

Processing Big Data with IBM and xPatterns

IBM LanguageWare and InfoSphere Streams Integration with xPatterns™



For additional information, call the Atigeo team at 425.635.3900 or visit our website at www.atigeo.com.

Much of the world's Big Data is similar to an iceberg, where most of it is frozen underwater. It's "frozen" because format and meta-data standards make it hard to flow from one place to another. Data is "underwater" when, reasons of competitiveness, privacy, or lack of resources prevent it from being shared. Furthermore, most enterprises are unable to take advantage of this semantically rich unstructured data they are capturing about their customers and users in their data centers.

Addressing the Growth of Unstructured Data with IBM and xPatterns

Integrating core IBM technologies with xPatterns™ technology is enabling a new pathway through the data processing stack; This enables the unprecedented conversion of extremely high volumes of unstructured data into actionable semantic understanding and knowledge. In many cases enterprises simply discard free-form unstructured text data, due to its large volume and the lack of tools to dynamically extract value from the data. This new pathway enables actions that organizations can take, creating insight about their customers or businesses. Imagine taking every customer email, chat session, and transcribed customer service calls, relating them all through semantic understanding, and enabling cross-channel analysis of all communications relating to a particular product, service or other customer issue; understanding the impact, sizing the cohort of affected users, and rapidly identifying the most successful messages, offers and strategies being used to result in a positive response to the customer issue.

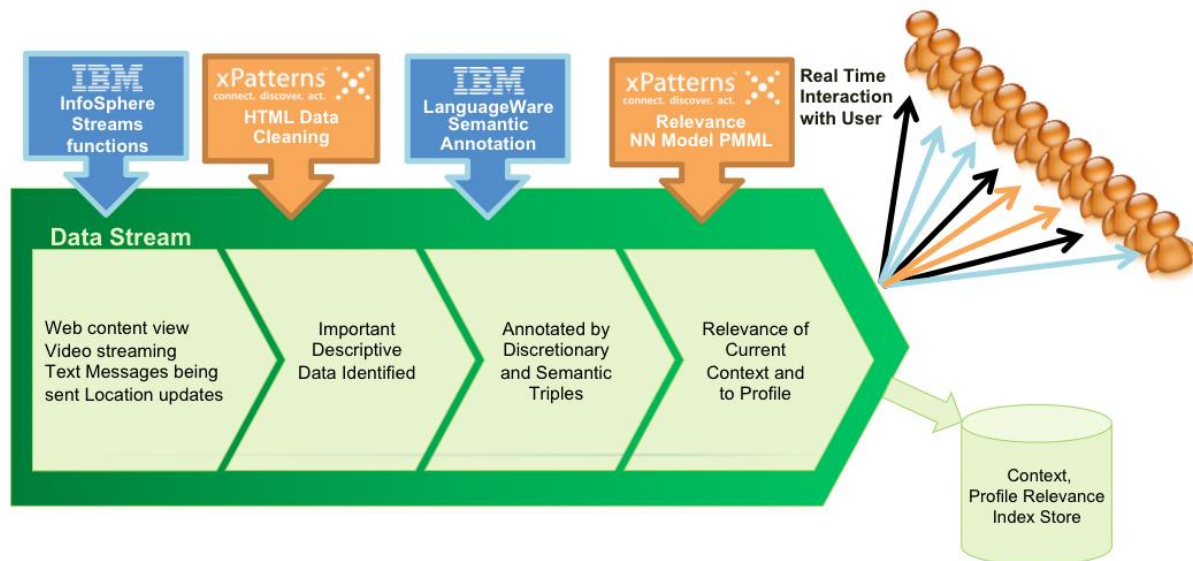


Diagram1: Extracting Semantic Meaning from Large Volumes of Streaming Data

As depicted in Diagram 1 above, IBM's InfoSphere Streams technology can be used to process extremely high, real-time, asynchronous data volumes. Integrating this data flow with xPatterns technology increases the relevance of current context to profile attributes in real-time. For more conventional data volumes the same data flow is true, but can be executed outside the Streams processing platform, using industry standard web services interfaces to allow the technologies to interact with inbound data sources, consumer systems and one another.



For additional information, call the Atigeo team at 425.635.3900 or visit our website at www.atigeo.com.

© 2009 Atigeo. All rights reserved. The information contained in this document is confidential and/or proprietary to IBM, Atigeo and/or its affiliates and/or its agents. The information contained herewith is intended only for use by the individual or entity to which it is addressed. – v12

InfoSphere Streams facilitates a web content view of streaming video, text messages, location updates, etc. while the xPatterns HTML data cleaning function determines the “meaning” of your content by indexing and semantic understanding - this identifies the important and descriptive data.

LanguageWare semantic annotation creates data for a domain expert corpus that is annotated by discretionary and semantic triples that xPatterns can determine the relevance of current content in an individual’s profile by using xPatterns neural network model. This also provides content into the profile relevance index store, which is used to learn and create accurate persona awareness.

The Technology

IBM InfoSphere Streams enables continuous and extremely fast analysis of information-in-motion to help improve business, insights, and decision making against extreme volumes of real-time data. This provides an execution platform and services for xPatterns, that ingest, filter, analyze, and correlate potentially massive volumes of continuous data streams.

