Smarter Content with a Dynamic Semantic Publishing Platform

The Semantic Technologies That Can Make Any Content Intelligent
SMARTER CONTENT WITH A DYNAMIC SEMANTIC PUBLISHING PLATFORM
The Semantic Technologies That Can Make Any Content Intelligent

A CASE FOR DYNAMIC SEMANTIC PUBLISHING
3

THE BUSINESS BENEFITS OF SMARTER CONTENT
7

THE ARCHITECTURE OF THE PLATFORM AND A LIVE SHOWCASE
10

WHAT DO YOU NEED TO KNOW WHEN CONSIDERING DYNAMIC SEMANTIC PUBLISHING SOLUTION?
20
INTRODUCTION

The question revolving around content management within and outside organizations is critical for many industries. It is a question of a major transition: from static to dynamic content creation, re-use, re-purposing and aggregation.

For years static content used to be a main way of representing knowledge. Today, we know that the need for knowledge in our modern networked societies can no longer be satiated by content, held hostage of the static format. What we are not certain about so far is what is the best approach to mapping information as to represent the actual, web-like and dynamic nature of knowledge.

Given the deluge of diverse and evolving data, dealing with heterogeneous kinds of content in an efficient way, at reduced operational costs, sounds quite challenging, if not impossible.

IMPOSSIBLE?
IT’S NOT IMPOSSIBLE, IT’S INEVITABLE.

“Everything changes and nothing remains still ...
and ... you cannot step twice into the same stream”

Heraclitus’ quote, as translated Seneca in Epistulae,
VI, 58, 23

As technologies are evolving, together with our ability to meet demanding computing power needs, we become more and more capable of modelling real-world complexity and representing relationships in web-like structures. Together with the shift in the way we perceive and consume information, come the tools we use to depict the relationships between the things in the world.
The dynamic nature of knowledge calls for dynamic solutions

Dealing with volumes of structured and unstructured data quickly and accurately is a promise that dynamic semantic publishing makes to those interested in producing, managing and curating content effectively.

A powerful blend of semantic technologies, this approach to content management has all the potential to power exceptional content experiences, created and delivered faster, in a more reliable way, at a low cost.

The automatic reuse and repurpose of content together with the need for automatic generation and linking of content, are among the main things driving the adoption of dynamic semantic publishing in the last years.

The first to recognize and utilize the advantages of a semantic approach to delivering dynamic, automated content were the BBC. In 2010 the organization decided to implement Semantic Web technologies on BBC’s FIFA World Cup 2010 portal. After seeing the benefits of dynamic semantic publishing approach, several other publishers followed suit: The New York Times, Associated Press, The Financial Times to mention but a few.

How does dynamic semantic publishing enhance content?

Dynamic semantic publishing can be used to create both journalist tools and customer facing solutions. It can power B2B applications or internal analytics engines. Relying on and producing rich metadata, a dynamic semantic publishing platform allows for:

- dynamic content aggregation
- improved navigation between linked content
- complex searches
- dynamic creation of custom content streams
- content reuse and repurposing
- user tailored recommendations

What will you find in this white paper?

In this white paper, you will read about the business benefits of having a dynamic semantic publishing platform and the problems it can solve. Additionally, you will find out what a dynamic semantic publishing platform is capable of through a live showcase demo, described in the last section of the paper. Last but not least, you will learn what to consider when choosing a dynamic semantic publishing platform.

And before we get to the business benefits of smarter content and the semantic technologies behind it, let’s answer one question:

Why is Dynamic Semantic Publishing Exciting?
Dynamic semantic publishing is exciting. It is a win-win solution for both readers and publishers. Harnessing the power of semantic technologies such as RDF, OWL, SPARQL, this approach enables a well-organized and cost-effective management of content assets.

On one hand, content creators can optimize their workflow and produce, curate and reuse richer publications. On the other hand, what's in for content consumers, are more relevant stories, interconnected content at a glance and information, as much relevant to their context, as demanded.
“BEING SMARTER
about how we do storytelling, using tools and machines to help
BETTER STRUCTURE
what we already know about information, will make us
MORE EFFICIENT and CREATIVE
with our journalism.”

Why journalism matters to the web

A dynamic semantic publishing platform does all the heavy-lifting of organizing and categorizing information thus minimizing the costs of manual content management and making the best use of the often hidden or very hard-to-retrieve knowledge within an organization. Users are being served exactly the content they are interested in, when and where they need it.

