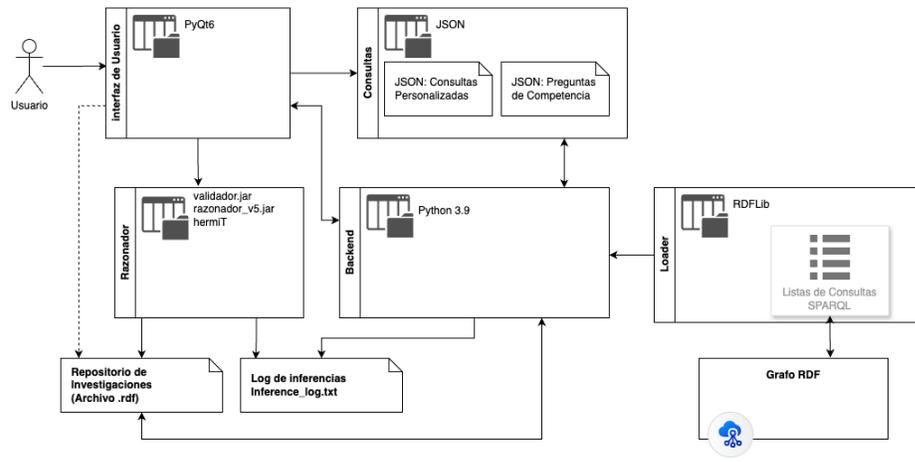


Figura 5.6: Diagrama de arquitectura de la herramienta de visualización

implicó la integración de una librería (OWL API) en lenguaje Java en nuestro sistema Python. ([The-OWL-API-Project, 2024](#))

En paralelo, contamos con un módulo de Consultas compuesto por archivos JSON que almacenan la información necesaria para ejecutar diversas consultas en lenguaje SPARQL. Optamos por el formato JSON debido a su amplia aceptación en la industria del software, lo que facilita su lectura y conversión a otros formatos de datos. Este enfoque no solo simplifica el manejo de las consultas, sino que también permite a los usuarios definir y guardar sus propias consultas, haciendo que la herramienta sea altamente flexible y fácilmente extensible.

Adicionalmente, contamos con un módulo denominado Loader, cuya implementación se basa en el uso de la biblioteca RDFlib. Esta herramienta resulta esencial para el manejo de ontologías en formato RDF, ya que permite cargar y manipular archivos en este formato, ejecutar consultas SPARQL para recuperar información específica de la ontología, y gestionar triplas mediante funciones que facilitan la adición, eliminación o modificación de datos en el grafo RDF. Su versatilidad y funcionalidad la convierten en un componente clave dentro de nuestra solución. ([RDFlib-Project, 2024](#))

Finalmente, en el centro del diagrama se encuentra un módulo denominado “Backend”, que actúa como el orquestador principal de nuestra aplicación. En este módulo se implementa toda la lógica de negocio necesaria para integrar y coordinar los distintos módulos mencionados previamente. En el diagrama se destacan explícitamente dos de estas interacciones: la primera corresponde a la interacción con el archivo .rdf cargado por el usuario, el cual es validado por el módulo del Razonador; la segunda, al registro de inferencias generado por el

Razonador, que es posteriormente procesado por el módulo Backend.

5.2.3. Resultado final y trabajo futuro

Luego de exponer el diseño y la arquitectura de la herramienta, en esta sección presentamos el resultado final de su implementación, acompañado de un recorrido por sus principales funcionalidades. Cabe destacar que, junto con la aplicación, se elaboró un manual de usuario que incluye una descripción detallada sobre el uso de la herramienta y las diversas funcionalidades que ofrece, proporcionando así una guía completa para su uso.