IBM’s LanguageWare provides a full range of text analysis functions which focus’ on mining facts from large repositories of text. By applying LanguageWare’s dictionary lookup technology one can classify supporting lookup functions; such as synonym expansion, hyphenation, approximate/fuzzy string matching, etc. - feeding the xPatterns corpus.

xPatterns delivers actionable intelligence from an enterprise’s disparate sources of structured and unstructured data. Building on a loosely-coupled architectural approach that facilitates straightforward communication and integration, xPatterns is able to represent data in unique ways, with emphasis on semantic understanding. In particular xPatterns provides a framework for modeling relevance through the mining of unstructured data, and uses Neural Network technology to dynamically learn and refine its model of relevance. The xPatterns Neural Network can be represented in the form of a PMML plug-in for InfoSphere Streams.

Furthermore, xPatterns facilitates better decision-making, smarter choices and more relevant connections by leveraging xPatterns rich user persona which has optional privacy-controls. The more the xPatterns is used, the smarter and more useful it becomes due to xPatterns unique technology that leverages machine learning, a type of artificial intelligence that learns through experience, and neural networks to help companies and people optimize the world in which they live, work and play.

LanguageWare/xPatterns Integration

LanguageWare provides documents annotated by dictionary term occurrence, semantic triples and higher-order semantic extraction rules for xPatterns to build a corpus. With xPatterns, we take a discovery approach to extracting a model of relevant themes and concepts among the semantic annotations by mining the corpus of text content provided by the LanguageWare data. xPatterns can then be used with a user’s xPatterns persona to determine relevant information for an individual, or can



For additional information, call the Atigeo team at 425.635.3900 or visit our website at www.atigeo.com.

© 2009 Atigeo. All rights reserved. The information contained in this document is confidential and/or proprietary to IBM, Atigeo and/or its affiliates and/or its agents. The information contained herewith is intended only for use by the individual or entity to which it is addressed. – v12

be queried independently to return either related semantic concepts or semantically-related content to a query.

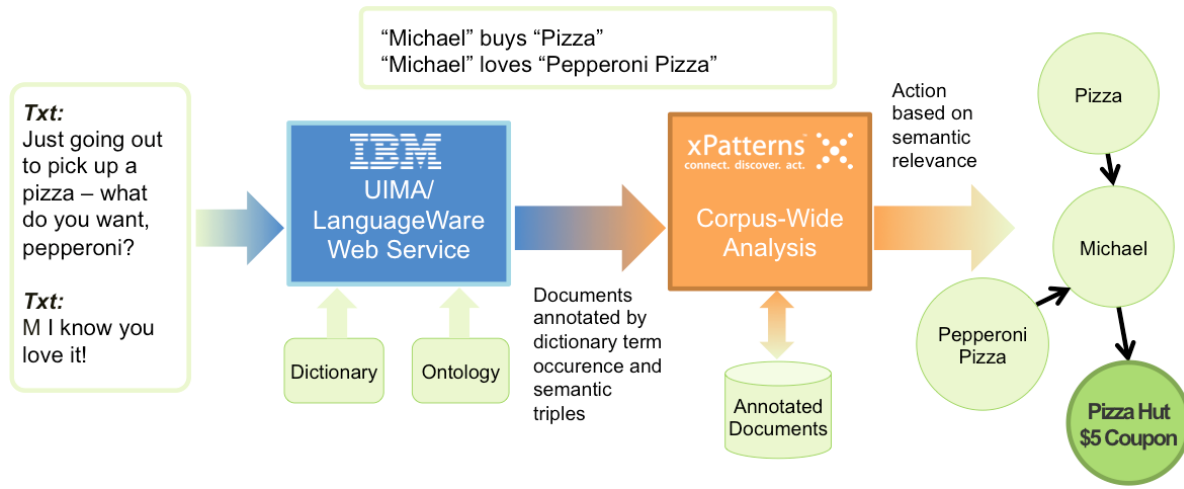


Diagram 2: LanguageWare and xPatterns Integration

Summary

- Extends the available analysis dimensions for visualization and reporting by business stakeholders.
- The consumer profile of the user and the annotated semantic relationships provide attributes that enable more relevant and actionable offers enriched by 3rd party ecosystem interactions through learning - the more xPatterns is used, the smarter and more useful it becomes.
- Empowers consumers to create and share their profile to receive the most relevant content and offers instead of being improperly targeted by historic behaviors and static profiles.
- Secure - rich consumer managed profiles, sets highest standard for legislative regulatory environments (i.e. Germany).
- xPatterns platform integrates easily with enterprise and third-party deployed solutions.
- It is possible to populate and take action on a rich user profile based on both explicit and implicit exhibitions of preference by the consumer.
- Produces measurable accelerations in performance against mission-critical KPI's and metrics. Creates new actionable attributes within the xPatterns persona manager to quantify KPI's.



For additional information, call the Atigeo team at 425.635.3900 or visit our website at www.atigeo.com.

© 2009 Atigeo. All rights reserved. The information contained in this document is confidential and/or proprietary to IBM, Atigeo and/or its affiliates and/or its agents.

The information contained herewith is intended only for use by the individual or entity to which it is addressed. – v12