![metaphacts logo]

![ontotext logo]

Major Government Agency Takes Their Digital & Print Library Services to the Next Level

Partnering with metaphacts & Ontotext

metaphacts and Ontotext have built a joint knowledge graph-powered solution that offers improved user experience for both data curators and library visitors through highly interlinked information across various libraries and archives

THE GOAL

A large Government Agency managing a country's national and public libraries wanted to improve the accessibility and usability of their digital and print resources for both data curators and library visitors.

The Agency's responsibilities cover the National Archive, the National Library and about 30 local public libraries. These house more than 1.7 million resources including digital assets (ebooks, audiobooks, collection of articles, etc.), print assets (DVDs, cassette tapes, books, etc.), authoritative data (terminology, gold standard of names, pronunciations in English, etc.) and more.

The customer required a solution that would help simplify data curation tasks and, at the same time, streamline the user experience for library visitors.

On the data curation side, the solution needed to:
» Provide data curators with one-point access and a 360° overview of all resources across all systems to identify data inconsistencies and ensure high data quality
» Map information across different systems, libraries and archives using one common semantic model for easier search, navigation and analytics
» Ensure a high data quality standard by classifying resources using consistent vocabularies based on authoritative data from the national library
» Support the easy discovery of resources and capture user flows across individual libraries and archives

From the end-user access and experience perspective, the solution needed to:
» Provide access via a single entry point to all available resources
» Deliver an easy and streamlined user experience across all libraries and archives
» Support easy discovery of resources and deliver targeted, context-relevant search results
» Offer recommendations across libraries and guide users as they navigate across systems
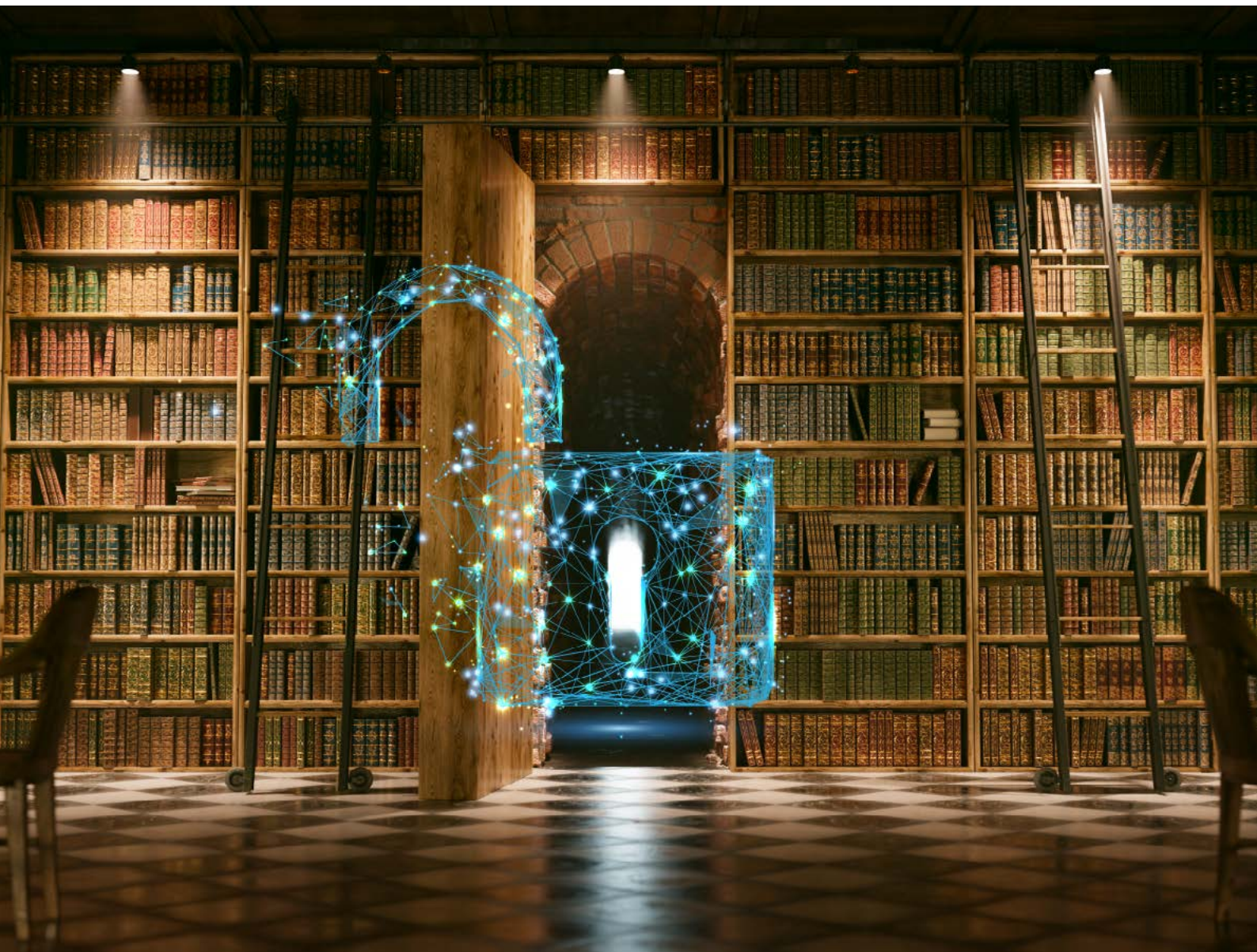
THE CHALLENGES

