LINKED DATA - THE NEXT 5 YEARS

From Hype to Action

Andreas Blumauer
CEO & Managing Partner
Semantic Web Company / PoolParty Semantic Suite
2

INTRODUCING SEMANTIC WEB COMPANY AND POOLPARTY

Semantic Web Company
- Founded in 2004
- Based in Vienna
- Privately held
- >30 employees, experts in text mining, data science & linked data
- SWC participates in EU-projects with a total funding of over € 17.0 million
- SWC named to KMWorld’s 2016 "100 Companies That Matter in Knowledge Management"

PoolParty Semantic Suite
- First release in 2009
- Current version 5.5
- W3C standards compliant
- Over 150 installations world-wide
- 50% of SWC’s revenue is reinvested into development of PoolParty
- To be installed on-premises or to be used as cloud service
- KMWorld listed PoolParty as Trend-Setting Product 2015 and 2016
“In the next 5 years we will see thousands of enterprises and many other organisations making use of linked data and semantic web technologies, but only when we are able to translate generic challenges and technical opportunities into specific, concrete and relevant use cases.”
Three generic challenges and hundreds of use cases

- Make informed decisions
- Link data and communicate across boundaries
- Context-aware and adaptive information supply

**Nutrition**
Food-Drug interactions
Inferring Cuisine - Drug Interactions

**Crisis Management**
Lack of data interoperability
Benefits of linked data for interoperability during crisis management

**Law**
Complex regulatory systems
Linked Legal Data

**Education**
Personalised learning pathways
A New Digital Education for Young Australians in the 21st Century

**Clean Energy**
GHG emissions
Well-founded Policy Making

**Healthcare**
Improving health literacy
The Semantic Content Engine of healthdirect Australia
How to communicate the technological advantages?

- Linked Data is an agile data integration technology
- RDF stores scale
- Semantic Web is standards-based
- Knowledge graphs can grow over time
- With Linked Data, unstructured and structured information can be handled
- Even complex queries can be executed with SPARQL
- Linked Data goes beyond relational data
- With Linked Data, you can load data, content, and metadata without the pain of upfront data modeling
- ...
From technologies over data to applications
Gartner’s Hype Cycle for Advanced Analytics and Data Science, 2015
Gartner’s Hype Cycle for Big Data, 2012
Escaping from the Trough of Disillusionment

Climbing the Slope of Enlightenment

Slope of Enlightenment:

More instances of how the technology can benefit the enterprise start to crystallize and become more widely understood.

Second- and third-generation products appear from technology providers.

More enterprises fund pilots.

Conservative companies remain cautious.

From: https://en.wikipedia.org/wiki/Hype_cycle
Climbing the Slope of Enlightenment: From Hype to Action - Some ingredients

- **Learning from the past**
  Main obstacles for potential users

- **Roll-out scenarios**
  Best practices for implementing linked data on a larger scale

- **Planning a semantic project**
  ‘Low-hanging fruits’ and a concrete action plan

- **Semantic Web eco-systems**
  Data and members, you would like to have linked

- **Complementary technologies**
  Other technologies as enablers for enterprise linked data

- **Integration & Network effects**
  How to engage end-users and the community
Some of our Customer References

- Credit Suisse
- Boehringer Ingelheim
- Roche
- adidas
- The Pokémon Company
- Canadian Broadcasting Corporation
- SwissRe
- Wolters Kluwer
- Bank of America
- HealthStream
- TC Media
- Techtarget
- BMJ Publishing Group
- CafePress
- Pearson - Always Learning
- Education Services Australia
- American Physical Society
- Healthdirect Australia
- World Bank Group
- Inter-American Development Bank
- Renewable Energy Partnership
- Wood MacKenzie
- Oxford University Press
- International Atomic Energy Agency
- Norwegian Directorate of Immigration
- Ministry of Finance (AT)
- Council of the E.U.
- Australian National Data Service

Some of our Partners

- Accenture
- EPAM Systems
- Tellura
- Term Management
- Taxonomy Strategies
- MarkLogic
- Solnet Solutions
- Wolters Kluwer
- Mekon
- Ontotext
Emerging technologies: Supporting the Pioneers

Top-down approach
“Big Bang”
→ Lack of down-to-earth approach

Bottom-up approach
“Quick win”
→ Lack of strategic relevancy

Revolutionists
“Make all things new”
→ Lack of implementation expertise

Sustainers
“Not invented here”
→ Lack of innovative capacity
13

How to disarm Killer Phrases

**Top-down approach**
- Application scenarios are unclear
  - It cannot be integrated
  - It cannot be managed

**Revolutionists**
- It doesn’t work with R
- It doesn’t support cognitive computing
- Doesn’t support NLQA

**Sustainers**
- It doesn’t scale
- It isn’t secure
- It isn’t well supported

**Bottom-up approach**
- It’s too expensive
- It’s too academic
- It cannot be sustained
What is ‘it’ really?

Understand and emphasize the fundamental differences
‘Things’ but not Strings:
Using a ‘Semantic Knowledge Graph’

Retina

HTTP://WWW.MY.COM/TAXONOMY/62346723

Image

http://www.my.com/images/90546089

Funduscope

HTTP://WWW.MY.COM/TAXONOMY/97345854

altLabel

Ophthalmoscope

HAS BROADER

Diagnostic Equipment

HTTP://WWW.MY.COM/TAXONOMY/4543567

prefLabel
Climbing the Slope of Enlightenment: Some ingredients

- Learning from the past
  Main obstacles for potential users

- **Roll-out scenarios**
  Best practices for implementing linked data on a larger scale