Figura 5.7: Pantalla inicial de la herramienta para visualizar la ontología

En primer lugar, en la figura 5.7 se muestra la pantalla inicial de la herramienta. Desde esta interfaz, el usuario puede cargar la ontología utilizando el explorador de archivos de su sistema operativo. Optamos por ofrecer la posibilidad de cargar diferentes archivos .rdf, permitiendo a los usuarios gestionar múltiples repositorios que utilicen la ontología en distintos proyectos. No obstante, es importante señalar que, para garantizar el correcto funcionamiento de la herramienta, el archivo cargado debe estar basado en la ontología que hemos diseñado. Las únicas variaciones permitidas se encuentran en las instancias incluidas, mientras que la definición de clases y relaciones debe adherirse estrictamente a la estructura que hemos establecido.

Tras seleccionar el archivo, el usuario debe esperar a que este sea validado por el Razonador y a que se genere un nuevo archivo con las instancias inferidas, lo cual habilita el uso completo de la herramienta. En este punto se da lugar a la mayor integración con el módulo del Razonador, descrito en la sección 5.2.2. Una vez finalizado este proceso, el usuario es dirigido a la siguiente pantalla,

ilustrada en la figura 5.8.

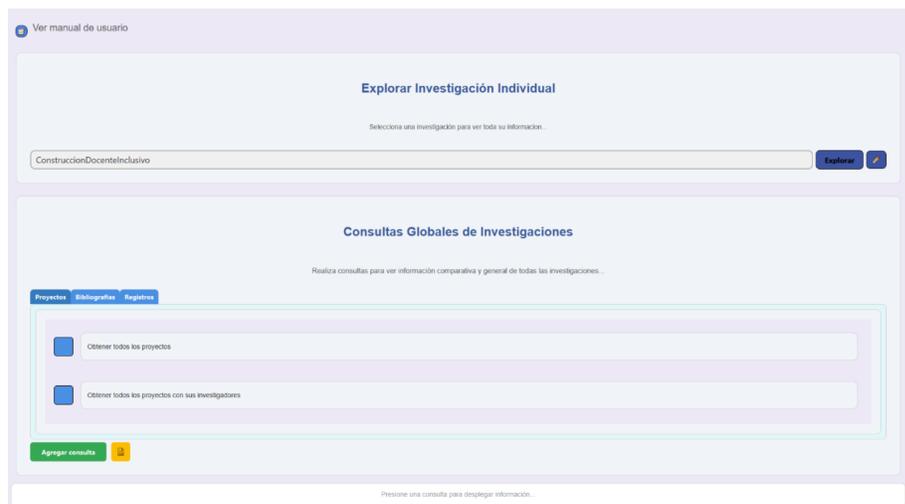


Figura 5.8: Pantalla del visualizador desplegada cuando una ontología es cargada con éxito.

Esta es la pantalla principal de la herramienta, donde se presentan dos opciones principales: explorar una investigación en particular o realizar consultas globales que abarcan todas las investigaciones. El caso de exploración individual será abordado a continuación. En cuanto a las consultas globales, el usuario tiene la posibilidad de ejecutar consultas predefinidas o de crear y guardar nuevas consultas personalizadas en lenguaje SPARQL, adaptándolas a sus necesidades específicas.

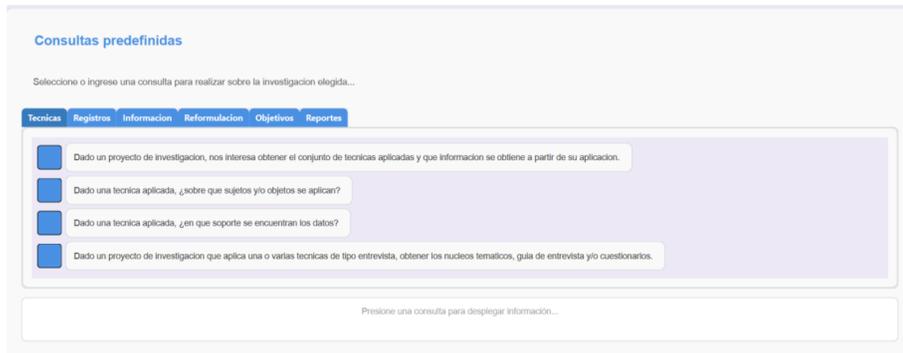


Figura 5.9: Pantalla del visualizador con la posibilidad de ejecutar preguntas de competencia para un proyecto seleccionado.

