Semantic AI
Bringing Machine Learning, NLP and Knowledge Graphs together

Andreas Blumauer
CEO & Managing Partner

Semantic Web Company
PoolParty Software Ltd
Graph-based Introduction
Current Status of Graphs & AI

Moving towards Semantic AI
## Largest change in market cap by company (2009 to 31 March 2018)

<table>
<thead>
<tr>
<th>Company name</th>
<th>Location</th>
<th>Industry</th>
<th>Change in market cap 2009-2019 ($bn)</th>
<th>Market cap 2018 ($bn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Apple</td>
<td>United States</td>
<td>Technology</td>
<td>757</td>
<td>851</td>
</tr>
<tr>
<td>2. Amazon.Com</td>
<td>United States</td>
<td>Consumer Services</td>
<td>670</td>
<td>701</td>
</tr>
<tr>
<td>3. Alphabet</td>
<td>United States</td>
<td>Technology</td>
<td>609</td>
<td>719</td>
</tr>
<tr>
<td>4. Microsoft Corp</td>
<td>United States</td>
<td>Technology</td>
<td>540</td>
<td>703</td>
</tr>
<tr>
<td>5. Tencent Holdings</td>
<td>China</td>
<td>Technology</td>
<td>483</td>
<td>496</td>
</tr>
<tr>
<td>6. Facebook</td>
<td>United States</td>
<td>Technology</td>
<td>383(1)</td>
<td>464</td>
</tr>
<tr>
<td>7. Berkshire Hathaway</td>
<td>United States</td>
<td>Financial</td>
<td>358</td>
<td>492</td>
</tr>
<tr>
<td>8. Alibaba</td>
<td>China</td>
<td>Consumer Services</td>
<td>302(1)</td>
<td>470</td>
</tr>
<tr>
<td>9. JPMorgan Chase</td>
<td>United States</td>
<td>Financial</td>
<td>275</td>
<td>375</td>
</tr>
</tbody>
</table>

**Source:** Collapsing the IT Stack: Clearing a path for AI adoption
Alan Morrison (Sr. Research Fellow at PwC)
“Knowledge graphs silently accrue ‘smart data’ — i.e., data that can be easily read and ‘understood’ by AI systems.”

“Once structured in the form of a knowledge graph, unstructured data can be queried, thereby preprocessing it for analysis.”
## The fast growing Graph Database Market

<table>
<thead>
<tr>
<th>Property Graph</th>
<th>RDF Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main use case</strong></td>
<td>Traverse a graph</td>
</tr>
<tr>
<td><strong>Typical applications</strong></td>
<td>Path Analytics, Social Network Analysis</td>
</tr>
<tr>
<td><strong>Standards</strong></td>
<td>No standards → Gremlin, Cypher, PGQL, ...</td>
</tr>
<tr>
<td><strong>Additional options</strong></td>
<td>Shortest path calculations</td>
</tr>
</tbody>
</table>
Core Principles
Foundations of Semantic AI
“Things but not Strings”: Semantic Knowledge

Graphs manage resources, not just terms

Retina

- The retina is the light-sensitive layer of tissue at the back of the eyeball. Images that come through the eye's lens are focused on the retina. The retina then converts these images to electric signals and sends them along the optic nerve to the brain.

- The retina usually looks red or orange because there are many blood vessels right behind it. An ophthalmoscope allows a healthcare provider to see through your pupil and lens to the retina. Sometimes, photos or special scans of the retina can show things that the provider cannot see just by looking at the retina through the ophthalmoscope. If other eye problems block the provider's view of the retina, ultrasound can be used.

Anyone who experiences these vision problems should get a retinal examination:

- Changes in sharpness of vision
- Loss of color perception
- Flashes of light or floaters
- Distorted vision (straight lines look wavy)

Watch this video about: Retina

Funduscope

Ophthalmoscope

Diagnostic Equipment
Core Principle

The Semantic Layer completes the Four-layered Data & Content Architecture
### Six Core Aspects of Semantic AI

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.</strong> Data Quality</td>
<td><strong>2.</strong> Data as a Service</td>
<td><strong>3.</strong> No black-box</td>
</tr>
<tr>
<td>Semantically enriched data serves as a basis for better data quality and provides more options for feature extraction.</td>
<td>Linked data based on W3C Standards can serve as an enterprise-wide data platform and helps to provide training data for machine learning in a more cost-efficient way.</td>
<td>Semantic AI ultimately leads to AI governance that works on three layers: technically, ethically, and on the legal layer.</td>
</tr>
<tr>
<td><strong>4.</strong> Hybrid approach</td>
<td><strong>5.</strong> Structured data meets text</td>
<td><strong>6.</strong> Towards self optimizing machines</td>
</tr>
<tr>
<td>Semantic AI is the combination of methods derived from symbolic AI and statistical AI. It is not only focused on process automation, but also on intelligence augmentation.</td>
<td>Most machine learning algorithms work well either with text or with structured data. Semantic AI is based on entity-centric data models.</td>
<td>ML can help to extend knowledge graphs, and in return, knowledge graphs can help to improve ML algorithms.</td>
</tr>
</tbody>
</table>
Data Quality

Training data is semantically enriched with help from semantic knowledge models.

