Andreas Blumauer
CEO, Semantic Web Company

Helmut Nagy
COO, Semantic Web Company
<table>
<thead>
<tr>
<th><strong>Semantic Web Company</strong></th>
<th><strong>PoolParty Semantic Suite</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ Founded in 2004</td>
<td>▶ First release in 2009</td>
</tr>
<tr>
<td>▶ Based in Vienna</td>
<td>▶ Current version 5.4</td>
</tr>
<tr>
<td>▶ Privately held</td>
<td>▶ W3C standards compliant</td>
</tr>
<tr>
<td>▶ &gt;30 employees, experts in knowledge modelling, text mining &amp; linked data</td>
<td>▶ Over 100 installations worldwide</td>
</tr>
<tr>
<td>▶ SWC participates in EU-projects with a total funding of over € 17.0 million</td>
<td>▶ 50% of SWC’s revenue is reinvested into development of PoolParty</td>
</tr>
<tr>
<td>▶ SWC named to KMWorld’s 2016 &quot;100 Companies That Matter in Knowledge Management&quot;</td>
<td>▶ PoolParty can be installed on-premises or used as cloud service</td>
</tr>
<tr>
<td></td>
<td>▶ KMWorld listed PoolParty as Trend-Setting Product 2015</td>
</tr>
</tbody>
</table>
SELECTED CUSTOMER REFERENCES AND PARTNERS

Customer References

- Credit Suisse
- Boehringer Ingelheim
- Roche
- adidas
- The Pokémon Company
- Canadian Broadcasting Corporation
- Red Bull Media House
- 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

Partners

- Accenture
- EPAM Systems
- Enterprise Knowledge
- Term Management
- Taxonomy Strategies
- MarkLogic
- Solnet Solutions
- Wolters Kluwer
- Mekon
- Tellura
Integration of structured and unstructured data
Agile schema linking through explicit semantic data models
Entity linking and document linking through highly precise entity extraction
Supervised learning and machine learning mechanisms (Taxonomy learning and corpus learning)
Data Quality assessment & improvement mechanisms
Outstanding user-friendliness
Fully standards-compliant
Maintainable enterprise knowledge graphs
Search, connect, visualize, and analyze - all based on standards-based graph databases and distributed APIs.
MAKE USE OF POOLPARTY SEMANTIC SUITE

TECHNICAL COMPONENTS OVERVIEW

Data Portals & Collaboration Platforms
- Recommender System
- Semantic Search
- Analytics & Visualization

Knowledge Engineering & Graph Management
- Taxonomy & Thesaurus Management
- Ontology Management
- Linked Data Management

Content Enrichment & Data Integration
- Concept Tagging
- Text Mining & Entity Extraction
- Data Linking & Mapping
SEE HOW IT WORKS

Benefiting from Agile Data Integration in a Nutshell
360-degree views over various content repositories
PoolParty Semantic Integrator - at a glance

https://youtu.be/l_LppfS3wxk

Unstructured Data

Structured Data

Semantic Integrator

Deep Data Analytics

ETL / Monitoring / Scheduling

MarkLogic

Virtuoso Universal Server

Ontotext

Stardog

Apache Solr

elasticsearch
PoolParty Semantic Integrator

High-level architecture
PoolParty Taxonomy Manager

PoolParty Taxonomy Manager allows the management of labels (e.g. synonyms), hierarchies and non-hierarchical relations, and mappings to other vocabularies.

Also more complex actions like merging of concepts, moving of subtrees or the creation of poly-hierarchies are supported.

PoolParty fully covers the SKOS standard of W3C incl. SKOS-XL and SKOS Collections.
PoolParty Ontology Editor

SKOS is based on a simple schema. This can be expanded by custom schemes based on ontologies.

Custom schemes can be used to map RDF graphs to relational data models.

For an increased interoperability, PoolParty provides a rich set of preconfigured ontologies like schema.org, FIBO or FOAF.
RDF based ETL

Data processing tasks can be modelled as pipelines: Make use of the intuitively usable graphical interface.

Versatile data integration platform: Link data from internal and external data sources in a central NoSQL linked data warehouse.