Objetos Marco Teórico Bibliografía Estrategia Metodológica Técnica Sujeto u Objeto Soporte Registro Información Metadatos Esquema de Clasificación Descriptiva Esquema de Clasificación Analítica Reporte

Bibliografía

	Nombre	Autor	Título	Relaciones
1	Bibliografía3	Adorjan, Motz, Vargas-Solar	Towards a human-in-the-loop curation: A qualitative perspective	InvestigacionEstrategiaEstudio Tiene bibliografía Bibliografía3
2	Bibliografía4	Ciesielski		Bibliografía4 Es parte del marco teórico Marco Teórico1 InvestigacionEstrategiaEstudio Tiene bibliografía Bibliografía4

Figura 5.10: Pantalla del visualizador para explorar instancias por clase de un proyecto seleccionado.

Al seleccionar una investigación individual, el usuario puede explorar manualmente todas las instancias cargadas, organizadas por clase (ver figura 5.9). Además, tiene la opción de realizar búsquedas rápidas utilizando las preguntas de competencia definidas para la ontología o formulando sus propias preguntas. En la figura 5.10 se muestra un ejemplo en formato de tabla con las instancias de la clase "Bibliografía". Todas las tablas generadas en la herramienta pueden ser descargadas, facilitando así su uso en trabajos externos.

En conclusión, vemos que las funcionalidades principales de la herramienta descritas anteriormente representan una base sólida para la prueba de concepto que nos propusimos desarrollar. A través del uso de la ontología, hemos logrado construir una solución que permite acceder a la información de distintos proyec-

tos de manera estructurada, haciendo uso de las funcionalidades que ofrece una ontología como razonamiento e inferencias y ofreciendo una interfaz de usuario adaptada al contexto de uso.

Sin embargo, como mencionamos al inicio, esta no es una herramienta integral ni una solución definitiva. Consideramos que la ontología desarrollada tiene el potencial de integrarse en un sistema más amplio, capaz de gestionar los diferentes proyectos de investigación de manera centralizada. Esto permitiría avanzar hacia una mejor gobernanza de los datos, fomentando una mayor contribución a la ciencia abierta mediante un acceso uniforme e intuitivo a la información.

En este sentido, una de las principales líneas de trabajo futuro consiste en colaborar con profesionales del diseño y la experiencia de usuario para crear una interfaz más profesional. Esto permitiría superar el carácter experimental de esta prueba de concepto, proporcionando una herramienta más accesible y comprensible para una audiencia más amplia.

Asimismo, esta evolución requeriría una escalabilidad en los demás módulos. Por ejemplo, el manejo de datos debería trasladarse a bases de datos especializadas, como triplestores ([Apache Software Foundation, 2024](#)), para optimizar la gestión de los grafos RDF. Del mismo modo, módulos clave como el Backend y el Razonador podrían migrar a servicios en la nube, garantizando un acceso independiente del entorno de cada usuario, posiblemente a través de una interfaz web. Estas mejoras serían fundamentales para consolidar el sistema integral que imaginamos y potenciar su impacto en el área.

Capítulo 6

Conclusiones y Trabajo Futuro

En este capítulo, se presentan las conclusiones finales de este trabajo, haciendo un repaso de los objetivos iniciales y una evaluación de los resultados obtenidos. A lo largo del proyecto, se plantearon diversas metas y enfoques que permitieran el desarrollo de una ontología, orientada a facilitar la comprensión, la aplicabilidad y la trazabilidad del proceso de investigación cualitativa. Aquí, revisaremos cómo se alcanzaron estos objetivos, identificando los aportes significativos, las limitaciones encontradas y las posibles extensiones para futuros trabajos.