One of the main challenges was that the data in all digital and physical public libraries and archives was stored and managed in diverse and heterogeneous systems and formats. This meant not only across but also within individual institutions through multiple library management and content management systems that often described the same data.

Each of these systems was structured differently and curated independently by different teams. Additionally, each library had a dedicated visitors portal with a separate UI. There was no way to interlink these systems, which locked the information in individual silos and made it difficult for data curators, as well as for library visitors to find.

As a result, curation teams did not have a comprehensive overview of the data they were working with and often created duplicate or inconsistent content.

Additionally, they could not follow user flows across systems and were therefore not able to understand the types of information needs their users had.

The information silos additionally contributed to a suboptimal user experience for the visitors of the various portals, which led to poor user retention. Visitors often resorted to search engines like Google instead of the actual library systems to find additional information about specific topics, authors or titles they were interested in, as it was not possible to expose links or make recommendations across portals.

On top of that, since the end-user search interfaces were powered through traditional keyword search, they failed to provide precise, context-relevant results to user queries. Instead, end users needed to manually sift through long lists of results that were missing the necessary context for disambiguation.
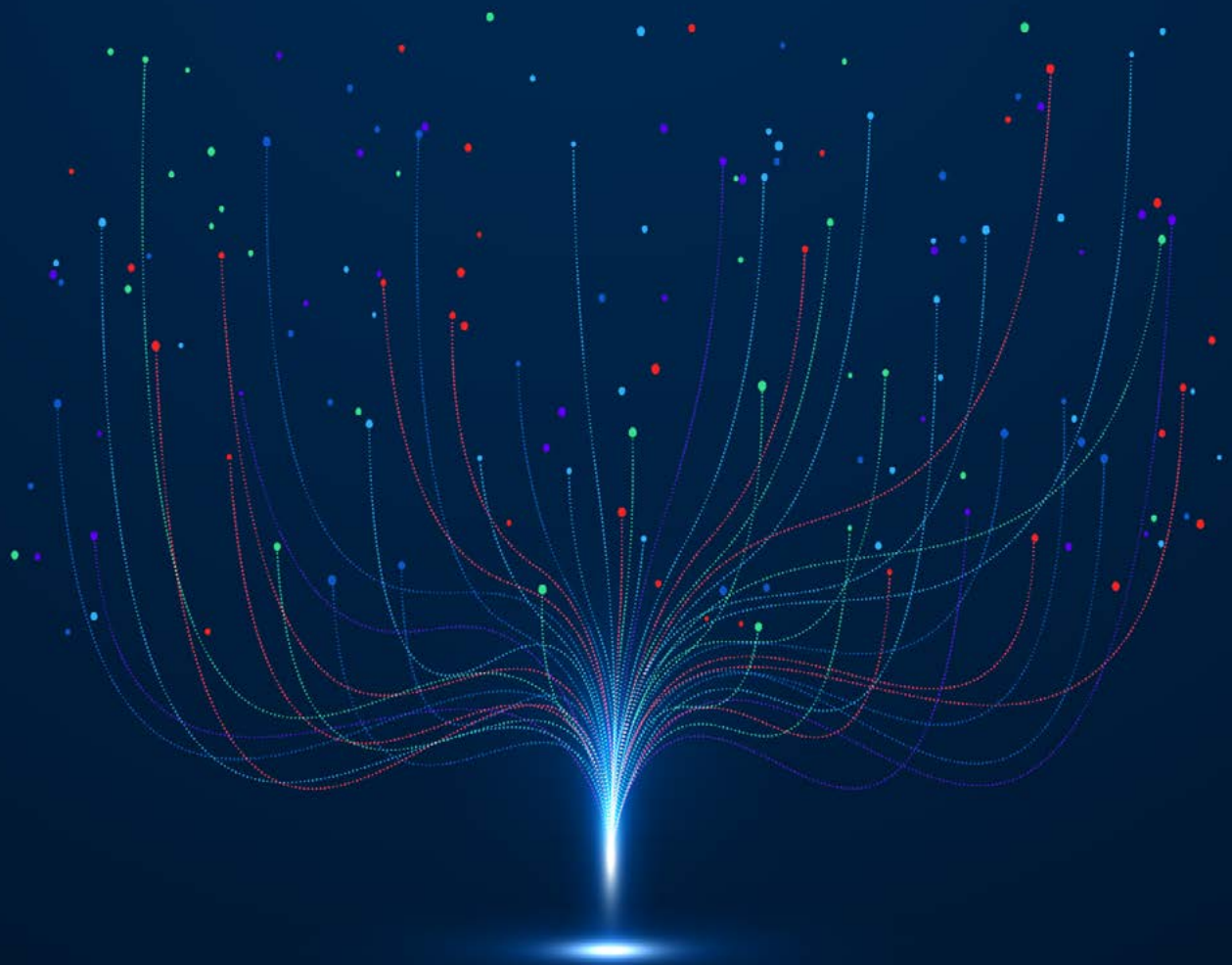
THE SOLUTION: A KNOWLEDGE GRAPH-BASED VIRTUAL LIBRARY EXPERIENCE

The solution provided by metaphacts and Ontotext brought all data together and created a cumulative knowledge graph that helped the Agency improve the accessibility and usability of their resources for data curators and, as a next step, will streamline the experience for library visitors.

Behind the Curtain

All (meta-)data from the various internal library systems, content management systems and archives was reconciled, aggregated and stored via a metaphacts developed pipeline in Ontotext's RDF database for knowledge graphs GraphDB. The data was further linked to external data sources such as Wikidata, the Library of US Congress, and VIAF via data federation with metaphacts' knowledge graph-driven platform metaphactory. This provided more context and enabled additional resolution of ambiguities.

The solution uses the GraphDB high availability cluster, which provides resilience, enterprise-grade performance and security. The GraphDB Connectors provide super fast full-text and faceted (aggregation) searches with multilingual support, whereas the GraphDB engine serves as the backbone for answering SPARQL queries.