Custom plugins: Your data processing pipelines are highly customizable by creating your own data processing units (DPUs).
PoolParty Extractor

PoolParty’s API provides a rich set of methods for text mining and entity extraction. This ultra-fast service makes use of your controlled vocabularies, therefore it is highly accurate for your specific domain.

The service will improve over time and learns from reference text corpora. It comes with a powerful disambiguation algorithm and can be used as part of a data processing unit (DPU).
GraphSearch

Semantic search at the highest level: PoolParty Graph Search Server combines the power of graph databases and SPARQL engines with features of ‘traditional’ search engines.

Document search and visual analytics: Benefit from additional insights through interactive visualizations of reports and search results derived from your data lake by executing sophisticated SPARQL queries.
The S tronic combines the sporty characteristics of a manual gearbox with the advantages of an automatic.

Gear-changes with the dual-clutch gearbox are performed easily and with no appreciable delay. Depending on the driver’s preference, the gears may be shifted in fully automatic mode or in manual mode using the gearshift paddles on the steering wheel.

The dual clutch forms the technical basis of the S tronic. One of the two clutches engages the odd-numbered gears and reverse, the other engages the even-numbered gears. The benefit: a gearbox that shifts from one gear to the next in less than 0.2 seconds and with no interruption in power flow.

‘Things’, not Strings in Semantic Graphs
How many cars were produced in 2014 in Germany?

<Production>
    <place>
        Győr
    </place>
    <product>
        Audi A3
    </product>
    <year>
        2014
    </year>
    <volume>
        200.000
    </volume>
</Production>
Entity linking based on taxonomies

How many cars were produced in 2014 in Germany?

<table>
<thead>
<tr>
<th>Year</th>
<th>Location</th>
<th>Type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>Győr, Audi A4</td>
<td>100000</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>Curitiba, Audi A3</td>
<td>50000</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>Ingolstadt, A3</td>
<td>75000</td>
<td></td>
</tr>
</tbody>
</table>

Audi A3

Ingolstadt

A3/Ing/14

2014

75000
User-friendly modelling of rules for entity extraction via taxonomies

The Audi Q3 is a compact crossover SUV made by Audi.

It is based on the PQ35 platform of Volkswagen.
The Audi TT is a small two-door sports car marketed by Volkswagen Group subsidiary Audi since 1998, assembled by the Audi subsidiary Audi Hungaria Motor Kft. in Győr, Hungary, using bodyshells manufactured and painted at Audi’s Ingolstadt plant. This changed with the third generation model that uses parts made entirely by the Hungarian factory.

The styling of the Audi TT began in the spring of 1994 at the Volkswagen Group Design Center in California. The TT was first shown as a concept car at the 1995 Frankfurt Motor Show.
INTEGRATION WITH VIRTUOSO

Benefiting from a large-scale RDF Store
YOUR BENEFIT

Native database capability and a virtual database
- Performant SPARQL engine
- Massive Linked Data Graphs
- Transactions
- Scaling to trillions of triples
- Federated environments
- Built-in inferencing

Semantic Middleware for Enrichment and Linking
- Superior user friendliness
- Semantic as a Service
- Standards-based technology
- Precise document classification
- Graph-based metadata management
- Beyond search

GRAPH BASED ANALYTICS
- Data Integration
- Data Enrichment
- Intelligent Search
- Linked Data
- Data Virtualization
INTEGRATION WITH MARKLOGIC

Benefiting from a Full Semantics Stack
MarkLogic and PoolParty at a Glance

DATA SOURCES

- JSON
- XML
- TXT
- BINARY
- GEO
- RDF

PoolParty
- Taxonomy and Vocabulary Management
- Ontology Management
- Text Mining / Entity Extraction
- RDF-based ETL Management

MarkLogic Database
- Database
- Search
- Application Services

SEARCH

ANALYTICS

DOWNSTREAM SYSTEMS
YOUR BENEFIT

MarkLogic

Operational and Transactional Enterprise NoSQL Database
- Fast Time to Results
- Ask Anything Universal Index
- Trusted Data and Transactions
- Enterprise-Grade Security
- Scale-Out Commodity Hardware
- Lightning Fast and Real-Time

poolparty

Semantic Middleware for Enrichment and Linking
- Superior user friendliness
- Semantic as a Service
- Standards-based technology
- Precise document classification
- Graph-based metadata management
- Beyond search