The goal of dynamic semantic publishing is aligned with a need, faced by every organization: the need for efficient knowledge management. For media and publishing companies this requirement is accompanied by the drive to stay relevant, retain audiences and bring out richer stories.

With dynamic semantic publishing the end-user’s reading experience and navigation opportunities are not only richer, but also behaviour-driven and contextualized. From the point of view of the publisher, the entire publication process is more productive, ranging from the authoring through the maintenance of the quality to the actual presentation to the reader.

The final outcome of this technology is the exciting opportunity to reuse and repurpose content in engaging new ways using the ability to personalize content streams, create content products, dynamically assemble information around user interests, and all that at a lowered operational cost and with higher efficiency.

What Does Dynamic Semantic Publishing Stand For?

The term is used to describe a set of technologies that enable the semantic enrichment (i.e. adding well-defined, machine-readable information to various content types) of content and power the automated aggregation and presentation of interlinked pieces of data.
How BBC used dynamic content to reduce content management costs and increase content reuse and visibility

In 2010, capturing the richness of the information around 2010 FIFA World Cup was challenging for BBC. There were too many pages and too few journalists to create and manage the content on the website. The editorial team needed a way to quickly and accurately generate large volumes of timely content about the matches, groups, teams and players without relying on the costly manual intervention.

An increase of content re-use, re-purposing and aggregation, at no additional cost

What helped BBC achieve the desired efficiencies in content creation and curation was dynamic semantic publishing. BBC started their journey by choosing flexible knowledge schema approaches such as ontologies and an enterprise-ready NoSQL graph database with the power of inference - Ontotext GraphDB.

After applying the process of producing not just content, but semantically enriched smart content, the organization managed to minimize the expensive editorial management of content assets (e.g. text, video, pictures and data) and to create a navigation of the website led by concepts that are important to the reader (e.g. teams, countries, players, etc.).

The dynamic semantic publishing architecture developed for BBC’s 2010 World Cup web site generated and published HTML based on dynamically selected content structured by embedded Linked Data identifiers, ontologies and associated inference.

Technically, W3C standardized RDF semantics was what triggered improvement in navigation, content re-use, re-purposing and search engine rankings. The dynamic semantic publishing approach also facilitated multi-dimensional entry points and a richer navigation, greatly improved user experience and levels of engagement.

What does W3C Standardized Mean?

The World Wide Web Consortium (W3C) is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards.

W3C standards define an Open Web Platform for application development that has the unprecedented potential to enable developers to build rich interactive experiences, powered by vast data stores, that are available on any device.

W3C publishes documents that define Web technologies. These documents follow a process designed to promote consensus, fairness, public accountability, and quality.
Making Sense of Text and Data 1 866-972-6686 © Copyright 2015 Ontotext. All Rights Reserved www.ontotext.com

Transforming the BBC’s static content management and publishing framework to a fully dynamic semantic architecture made it possible for the British Broadcasting Company to:

- deliver products and content in a smarter way, at lightning speed and finer grain of detail
- improve the breadth and depth of the navigation across the content
- create content and reuse it across divisions
- enhance the visibility of BBC content through better SEO
- reduce production costs

User interests

Authoring

Editing

Dynamic Topic Pages

Personalized News

A CASE FOR DYNAMIC SEMANTIC PUBLISHING
The Semantic Technologies That Can Make Any Content Intelligent

News

Personalized News

Editing Authoring

Dynamic Topic Pages

User interests

Greece

Greece

Greece

Financial News

Political News

News

WWW

Greece

Greece

Greece

Political

News

Financial

News

WWW
THE BUSINESS BENEFITS OF SMARTER CONTENT

One of the most enticing advantages that semantically enriched content provides for organizations is interconnectedness. By semantic enrichment data becomes reusable, easy-to-search and ready to be interlinked in innumerable combinations. This opens golden opportunities for efficient content repurposing and relevant delivery.

Automated surfacing and repurposing of the content allows for dynamic streams and topical pages, easy browsing through the content, automatically created links, related to each topic (or entity) and to the relevant places and things.
With the help of semantic technologies, content management enters a new era where content enrichment, personalization and semantic search efficiently serve business goals.

Semantically enriched content brings the benefits of:
- Faster and more accurate research
- Relevant, up-to-date, contextualized information
- Easier searching for, presenting and grouping information
- Reduced operational costs for content management
- Interrelated ways of navigating through content
- Complex searches that go beyond keywords
- Depth and breadth of information

The network view of numerous connections, offered by a dynamic semantic publishing platform, solves many of the content challenges businesses face today. Without structure, proper tagging and classification methods, any content is just an unusable pile of hard to find information. What dynamic semantic publishing does is turn that pile into an opportunity.