PoolParty Semantic Classifier combines machine learning algorithms (SVM, Deep Learning, Naive Bayes, etc.) with Semantic Knowledge Graphs.
Data as a Service

Proposal for a Cognitive Computing Platform Architecture

Unstructured Data → Machine Learning

Structured Data → Semantic Layer

Knowledge Graphs → Semantic Layer

Cognitive Applications
No black-box Explainable AI (XAI)

Explaining a text classification prediction made by PoolParty Semantic Suite, highlighting positive concepts and terms.
Towards self-optimizing machines

The Singularity is just around the corner 😊

Mike Bergman (2014): Knowledge-based Artificial Intelligence

- Greater scope
- Fewer errors

- More features
- More precision

KBAI Virtuous Circle
How to build a Knowledge Graph?

Anatomy of an Enterprise Knowledge Graph
The Peggy Guggenheim Collection is a modern art museum on the Grand Canal in the Dorsoduro sestiere of Venice, Italy.
Linked Data Life Cycle

How to build Enterprise Knowledge Graphs?

**ANALYSIS**
Insights from linked data
Discover and analyze complex - and often hidden - relations in your data by means of customized data portals and state-of-the-art data visualization based on Semantic Data Lakes.

**PROVISIONING**
Data-driven personalization
Use semantic technologies to make your content and data management workflows smarter and more dynamic. Create personalized digital experiences. Feed your bots and machine learning algorithms with higher-quality data.

**ENRICHMENT**
Actionable data
Provide automatic tagging functions based on controlled vocabularies to ensure precise content classification. Semantic models help to enrich all content and data consistently and based on standards.

**INGESTION**
Avoid data migration
PoolParty is a semantic middleware that enriches and contextualizes your data with relevant information. Link and integrate different datasets by applying semantic knowledge models. Enterprise data gets smarter while remaining in the original storage system.

**CLEANING**
Data quality management
PoolParty allows organizations to ensure that their data is consistent, reliable and usable for analysis. Benefit from highly automated ‘Data Repair’ features.

**AUTHORING**
Knowledge graph capabilities
Enable collaboration and develop synergies between your departments by establishing defined knowledge graphs and controlled vocabularies.

**LINKING**
The power of connected data
Link internal and external data resources and provide your organization with a customized and relevant knowledge platform.
Use Cases for Enterprise Knowledge Graphs

Semantic Layer, Linked Data Life Cycle
1. Deal with **hierarchical or highly connected datasets** more efficiently
2. Gain new insights based on **entity-centric views** (in contrast to document-centric views)
3. Understand and calculate **causalities and the effects in a knowledge domain**
4. **Integrate heterogeneous data sources** (structured & unstructured) based on a “schema-late” approach
5. Create **federated (unified) views across multiple data silos** within the enterprise
The potential of AI spans across industries and function. The digital workplace is becoming enhanced by smart applications. Business leaders have to decide where to start. They need to keep pace with the technological advancement and how their competitors embrace it.
As a researcher in pharmaceutical industry, I want to plan new experiments more efficiently. I want to know what’s already available. I’m interested in former experiments where

- certain genes were tested
- under specific treatment conditions
- in a target therapeutic area
- with help from categorisation systems like ‘disease hierarchies’

→ Linking Structured to Unstructured Data and to Industry Knowledge Graphs
Knowledge Graphs serve as means to enrich unstructured information to provide a rich set of additional access points to document repositories.
With a knowledge graph, search can understand what doctors are really asking for and serve them the information they meant to search for:

**Insomnia**

### Search Results Without Graph

- May be treated with...
  - zolpidem
- Related to...
  - prednisone

### Search Results With Graph

- May be caused by...
  - prednisone
  - tolcapone
  - restless legs syndrome
  - iron deficiency
- May be treated with...
  - zolpidem
  - pregabalin
  - iron sulfate
Andreas Blumauer
Director, PoolParty Software Ltd
CEO, Semantic Web Company GmbH

- andreas.blumauer@semantic-web.com
- http://linkedin.com/in/andreasblumauer
- https://twitter.com/semwebcompany