Además de reflexionar sobre los logros obtenidos, se abordan las dificultades que surgieron durante el desarrollo del proyecto. Estas experiencias ofrecen una oportunidad para plantear distintos aspectos que podrían mejorarse, ya sea por limitaciones de tiempo, recursos o enfoques adoptados. Finalmente, se exploran algunas propuestas para futuras líneas de investigación y desarrollo que podrían enriquecer aún más el trabajo realizado, ampliando su aplicabilidad y profundizando en los aspectos que quedaron pendientes.

6.1. Evaluación de Resultados y principales aportes

Desde el inicio, el proyecto se planteó como una respuesta a las necesidades específicas de los investigadores cualitativos en la organización y estructuración de datos complejos. A través de un proceso iterativo de diseño y validación, nuestro objetivo principal fue desarrollar una ontología que no solo permitiera gestionar de forma efectiva los datos cualitativos, sino que también permitiera la obtención de información relacionada a un proyecto dado. A continuación,

evaluamos los logros alcanzados en función de los objetivos específicos establecidos y su grado de cumplimiento.

Un logro fundamental del proyecto fue la creación de una ontología que facilita la organización y gestión de datos cualitativos. La estructura desarrollada permite a los investigadores conectar conceptos y realizar consultas semánticas que eran difíciles de ejecutar con herramientas genéricas, como hojas de cálculo o procesadores de texto. Con esta ontología, los usuarios finales pueden seguir la evolución de un proyecto, rastrear relaciones entre entrevistas y categorizar respuestas de una manera mucho más eficiente, minimizando la carga de trabajo manual y reduciendo el margen de error asociado a la falta de interconectividad en las herramientas convencionales.

Además, la ontología no solo cumple con el propósito de facilitar la gestión de datos, sino que también actúa como una herramienta didáctica. Su diseño permite a los investigadores comprender mejor cómo sus datos y conceptos se alinean con el modelo propuesto, ofreciendo una estructura centralizada y coherente que apoya la gobernanza de datos en investigaciones que manejan información sensible. Esta centralización también contribuye a crear un ambiente de trabajo más amigable y seguro, donde la semántica de los datos se mantiene intacta y coherente a lo largo de todas las etapas del proyecto.

La capacidad de realizar consultas relevantes y rastrear datos en las distintas fases de la investigación fue otro de los objetivos clave que logramos satisfacer. El modelo no solo permite responder a preguntas de competencia definidas al inicio, sino que también soporta la integración de datos provenientes de diferentes etapas del proceso investigativo, manteniendo la coherencia semántica necesaria para asegurar una interpretación precisa y fiable. Esta funcionalidad no solo simplifica el acceso a información relevante, sino que también optimiza la trazabilidad de los datos en el contexto cualitativo.

Finalmente, uno de los aspectos más destacados de la ontología es su carácter reutilizable. La estructura permite a los investigadores abstraerse de volver a definir conceptos básicos como “entrevista” en cada proyecto, facilitando su aplicación en investigaciones futuras. Esto no solo agiliza el inicio de nuevos proyectos, sino que permite a los equipos de investigación concentrarse en los aspectos más críticos de su análisis, respaldados por una base sólida. La posibilidad de acompañar cada etapa del proyecto a través de una versión ajustable de la ontología también se ha cumplido, permitiendo no solo modificaciones y actualizaciones en los datos, sino también ajustes en la estructura, como axiomas de TBox y RBox, la definición de nuevas subclases, subpropiedades o extensiones más importantes, según los requerimientos específicos de cada fase.

En resumen, los resultados obtenidos reflejan un cumplimiento de los objetivos propuestos, contribuyendo a resolver las deficiencias de los modelos actuales y ofreciendo una herramienta que se ajusta a las necesidades específicas de los

investigadores cualitativos.