To harmonize data and support data sharing across data sources, the solution is built on one modular semantic model that provides a general description of the entities in the knowledge graph based on schema.org. The schema.org ontology has been extended with relevant concepts from other ontologies, such as Bibframe, to include more fine-grained bibliographic references. Additionally, structured vocabularies for classifying resources based on authoritative data (e.g., along media types, historic periods, etc.) have been linked to the semantic model.

metaphactory's visual interface for semantic knowledge modeling is used for visualizing, extending and validating this semantic model and the controlled vocabularies. Additionally, metaphactory's interactive graph visualization component allows data curators to visualize and explore instance data.
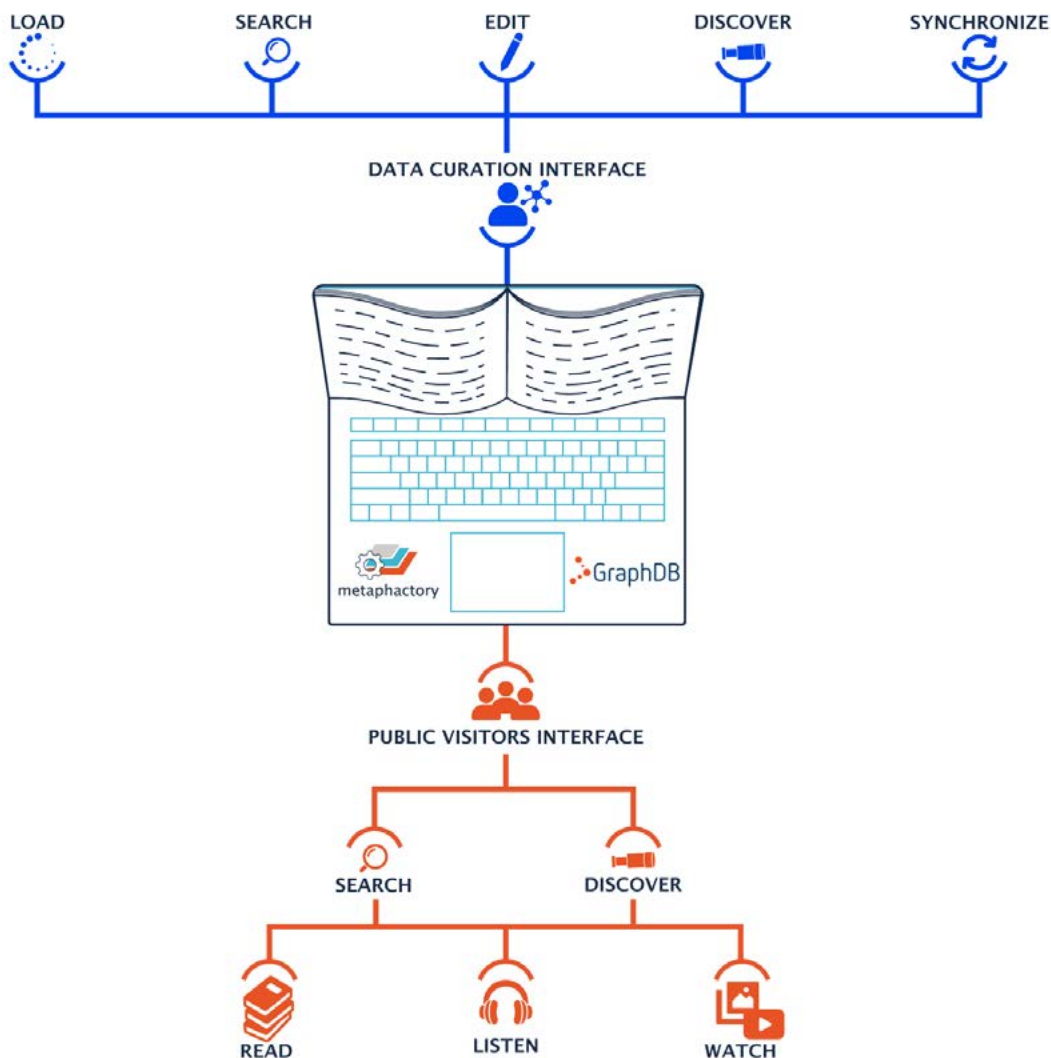
The Visitors Interface for library users that will be released in a next step will be powered by metaphactory's low-code platform for building end-user-oriented knowledge graph applications. It will offer an intuitive search interface as well as multiple paradigms for exploring and visualizing information.