- Planning a semantic project
  ‘Low-hanging fruits’ and a concrete action plan

- **Semantic Web eco-systems**
  Data and members, you would like to have linked

- Complementary technologies
  Other technologies as enablers for enterprise linked data

- Integration & Network effects
  How to engage end-users and the community
Build capacities first

Top-down approach
- Define benefits and KPIs
- Specify a target architecture
- Define roles and responsibilities

Revolutionists
- Realistic targets
- Semantics as a ‘horizontal’ technology
- Envision something ‘really game-changing’

Sustainers
- Software evaluation
- Benchmark tests
- Make a PoC

Bottom-up approach
- Develop a Linked Data Strategy
- Understand the benefits of standards
- See the network effects, overcome departmental thinking
1.1 How to use Semantic Technologies for business and technology challenges

In this learning tutorial you will get an overview of tools & methods to develop semantic applications. You will also learn in which business domains it can be useful to embrace semantic solutions.

- Become productive quickly
- Three e-learning tracks
- Certification program
- Complemented by tailored workshops

https://www.poolparty.biz/academy/
Climbing the Slope of Enlightenment: Some ingredients

- Learning from the past
  Main obstacles for potential users

- Roll-out scenarios
  Best practices for implementing linked data on a larger scale

- Planning a semantic project
  ‘Low-hanging fruits’ and a concrete action plan

- Semantic Web eco-systems
  Data and members, you would like to have linked

- Complementary technologies
  Other technologies as enablers for enterprise linked data

- Integration & Network effects
  How to engage end-users and the community
Planning a Semantic Project:
Some building blocks in a multi-stakeholder environment

- Describe the use case / business case
- Specify target applications and KPIs
- Develop and implement in an agile way

**Software Development**
- Analyse available and relevant data sources
- Reuse / build taxonomies and ontologies
- Make data available, focus on data quality
- Make architectural decisions
- Develop a URI strategy
- Transform, map and link data

**Data Engineering**
- Transfer knowledge to pioneers / key people
- Make workshops / provide e-learning to project team
- Explain benefit to the end-users, engage the community

**Capacity Building**
Aligned Software & Data Development Life Cycle

http://aligned-project.eu/
Climbing the Slope of Enlightenment:

Some ingredients

- Learning from the past
  Main obstacles for potential users

- Roll-out scenarios
  Best practices for implementing linked data on a larger scale

- Planning a semantic project
  ‘Low-hanging fruits’ and a concrete action plan

- **Semantic Web eco-systems**
  Data and members, you would like to have linked

- Complementary technologies
  Other technologies as enablers for enterprise linked data

- Integration & Network effects
  How to engage end-users and the community
Identify Data Sets and Ontologies

Example for a Linked Data Portfolio

Degree of Maintenance Service Level

Focus on commercialization
Current Allotrope Foundation members include:

- AbbVie
- Amgen
- Baxter
- Bayer
- Biogen
- Boehringer Ingelheim
- Bristol-Myers Squibb
- Eli Lilly
- Genentech/Roche
- GlaxoSmithKline
- Merck & Co.
- Pfizer

Utilize, implement, and integrate data standards into analytical laboratory workflows.

Contains:
- Method, instrument, sample, process, result, etc.
- Data cube metadata
- Data package metadata
- ...

Analytical data represented by one- or multidimensional arrays.

Analytical data represented by arbitrary formats, incl. native instrument formats, images, pdf, video, etc.

Specifically designed to store and organize large amounts of numerical data.

http://www.allotrope.org/
Best practice: Australian National Data Service (ANDS)

Current ANDS partners include:

- Over 40 research institutions
- 12 NCRIS data intensive partners
- 6 high performance computing partners
- 17 partners from public sector
- 5 partners from cultural heritage

When research communities agree to use common language for the concepts in datasets, then the discovery, linking, understanding and reuse of research data are improved.
Climbing the Slope of Enlightenment:

Some ingredients

- Learning from the past
  Main obstacles for potential users

- Roll-out scenarios
  Best practices for implementing linked data on a larger scale

- Planning a semantic project
  ‘Low-hanging fruits’ and a concrete action plan

- Semantic Web eco-systems
  Data and members, you would like to have linked

- Complementary technologies
  Other technologies as enablers for enterprise linked data

- Integration & Network effects
  How to engage end-users and the community
Change is the only constant:
Gartner’s Hype Cycle for Emerging Technologies, 2016
Data Integration & Mapping of Knowledge: Basis for a Cognitive Approach

- Map Metadata with Knowledge Graph
- New Information within Documents can be integrated into the Knowledge Model

- Dividing Documents into Building Blocks
- Information Extraction
- PowerTagging
Change the way how you think that change will happen:

When is the AI revolution coming?

Graph by Jeremy Howard from his TED talk “The wonderful and terrifying implications of computers that can learn.”
Climbing the Slope of Enlightenment:

Some ingredients

- Learning from the past
  Main obstacles for potential users

- Roll-out scenarios
  Best practices for implementing linked data on a larger scale

- Planning a semantic project
  ‘Low-hanging fruits’ and a concrete action plan

- Semantic Web eco-systems
  Data and members, you would like to have linked

- Complementary technologies
  Other technologies as enablers for enterprise linked data

- Integration & Network effects
  How to engage end-users and the community
Make services easily accessible
Andreas Blumauer
CEO, Semantic Web Company

- a.blumauer@semantic-web.at
- http://at.linkedin.com/in/andreasblumauer
- https://twitter.com/semwebcompany
- https://ablvienna.wordpress.com/