6.2. Dificultades y Limitaciones Encontradas

Una de las principales dificultades encontradas en este trabajo fue la complejidad inherente al dominio de la investigación cualitativa. Entrar en esta área y comprenderla en profundidad representó un desafío considerable, especialmente en la tarea de identificar y diferenciar los conceptos fundamentales que deberían formar parte de la ontología. La naturaleza abstracta y diversa de estos conceptos planteó una barrera conceptual inicial, en la que se debió discernir cuidadosamente qué elementos eran realmente esenciales para el modelo y cuáles no aportaban valor suficiente para su inclusión.

Es importante resaltar la valiosa contribución del docente Alén Pérez, nuestro usuario experto, cuya cercanía y conocimiento fueron fundamentales para que pudiéramos ingresar rápidamente al dominio y enfocarnos en los aspectos más relevantes para nuestro desarrollo. Gracias a su guía, logramos abstraernos de detalles innecesarios y concentrarnos en lo esencial. Esto resultó crucial dada nuestra aproximación bottom-up, en la que superar la barrera conceptual inicial suele representar uno de los mayores desafíos.

Además, otro desafío importante fue lograr un equilibrio entre especificidad y flexibilidad en la estructura de la ontología. Dentro de la comunidad de investigadores cualitativos, existe una preocupación legítima de que una definición excesiva de conceptos podría imponer limitaciones a la interpretación y análisis. Diseñar una estructura que fuera lo suficientemente detallada para proporcionar una organización efectiva, pero al mismo tiempo lo suficientemente abierta como para no restringir la libertad interpretativa de los investigadores, exigió un esfuerzo de conceptualización significativo y reiteradas revisiones.

Otro aspecto desafiante fue el acceso y comprensión de la información necesaria para cargar instancias en la ontología y validar su funcionalidad. No solo resultó complejo obtener los datos, sino que también fue difícil interpretarlos y categorizarlos desde nuestra posición, ya que no somos especialistas en el área de la investigación cualitativa. Al trabajar con los documentos suministrados, nos enfrentamos a la tarea de descifrar qué representaba cada elemento y cómo estructurarlo en la ontología de manera coherente. Esta experiencia nos permitió, además, comprender de primera mano las dificultades que pueden surgir cuando alguien ajeno a la investigación intenta navegar y entender los distintos componentes de un estudio cualitativo. La falta de familiaridad con la terminología y estructura de estos archivos resaltó la importancia de tener herramientas que faciliten la interpretación y organización de este tipo de datos, algo que nuestro trabajo busca precisamente abordar y mejorar.

6.3. Posibles Extensiones y Trabajo Futuro

A continuación, se presentan algunas posibles extensiones y líneas de trabajo futuro que podrían aprovechar y expandir los resultados de este proyecto. La ontología desarrollada aquí ofrece una base conceptual sólida que puede ser enriquecida y aplicada de diversas maneras para potenciar su utilidad en el ámbito de la investigación cualitativa. Tanto la propia estructura de la ontología como su potencial integración en sistemas de gestión y análisis ofrecen un amplio espectro de oportunidades para continuar explorando y mejorando las herramientas a disposición de los investigadores.

Una extensión inmediata de este trabajo se centra en la propia ontología. Aunque logramos una conceptualización del dominio, la ontología podría enriquecerse y ampliarse aún más. En primer lugar, se podrían añadir propiedades adicionales (data properties) y especificaciones más detalladas para cada uno de sus elementos, lo cual permitiría capturar con mayor precisión las características de cada componente. Sin embargo, es fundamental equilibrar esta especificación con la flexibilidad, evitando que una definición excesiva limite la aplicabilidad y adaptabilidad de la metodología. Adicionalmente, podrían incorporarse más técnicas de recopilación de datos y ampliarse algunos conceptos según las necesidades, como el modelado de los investigadores, que en etapas previas consideramos podría beneficiarse de una ontología externa específica para representar sus atributos y roles en mayor profundidad.