FULL SEMANTICS STACK
- Data Integration
- Data Enrichment
- Intelligent Search
- Deep Analytics
- Data Governance
TECHNICAL DEEP DIVE

Learn more about some core workflows of PoolParty Semantic Integrator
Unified Views on structured & unstructured information
Pipelines and Data Processing Units (DPUs):

Transform all of your data into RDF and link it
Every Data Pipeline:
- Represents one data processing task
- Contains Data Processing Units (DPUs)
- Defines the data flow between those DPUs

Data Pipelines may be:
- Designed
- Debugged
- Executed
- Scheduled
Every Data Processing Unit:
- Declares mandatory or optional inputs
- Encapsulates certain business logic that processes the data, e.g. transforms CSV data to RDF data
- Produces certain outputs
- DPUs may also provide a configuration dialog, so that the DPU may be configured by a pipeline designer
- Administrators may set up default configurations

Four types of DPUs:
- Extract
- Transform
- Load
- Quality Assessment
Data Units

A container for data being consumed or produced by a DPU
We distinguish input and output data units

3 types of data units:
- RDF data unit
  - One or more RDF graphs
- Files data unit
  - One or more files
- Relational data unit
  - One or more relational tables
Example Pipelines

Benefit from Agile Data Integration
RDBMS to RDF transformation: Mapping of DB table to Ontology
RDBMS to RDF transformation: UnifiedViews transformation pipeline
Enrichment via Taxonomy: Tagging of content

The main objective of the iART project is to bring visual art into bespoke clothing so as to make apparel products more valuable compared to other garments in the market and to support artists to exploit their work. To reach this objective a set of innovative ICT solutions has to be developed first. This new approach will also change the business relationship along the conception and design phase in the clothing sector from B2B over B2C to C2C, where the artist and the consumer will connect through enabling platforms like iArt. iArt will be built upon an existing B2C webshop for bespoke shirts Bivolino.com. In particular, the specific ICT developments of the project are: 1. Upgrade the configuration toolkit from 2.5D to 2.5D/3D by including shadowing and rotating features without the need of plug-ins and ease the extension of the catalog to clothing accessories and nightwear fitting...
Enrichment via Taxonomy: UnifiedViews transformation pipeline
RDBMS to RDF transformation: GraphSearch

http://unifiedviews.poolparty.biz
<xml version="1.0" encoding="UTF-8"/>
<documents rows="10" os="0" page="1" total="48162">
  <doc id="2854478">
    <abstracts><![CDATA[<em>Poverty</em> in Guatemala is high and deep. In 2000, over half of all Guatemalans lived in <em>poverty</em>. ...]]>
    <subtopic>Health Monitoring & Evaluation, Achieving Shared Growth, Poverty Assessment, Health Economics & Finance, Services & Transfers to Poor</subtopic>
    <count>Guatemala</count>
  </doc>
</documents>
XML to RDF transformation: UnifiedViews transformation pipeline
Poverty in Guatemala is high and deep. In 2000, over half of all Guatemalans lived in poverty. 

Abstracts:

Poverty in Guatemala is high and deep. In 2000, over half of all Guatemalans lived in poverty.

Subtopics:

- Health Monitoring & Evaluation
- Achieving Shared Growth
- Poverty Assessment
- Health Economics & Finance
- Services & Transfers to Poor

Count:

Guatemala

Related terms:

- Central America
- Health Monitoring & Evaluation
- Labor Markets
- Shared Growth
- Achieving Shared Growth

Topical Taxonomy

Enrichment via Taxonomy: Mapping to taxonomy
Enrichment via Taxonomy: UnifiedViews transformation pipeline
XML to RDF transformation: GraphSearch

http://marklogic-demo.poolparty.biz
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/

Sorry, no Facebook!
Helmut Nagy
COO, Semantic Web Company

▸ h.nagy@semantic-web.at
▸ https://at.linkedin.com/in/helmutnagy
▸ https://twitter.com/semwebcompany
▸ https://blog.semantic-web.at/