When to consider a dynamic semantic publishing platform?

The difficulties which organizations are trying to overcome when deciding to enhance their content management and publishing frameworks are usually related to inefficient production workflows, static, hard-to-reuse content, content silos etc. All these issues and many more can be resolved with the implementation of a dynamic semantic publishing platform.

Inflexible legacy CMSs
Disparate Metadata Schemas and Systems
Static Digital Content
High Content Production Costs
Content Silos
Content Integration Issues
Lack of User Analytics
Lack of Content Analytics
Redundancies in Production Workflows
Under-utilized Archival Material

A dynamic semantic publishing platform can:
- Automatically create relationships between Things (people, places, events etc)
- Automatically add tags, concepts and topics to content
- Dynamically aggregate topic and profile pages
- Create custom content streams
- Reuse and repurpose content, based on variables
- Provide tailored recommendations
- Suggest related content and data as the user types
You know that a dynamic semantic architecture will benefit your organization when:

- You are unable to search and link your content efficiently
- You need significant time to create content/products
- You don’t have a finer grain of understanding of your content and its usage
- Your model of knowledge doesn’t change as your organization changes
- You can’t empower efficiently your authors and editors to create high quality content despite of the fact that you have tons of marvelous legacy data
- You can’t enrich efficiently your taxonomies, thesauri and vocabularies.
- You start losing end users because of the lack of dynamic content, good search and personalized recommendations

THE SEMANTIC APPROACH ENABLES YOUR ORGANIZATION’S IT INFRASTRUCTURE TO “KNOW” AND “REMEMBER” what things in the world your content is about and allows users to explore the content in a more intuitive way at a finer granularity of detail.
THE ARCHITECTURE OF A DYNAMIC SEMANTIC PUBLISHING PLATFORM AND A LIVE SHOWCASE

The dynamic semantic publishing platform architecture is oriented towards availability and consistency (as in the CAP theorem). It consists of databases: proprietary graph database - Ontotext GraphDB, and open source document store - Cassandra, two core computing services Concept Extraction Service (CES) and Recommendation engine and a set of (micro) services - APIs on top of the databases and computing components. All services communicate via HTTP and provide a RESTful interface.

The database layer is horizontally scalable, which is ensured by GraphDB replication cluster and Cassandra's ring topology. CES employs memory sharing for vertical scalability and also a high-availability cluster, which can be used for horizontal scaling. The Recommender is basically a REST API relying on Apache Solr (for full text search) and Cassandra (for document retrieval), which can both be scaled according to the system load.
A live showcase of a Dynamic Semantic Publishing Platform

News on the Web ([NOW.ontotext.com](http://NOW.ontotext.com)) is a free public service, showcasing the opportunities a dynamic semantic publishing platform opens up before media & publishing companies. It is a live showcase of some of the basic capabilities of dynamic semantic publishing.

**NOW** has an RSS crawler which feeds documents from multiple sources into a custom processing component, which annotates the content and then stores it in GraphDB via the Concept API. This is an easy to develop approach when the developers and the core component creators work under the same roof. Since all the components provide RESTful APIs, it is relatively easy to embed the dynamic semantic publishing platform into an architecture based on distributed messaging systems, such as Apache Kafka.

In NOW have been combined several Open Linked datasets, including [DBPedia 2015](http://dbpedia.org), [WikiData](http://wikidata.org), and [GeoNames](http://geonames.org), to create a high-coverage general purpose dataset, containing over 4 million People, Locations, Organizations, Animals, Plants and other Things.

The dataset is used to perform automated concept and relation extraction, which basically generates the semantic fingerprint of an article. Further this fingerprint serves for building links between content, suggest similar content, provide facet and hybrid (concept + FTS) search.

On top of this, the platform can also recognize things in the articles, which are not yet present in the dataset. This allows the dataset to be extended with automatically extracted data, for example, when new companies are found, or new people have become popular in the news.