Además, otra extensión relevante sería integrar la ontología como el núcleo de un sistema de gestión para investigaciones cualitativas. Utilizando la ontología en conjunto con motores de gestión de ontologías integrados con razonadores, como por ejemplo Jena, que almacena ontologías en triplestores ([Apache Software Foundation, 2024](#)), sería posible desarrollar un sistema que permita a los investigadores gestionar sus proyectos, realizar consultas y organizar sus datos en un entorno intuitivo y accesible. Este sistema, con una interfaz amigable, facilitaría la visualización y consulta de relaciones semánticas en los datos, optimizando así el trabajo de los investigadores durante todo el proceso de análisis.

Por otro lado, también se podría explorar el uso de inteligencia artificial para realizar consultas en lenguaje natural sobre investigaciones almacenadas en el sistema. Al integrar un modelo de lenguaje como un LLM, los usuarios podrían hacer preguntas en lenguaje humano, las cuales serían traducidas a consultas SPARQL sobre el triple store. Esto permitiría una recuperación de información sencilla y eficiente, acercando la tecnología a investigadores que no posean conocimientos avanzados en ontologías o lenguajes de consulta.

En conjunto, estos posibles desarrollos sentarían las bases para un sistema integral de gestión y análisis de datos cualitativos, en el cual la ontología no solo organiza y centraliza el conocimiento, sino que también impulsa herramientas avanzadas para facilitar la interpretación y aplicación de dicho conocimiento.

En conclusión, el presente trabajo representa una base que, aunque inicial, ofrece un amplio potencial para futuros desarrollos en el ámbito de la investigación cualitativa. La ontología y la conceptualización logradas no solo establecen un marco comprensivo y adaptable, sino que también sientan los cimientos para la creación de sistemas más integrales de gestión y análisis de datos. Al centrarse en las necesidades reales de los investigadores, este proyecto ha conseguido articular una herramienta con capacidad de evolución y expansión, destinada a seguir creciendo y adaptándose a medida que se enriquezca con nuevos elementos y técnicas.

Las posibilidades de extensión son vastas. Desde la integración de funcionalidades avanzadas en sistemas de almacenamiento y consulta hasta la incorporación de tecnologías como la inteligencia artificial para facilitar el acceso y uso de la ontología en un entorno más accesible e intuitivo con herramientas como la transcripción de consultas en texto plano a SPARQL. Con una estructura, propósito y alcance debidamente fundamentados, el trabajo aquí desarrollado no solo responde a desafíos actuales, sino que también se convierte en un punto de partida para la creación de herramientas de apoyo más complejas que permitan a los investigadores cualitativos gestionar sus proyectos de manera más efectiva y profunda.

Referencias

- Academy, B. S. O. (2021). *Investigación cualitativa y cuantitativa: características, ventajas y limitaciones*. <https://www.santanderopenacademy.com/es/blog/cualitativa-y-cuantitativa.html>. (Accessed: 2024-10-08)
- Adorjan, A., Vargas-Solar, G., y Motz, R. (2022). Towards a human-in-the-loop curation: A qualitative perspective. En *2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)* (p. 1-8). doi: 10.1109/AICCSA56895.2022.10017577
- Apache Software Foundation. (2024). *Apache jena*. <https://jena.apache.org/>. (Accessed: 2024-11-28)
- Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., y Patel-Schneider, P. F. (2007). *The description logic handbook: Theory, implementation and applications* (2nd ed.). Cambridge: Cambridge University Press. (Available at: <https://www.cambridge.org/core/books/description-logic-handbook/>)
- Berní, A., y Yesica, V. (2023). *La construcción de un docente inclusivo. competencias y desafíos. una mirada hacia la discapacidad. aportes desde una investigación cualitativa en el departamento de Cerro Largo, Uruguay*. Descargado de <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/40610> (-)
- Conde, A. B. (2014). *El encuentro inicial: Una oportunidad de facilitar la simbolización*. Descargado de <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/4468> (-)
- Hanchard, M., y Merrington, P. (2019). Developing a computational ontology from mixed-methods research: A workflow and its challenges. En L. Pitcher y M. Pidd (Eds.), *Proceedings of the digital humanities congress 2018*. The Digital Humanities Institute. (Available online at: <https://www.dhi.ac.uk/openbook/chapter/dhc2018-hanchard>)
- Hocker, J., Bipat, T., McDonald, D. W., y Zachry, M. (2021). Evaluating qualico: an ontology to facilitate qualitative methods sharing to support open science. *Journal of Internet Services and Applications*, 12(5), 5. Descargado de <https://doi.org/10.1186/s13174-021-00135-w> doi: 10.1186/s13174-021-00135-w
- Keet, M. (2020). *An introduction to ontology engineering*. Cape Town, South Africa: Self-published. (Licensed under a Creative Commons Attribution