The Data Curation Interface Based on metaphactory

The single Data Curation Interface (based on metaphactory) offers full control over the entire data space. It presents data curators with aggregated views of resources, derived from data that has been integrated and reconciled across all individual library systems. This allows data curators to also easily investigate data across different systems, libraries and archives, augment missing information, identify the source of inconsistencies and update any wrong values to ensure a high-quality standard for end users.

The (meta-)data that is materialized in the knowledge graph and provides a comprehensive overview of all individual library systems is updated overnight and ingested and reconciled into the knowledge graph. Now data curators have the most up-to-date state of all

data sources and can reverse entire systems or individual resources to a previous state if they need to. The use of one semantic model and unified vocabularies has helped standardize the way data is described, thus reducing human errors in data curation. Additionally, it has also streamlined the search experience for data curators (and also provides the same benefits for end-user search).

Data curators can now easily search across all data sources and understand how resources are related across systems. They can do that either by expressing a targeted query using a known concept or by starting with an area of interest and further refining the search to uncover previously unknown relationships. This greatly facilitates knowledge discovery and supports the enablement of new services.

Another important functionality is that all records can be downloaded in different RDF formats (including JSON-LD) and are automatically published to search engines as structured data. This is extremely relevant for SEO of the planned public Visitors Interface, since it supports structured data crawling by search engines and will help boost page ranking in search results.

The Public Visitors Interface

The customer is in the validation phase of a knowledge graph-based Visitors Interface that will be made available to the general public within the next months. This interface will greatly benefit library visitors by providing both in-depth exploration on topics of interest, as well as "horizontal" exploration across individual library portals.

Users will be able to access the entire collection of information from a single entry point, seamlessly navigating across libraries and archives and exploring resources presented in unified templates and result cards. The portal's UI will provide a complete picture of any information they need (including where a resource is available) and offer unified interaction patterns.

Through intuitive search and exploration interfaces based on semantic and faceted search, the Visitors Interface will provide end users with precise results for their questions. This will help reduce clutter, significantly improve the quality of search results and surface the information users are looking for in seconds.

Moreover, recommendations of related resources (regardless of where they are located) will help improve the end-user experience by interpreting user queries to provide context-sensitive suggestions. The ability to provide suggestions for mistyped keywords will further enhance the user experience and ensure that users find the information they are looking for.

BUSINESS BENEFITS

Significantly improved user experience for data curators, including through higher data quality

Time and cost savings by eliminating data redundancies and avoiding duplicate work

Faster enablement of new services for library visitors

Improved end-user experience planned, leading to better retention rate, longer engagement, higher number of visitors, etc.

SEO support by publishing data as a lightweight Linked Data format for better discovery and understandability by search engines

Standardized solution that can be deployed for other libraries

WHY CHOOSE METAPHACTS & ONTOTEXT

With the new solution developed by metaphacts and Ontotext, the Agency now has fast and easy access to and control over the information residing across multiple public libraries and archives via a knowledge graph where the (meta-)data is constantly updated.

Combining a highly scalable and robust RDF database like GraphDB with metaphactory's powerful capabilities for semantic modeling, intuitive search and data exploration will be game-changers for both data curators and library visitors. The solution has enabled the Agency to offer smooth, consistent and user-friendly interaction with huge volumes of data and access to knowledge that was previously locked in disconnected systems. All this contributes to the Agency's goal to provide a more personalized and omnichannel learning experience to all their visitors.

Your notes:

Demo
Open Knowledge Graphs
» https://wikidata.metaphacts.com

Get started
Get started with metaphactory for free
and start building your Knowledge Graph
today!
» https://metaphacts.com/get-started

Tutorial
Get a 4-week trial of metaphactory with your
tutorial when you sign up
» https://metaphacts.com/tutorial

metaphacts

ontotext