### Here’s a list of the features and the semantic technologies that are showcased in NOW

- Concept Extraction
- A Novel Concepts Tagger
- Content Aggregation
- Recommendations and Personalization
- Faceted, Hybrid and Complex Searches
THE ARCHITECTURE OF THE PLATFORM AND A LIVE SHOWCASE
The Semantic Technologies That Can Make Any Content Intelligent

SOLUTION APPLICATIONS

SEMANTIC SERVICES

CONTENT & DATA

www.ontotext.com ● Making Sense of Text and Data ● © Copyright 2015 Ontotext. All Rights Reserved ● 1 866-972-6686
“Concept extraction” or “semantic annotation” is the process of identifying in text mentions of concrete concepts, named entities, numbers, datasets and similar objects. For the showcase, have been extracted common concepts such as People, Locations and Organizations and the various relationships between them, for example the role of a person within an organization, competition between organizations, partnership and customer relations, mergers and acquisitions, etc. All recognized mentions of concepts and relations in texts are called annotations.

80% OF THE TOKENS in every article are **TAGGED**, which **ALLOWS DISAMBIGUATION** between candidate annotations to be much better, but, for simplicity’s sake, only a few annotation types are on display.
The showcase platform uses a mixture of in-house developed dynamic Gazetteers, Machine Learning (ML) classification and disambiguation models, as well as standard Natural Language Processing (NLP) components. There is also a novel concepts tagger, capable of recognizing concepts that do not exist in the database yet.

And since the world changes, there's a built infrastructure that allows to continuously extend the ML - machine learning training and test data and to re-train the models, in order to make this component of the system learn and become better over time.

We now know that Chung works as an analyst in Nomura. And what he said about the smartphone market.

"And since the world changes, there's a **BUILT INFRASTRUCTURE** that **ALLOWS** to continuously extend the **MACHINE LEARNING TRAINING**"
Content Aggregation
Or topic pages or profiles

One of the challenges in the publishing industry is to maintain large numbers of aggregated pages, as people can’t do it manually on such a scale. For example, the BBC did succeed to maintain rich and dynamic, always up to date, web pages for each of the 10 000 athletes that participated 2012 London Olympic games. These web pages combined related news and various statistics about each athlete. Without this core technology, they wouldn’t have been able to do it (BBC’s blog on this topic).

At the moment, NOW has around 150 000 topic pages People, Locations and Organizations. These pages present concept descriptions (from the reference knowledge graphs), how concepts trend in the news, recent articles about the concepts, what concepts they usually co-occur with, as well as other extracted information such as what people were quoted saying in the news. As a result, new ways of navigating through content are provided, as there are many more links to follow than in the traditional news websites.

For example, the topic page of Vladimir Putin shows a lot about who he is, what his religion is, when he was born, who his parents are, and so on.
The recent news about him and at which points in time he peaked in the news is also available to view. The mentioned peaks usually coincide with major events and can enable data journalists to summarize the most important events of the year in which Mr. Putin participated.

The associated entities view informs about the usual context in which Vladimir Putin is mentioned. In NOW there are 323 news that mention both Putin and the crisis in Ukraine. In August 2015, Angela Merkel is represented by a smaller circle, as she was mentioned less frequently together with Putin, because she could not talk to him on the phone every other day. All of these bubbles are clickable, so if one is interested in the topic page of the United Nations, they could easily navigate through.
Recommendations and Personalization
How to serve relevant content and ads

The dynamic semantic publishing platform includes a recommendations engine, which can be used to serve relevant content and ads, either based solely on context, or on user behaviour, or both.

The behavioural aspect of the engine feeds from the semantic fingerprints that people leave when they browse content on the web. The system also incorporates a time component, responsible for the so called recency boost. It can be configured according to the use case - a news website usually tends to require fresh content in the top results, whereas a librarian tool opts for the most relevant matches.

For contextual recommendation consider the following example - an article about the slight increase of US gas prices is matched with relevant articles using the most relevant terms via classical bag of words techniques such as term weighting (tf-idf) and hashing tricks, calculation for similarity between the concepts and key-phrases extracted from these articles, and the recency component which is configured to prefer fresh news.

What’s more, the recommendations engine opens the door for creating custom content streams for each user, or dashboards, showing what the different trending topics are along with what’s happening with the things they’re interested in (i.e. F1 season is resuming soon, news about the teams and the drivers are starting to reappear), where they can start exploring the (latest) content.

Another interesting option is to provide user interface mechanisms that show what people in the same geo-spatial region or with similar interests are reading. All these collaborative filters, basically cluster users around different properties and this comes in very handy when serving targeted ads.

In NOW is shown the contextual recommendation flavour of the engine, as there is no tracking of user behaviour (You can try the behavioural recommendations on Financial Times’ website).
Faceted search suggests concepts to search for and allows the user to drill-down - when concepts are selected, the suggestions are updated instantly to include only concepts co-occurring in text with the already selected ones. Faceted search is one of the cool features supported efficiently by all Lucene-based search engines - Lucene, Solr and ElasticSearch.

