

PoolParty - Technical White Paper

Introduction

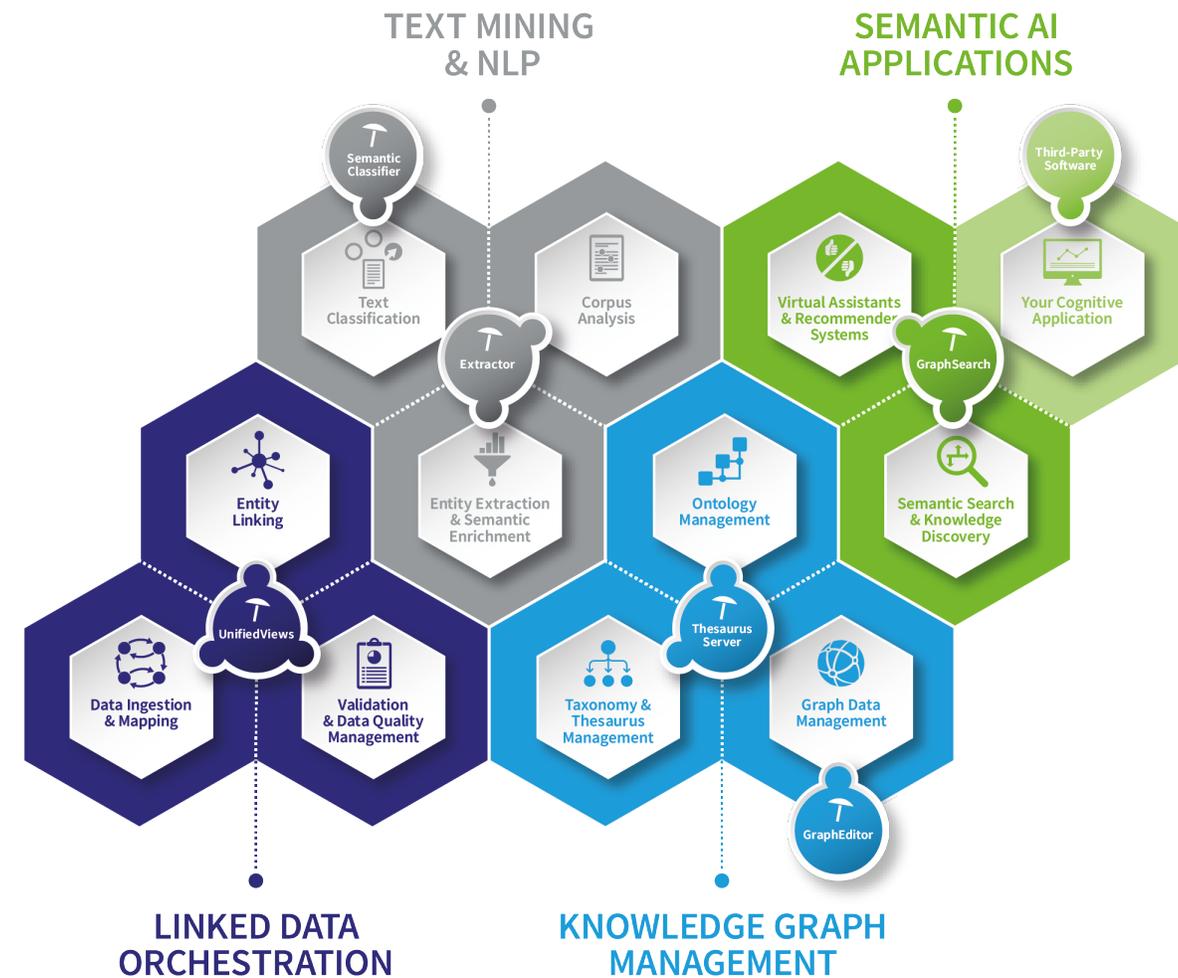
PoolParty® Semantic Suite (<https://www.poolparty.biz/>) is an AI platform based on semantic technologies and machine learning. It helps organizations to build and manage knowledge graphs as a basis for various AI applications. As a semantic middleware, PoolParty extracts the semantic meaning from your data and links your business objects and content assets automatically. Make your data actionable and benefit from smart applications!

**PoolParty
Semantic
Suite -
Components
and Features**

**PoolParty
Technical
Overview**

PoolParty technology platform consists of several components and can be configured and extended to meet individual requirements.

PoolParty Thesaurus Server supports web-based taxonomy and ontology management. It is completely built on top of W3C's Semantic Web standards (<http://www.w3.org/standards/semanticweb/>).



In its core, PoolParty uses the Resource Description Framework (RDF) to represent SKOS and any other ontology (FOAF, FIBO, Schema.org, etc.). For this reason, an RDF graph database (triple store) is used as its technological basis. Compared to other systems, which remain to be based on relational databases, PoolParty is ready to consume and to publish Linked Data out-of-the-box. Besides the possibility to publish any PoolParty based thesaurus and ontology via a Linked Data front-end, the system offers a SPARQL endpoint (<http://www.w3.org/TR/rdf-sparql-query/>) to execute queries over each thesaurus project. This technology can be used to integrate knowledge graphs with content platforms (Wikis, CMS, etc.) or search engines.

PoolParty GraphEditor complements Thesaurus Server functionalities and supports data engineers with functionalities to create, maintain and edit all types of knowledge graphs. GraphEditor allows to create ontology-based custom views on RDF data. This component also supports bulk editing and lets data engineers interact with RDF data without the need of having deep knowledge in SPARQL.