- 4.0 International License (CC BY 4.0). Available at: <https://people.cs.uct.ac.za/~mkeet/OEbook/>)
- Limited, R. C. (2024). *Pyqt6*. <https://pypi.org/project/PyQt6/>. (Accessed: 2024-11-28)
- Mack, N., Woodsong, C., M.MacQueen, K., Guest, G., y Namey, E. (2005). *Qualitative research methods: A data collector's field guide*. Family Health International.
- Mahlaza, Z., y Keet, C. M. (s.f.). *Ontoclean in owl with a dl reasoner – a tutorial*. https://eva.fing.edu.uy/pluginfile.php/491210/mod_resource/content/1/ontocleantutorialOE19.pdf. (Accessed: 2024-10-26)
- Ontology Engineering Group (OEG). (2024). *Ontology engineering group*. <https://oeg.fi.upm.es/index.html>. (Accessed: 2024-10-19)
- Oxford University Computing Laboratory. (2024). *Hermit reasoner*. <http://www.hermit-reasoner.com/>. (Accessed: 2024-11-29)
- Poveda-Villalón, M., Gómez-Pérez, A., y Suárez-Figueroa, M. C. (2014). OOPS! (Ontology Pitfall Scanner!): An On-line Tool for Ontology Evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10(2), 7–34.
- RAE. (2024). *Real academia española*. <https://dle.rae.es/ontolog%C3%ADa?m=form>. (Accessed: 2024-08-11)
- RDFlib-Project. (2024). *Rdflib*. <https://rdflib.readthedocs.io/en/stable/>. (Accessed: 2024-11-28)
- Sampieri, R. H., Collado, C. F., y del Pilar Baptista Lucio, M. (2014). *Metodología de la investigación* (6ta edición ed.). McGraw-Hill Education.
- Simperl, E., Mochol, M., Bürger, T., y Popov, I. O. (2010). Achieving maturity: the state of practice in ontology engineering in 2009. *International Journal of Computer Science and Applications*, 7(1), 45–65.
- Staab, S., y Studer, R. (2009). *Handbook on ontologies*. Springer.
- Stanford Center for Biomedical Informatics Research. (2024). *Protégé*. <https://protege.stanford.edu/>. (Accessed: 2024-10-19)
- Stewart, L. (2024). *Using caqdas for qualitative research*. <https://atlasti.com/research-hub/qdas-or-cqdas>. (Accessed: 2024-09-06)
- Studer, R., Benjamins, V., y Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, 25(1), 161-197.
- The-OWL-API-Project. (2024). *Owl api*. <https://owlapi.sourceforge.net/>. (Accessed: 2024-11-28)
- Universidad de la República. (2024). *Colibri - universidad de la república*. <https://www.colibri.udelar.edu.uy/jspui/>. (Accessed: 2024-11-29)
- W3C. (2024). *Owl web ontology language*. <https://www.w3.org/OWL/>. (Accessed: 2024-11-07)