Using a graphdatabase GraphDB Connector, search indices are seamlessly updated, so whenever metadata and content are added to GraphDB, such full-text search indices are updated automatically within the same transaction.

Since the concepts in a document are indexed, later when searching, suggestions of other concepts can be limited to only such that are mentioned together with the ones already selected. In this way relevant suggestions in the facets are being ensured and there is no chance to hit a search that will result in an empty list of documents.

As an example, you might start wondering what Albert Einstein, New Jersey and Neil Armstrong have in common, especially when you see Nike popping up in the suggestions. Apparently, Nike was sponsoring an event where the two greats were honored.
Next, you can mix the faceted search query with a full-text search (FTS) query to find documents containing particular concepts and phrases.

Finally, you can do much more complex queries when there is rich metadata associated with the free text. For example, we can query for a person's opinion about a particular topic within a given time range. This is not only useful when you're having an argument at dinner, but also to conduct journalistic fact drills. It also allows you to see how a person's opinion or situation changes over time and so do the statements of that person.

For instance, you can trace Obama's stance against the Iran deal or you can ask what a macroeconomist forecasts about the yields of German bonds for the two quarters after the European Central Bank started its quantitative easing programme. There are no examples in NOW, however, as it is far from trivial to expose this functionality in a generic way. The potential for such searches is there though, they just require the implementation of custom (micro) services to do the job.

“You can do much more complex queries when there is rich metadata associated with the free text.”
WHAT DO YOU NEED TO KNOW WHEN CONSIDERING DYNAMIC SEMANTIC PUBLISHING SOLUTION?

Dynamic semantic publishing solutions are based on rich metadata descriptions of individual assets, and allow you to automatically publish thematic and personalised content streams.

Defining the bespoke parts of your solution should always start with your business objectives. Imagining the functionality of your final product is then easily reverse engineered into:

- the recommendation, search, navigation, aggregation functionality needed
- the metadata you need as asset descriptions and how to obtain it
- the base domain models and datasets as your starting knowledge graph

There are a number of important decisions to be made when making an investment to move from a more traditional digital publishing workflow towards a dynamic semantic alternative.
Having a domain model addressing the functional aspects of your business objectives and having a quality dataset is key for achieving optimal results. You can build a data set from free data sources such as Linked Open Data, from your own data, or buy a high quality curated one.

As a quick-start package for the Ontotext publishing platform we chose to carefully select parts of public datasets like DBpedia and WikiData. The resulting enrichment coverage and quality can be observed in our public showcase - News On the Web. For general or open domain enrichment and recommendations this approach gives great results for popular entities.

Many of our clients are focussed in a specific domain (e.g. financial news, scientific content, or entertainment) and in these cases they very often choose to build a domain model and a data set matching perfectly their view of the world. Sourcing your data from your own databases or buying a data feed from a third party gives you better control but it is usually more expensive. This was the route taken by established media organisations such as the BBC, Financial Times and Euromoney. Choosing to buy a high quality dataset saves you time for updating, merging and complementing the data but you have to consider that these data sets are usually not in RDF format and might be as well bigger (i.e. more noisy) than you need.

Third, the best concept extraction solutions to date seem to be done by mixing the two main approaches in the field - machine learning and rules. For the statistical models created with machine learning there needs to be a good corpus to train and test the machine models with. Again, doing this of high or low quality is essential for the outcome. Having a poorly selected corpus will result in very promising quality of extraction metrics during the development phase and a bafflingly crappy performance in the wild. On the contrary, a well selected corpus will make for reliable metrics during development which are very close to reality.

**Common mistakes:**

- underestimating the cost of building a domain model and assembling a representative data set
- assembling the data sets without constant awareness that they will be used for semantic enrichment and the specific requirements associated with it
A major decision in the adoption of a semantic publishing approach is deciding on how you will obtain the capability to describe your content assets through semantic metadata. This metadata or so-called concepts have additional links to other concepts and thus represent a rich graph structure. For instance, Barack Obama is linked to the Democratic Party (his political orientation), other Democrats such as Joe Biden, a place of birth, a spouse, children, education, etc. Enriching your content in this way opens up the opportunity to dynamically create new content-based product offerings, while your consumers can benefit from adaptive content streams, personalised through their choices and behaviour. However, you have to consider that a well-performant semantic enrichment always relies on a quality dataset, gold standards for training and evaluation, and a continuous feedback mechanism.