PoolParty Extractor supports highly scalable and precise entity extraction based on knowledge graphs as well as machine learning, which can be combined, put in series, or even used as parts of more complex rules and constraints for sophisticated text mining tasks. Its ability to transform structured and unstructured information into RDF offers new options for data analytics.

PoolParty Semantic Classifier works well together with Extractor while classifying whole text fragments or documents. It is based on machine learning algorithms like SVM, Deep Learning, Naive Bayes, and some others. It's well proven that Semantic Classifier is able to outperform other tools of this kind when using controlled vocabularies to label training documents based on an established domain knowledge model.

PoolParty UnifiedViews supports automation of various data management tasks along the whole Linked Data Life Cycle. Typical tasks fulfilled by UnifiedViews are data ingestion, data transformation, enrichment, entity linking, or data quality assurance. UnifiedViews provides a large library of Data Processing Units (DPUs) that can be used as parts of whole data processing pipelines. Configuration of such pipelines can be managed in a user-friendly way by using a graphical editor. Pipelines can be triggered automatically, scheduled, and monitored. By that means, linked data orchestration can be highly automated.

PoolParty GraphSearch serves as a component to make heterogeneous data better accessible to users or other third-party applications. Input sources can range from document repositories over spreadsheets to relational data. GraphSearch delivers integrated views on business objects (entities) while using knowledge graphs and linked data. Its API provides several methods to set up systems like semantic search, recommender engines, data portals, or chatbots. GraphSearch works with traditional, document-centric search technologies like Solr or Elastic, and can also make use of RDF graph databases.

As a result, PoolParty Semantic Integrator, which contains all components as mentioned above, is the most complete semantic middleware on the global market. It serves as a solution for data integration, text mining, semantic search, knowledge discovery, and data analytics resulting into a highly structured Semantic Data Lakes (Linked Data Warehouses) powered by SPARQL engines and reasoning in its core.

In addition to full support of SPARQL, PoolParty APIs offer 'traditional means' to integrate semantics into enterprise information systems and web platforms. Based on RESTful services and JSON, developers can make use of all CRUD methods necessary to maintain knowledge graphs from within a third-party application like a CMS.

PoolParty integrations have been implemented with content platforms like Drupal, SharePoint, Confluence, Alfresco, or WordPress. As an additional result, guidelines have been developed, which can be reused for other integration projects. PoolParty was also successfully integrated into search engines like Solr, Elasticsearch, Mindbreeze, Sinequa or Intrafind.

PoolParty System Architecture

PoolParty Semantic Suite serves as a semantic middleware (see diagram below) providing various APIs and GUIs

- to ingest data from various sources,
- to transform, enrich, and link those,
- to provide eventually high-quality, semantically enriched data as a basis for integrated views on all types of entities (business objects)

PoolParty components can be combined to specific product bundles offering various integration options.

PoolParty Semantic Suite - System Architecture

In the following chapters an overview of the different components is provided and integration options are outlined:

- [PoolParty Thesaurus Server](#)
- [PoolParty Extractor](#)
- [PoolParty GraphSearch Server](#)
- [PoolParty UnifiedViews](#)
- [Integrating PoolParty](#)

Summary

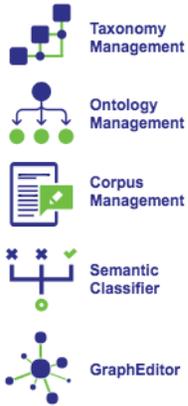
As a semantic middleware, PoolParty offers a wide range of options to benefit from semantic technologies and machine learning along the whole Linked Data Life Cycle. The major topics covered are: semantic search, taxonomies, ontologies, knowledge graph management, text mining, NLP, data integration and linked data. In its core, PoolParty uses state-of-the art semantic web technologies, which are built on top of open W3C standards. Professional metadata management is the key for efficient information management in any data-driven organization. PoolParty combines methodologies from the Semantic Web with machine learning and text mining algorithms, as well as with approaches for collaborative knowledge engineering. As a result, organizations benefit from a better data quality, integrated views on data, better reuse of existing knowledge, and end-users who love to work with smarter applications with great user experience.

Linked Data Life Cycle

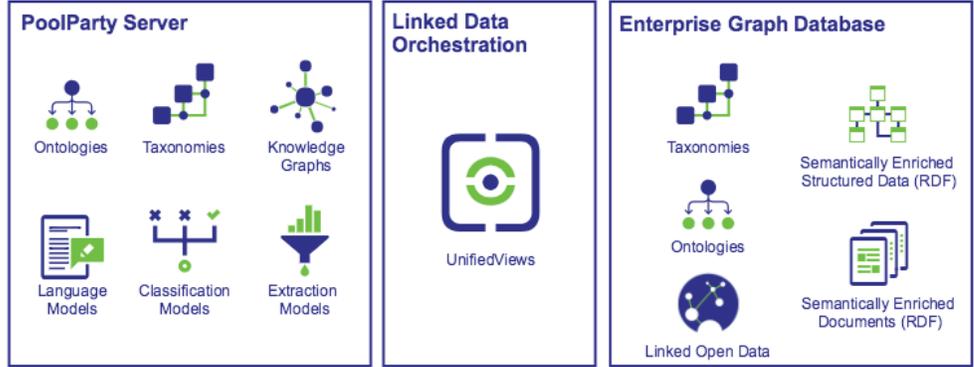
Semantic AI applications



PoolParty GUI



PoolParty Semantic Middleware (Semantic Integrator)



Data Sources