Some organisations produce very low volume high-quality content and in such cases it is natural to go with manual annotation by a subject matter expert. However, in most cases the content feeds that needs to be enriched are of much higher volume and an automatic approach is way more efficient. In such cases, you need to invest time from subject matter experts in building the initial training sets from which the automatic concept extraction will learn. If you want to ensure quality of the enrichment in the long run you can design a mechanism where a curator is approving or rejecting the suggestions from the automatic process for a portion of the annotated content.

Common mistakes:
- underestimating the effort to produce and maintain a high quality machine readable gold standard for training and evaluation
The best semantic enrichment solutions to date are done using a hybrid approach of machine learning and rule-based techniques. The rule-based approach is good when you have a reasonably good dictionary of entities and very clear repetitive patterns surrounding the mentions in your content. This approach is usually cheaper to start with, but harder to manage in the long run. If your content is too complicated for a human to derive rules then machine learning should be your choice as it takes into account much more features from both the content and the graph of your data. Still, you have to know that statistical machine learning algorithms give excellent results when there are enough manually curated examples to train them and good feature engineering.

Common mistakes:
- start with rule-based approach and get lost in the interdependency of the multitude of rules
- fail to build a hybrid workflow combining the benefits of both approaches
Search, Recommendations and Personalisation
What matters for your users and what matters for you?

One of the most prominent applications of dynamic semantic publishing is to use it for recommendations and personalisation. The metadata fingerprints of the content and capturing user activity allows us to build a metadata-driven user interests and behavioural profiles. Many publishers know some demographic characteristics of their readers, or explicitly ask them to express interests in various topics out of a predefined set of areas. All this data plays an important role, but remain quite static and way too general. By capturing user activities and aggregating a dynamic metadata driven profile we are able to personalise the user experience and offer:

- personalised landing page - just the content relevant to your interests (this is the most unexploited benefit so far, as many publications still tend to prefer a manually curated of what is considered to be the main topics for the day)
- contextually and behaviourally augmented recommendations of relevant content while you are reading or viewing an asset
- identification and potential use of interest groups based on post processing of the accumulated profiles

You first need to consider what type of recommendations and personalisation your solution needs, based on this it will be obvious which are the user activities you would like to capture: reading, searching, choosing a search result or a recommended article, liking, posting, commenting. The recommendation engine should be built in a flexible way, so that you can add new actions over time and most notably to be able to change the different weights used in the recommendation algorithms. A typical solution including our personalisation engine has runtime control of over a dozen of coefficients influencing the results.

Common mistakes:
- over simplifying the user interests and locking the user in a limited subset of your content; instead one should surprise and open those limits by injecting surprise recommendations
- failing to integrate the static (demographic) characteristics and the dynamic behaviour driven profile
CONCLUSION

Well-structured, semantically rich content is a strategic asset. Having a system that lets you dynamically and automatically interconnect content objects such as tags, concepts, terms etc can easily turn unstructured information into knowledge and insights.

A dynamic semantic publishing platform creates reusable and reconfigurable pieces of content that in turn enable innumerable combinations of resources. The live showcase NOW is only an instance of this technology, created with specific datasets and specific configurations. It is based on a readily prepared solution for news and media and it is pre-loaded with defaults – with a default ontology, with hand-picked datasets, with already trained extraction pipeline.

This type of filtering and decisions of what to be processed and what not to be processed pretty much lies with the client. Depending on the client's needs a dynamic semantic publishing platform can be fed with proprietary datasets and assembled in a unique way that best serves the needs of the user.

To learn more, visit www.ontotext.com/customers/ and check out the dynamic semantic publishing solutions Ontotext provided throughout the years across media & publishing, life sciences, government, financial services, telecommunications, cultural heritage and other industries.

Get in touch with us for a tailored demo on NOW.ontotext.com, and a friendly chat, during which we’ll discuss your unique challenges and identify the biggest opportunities for you to turn your assets into smart content using a dynamic semantic publishing approach.

EXCITED BY WHAT DYNAMIC SEMANTIC PUBLISHING HAS IN STORE FOR YOUR ORGANIZATION?
You can also reach us via email at info@ontotext.com and directly by calling 1-866-972-6686 (North America), or +359 2 974 61 60 (Europe).

WE LOOK FORWARD TO HELPING YOU MAKE SENSE OF YOUR DATA AND CONVERT YOUR CONTENT INTO REVENUE