Taxonomies, Knowledge Graphs, and AI: Delivering Targeted Learning Recommendations

Heather Hedden
Semantic Web Company

Semantic Web Company (SWC) and PoolParty

SWC is developer / vendor of **PoolParty Semantic Suite**

Most complete and secure **Semantic Middleware / Semantic AI platform** on the Global Market

W3C standards compliant

ISO 27001:2013 certified

First release in 2009

Current version **8.0**

On-premises or **cloud-based**

Over **200** **installations world-wide**

Semantic AI: Fusion of Graphs, NLP, and Machine Learning

Named as Visionary in **Gartner's Magic Quadrant** for Metadata Management Systems 2019-20

KMWorld listed PoolParty as one of the **Trend-Setting Products** 2015 - 2020 and listed SWC in the **AI 50** list of companies in 2020

© Semantic Web Company 2020

Outline

- Knowledge Graphs: Why, How, and What?
- Taxonomies: Knowledge Graph Foundation
- Text Mining: Term & Concept Extraction
- Ontologies: Knowledge Graph Structure

© Semantic Web Company 2020

Knowledge Graphs

Why, How, and What?

Better decisions
Customer satisfaction
Knowledge discovery
Structured and linked content yields knowledge
But how do I create it?
poolparty

Knowledge Graph Benefits

Better decisions

1. Fast, assisted, and precise **search and access** to all content and data.

2. Learn from content and data **visualized** in a **networked** and **contextualized** way.*

*e.g., Customer 360 Views

Customer satisfaction

3. **Personalized access** to all relevant content and data, inside and outside the enterprise.

4. Intelligent **recommender systems** uncover hidden information.

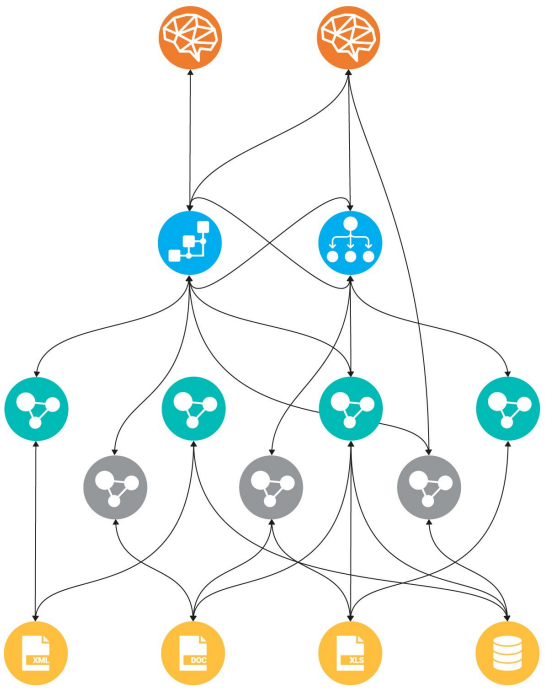
Knowledge discovery

5. Integrated **knowledge management systems** based on **knowledge graphs.**

6. **Analyze** and **explore** large amounts of unstructured data and documents.

© Semantic Web Company 2020

What is a Knowledge Graph?



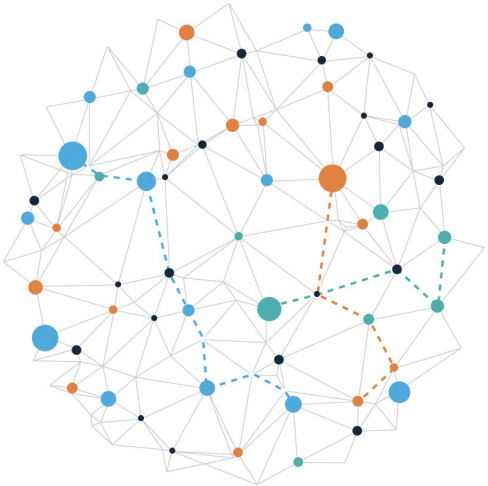
SEMANTIC AI APPLICATION

CONCEPTUAL AND LINGUISTIC MODEL

(VIRTUAL) DATA GRAPH

CONTENT & DATA LAYER

ENTERPRISE KNOWLEDGE GRAPH



An Enterprise Knowledge Graph contains business objects and topics that are closely linked, classified, semantically enriched, and connected to existing data and documents.

© Semantic Web Company 2020

What is a Knowledge Graph?

The Knowledge Engineer's perspective

A **model of a knowledge domain** created by subject-matter experts with the help of intelligent machine learning algorithms.

The Data Architect's perspective

As an additional **virtual data layer**, the Knowledge Graph lies on top of your existing databases or data sets to **link all your data together at scale.**

The Data Engineer's perspective

A **structured and common interface** for all of your data that enables the creation of smart multilateral relations throughout your databases.

© Semantic Web Company 2020

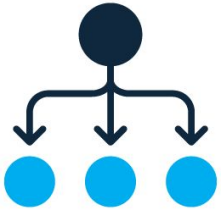
Taxonomies

Knowledge Graph Foundation

Taxonomies

Taxonomies + Ontology + Instance data = Knowledge graph

Why taxonomies (controlled vocabularies)?



- Provide consistent metadata
- Support consistent tagging
- Improve retrieval *precision* by focusing on concepts, not just words
- Improve retrieval *recall* by not missing content/data due to different synonyms or variant names
- Provides context for understanding through a hierarchical structure

© Semantic Web Company 2020

Taxonomy Standards

SKOS (Simple Knowledge Organization System)

- A World Wide Web Consortium (W3C) standard
- Represents knowledge organization systems using the Resource Description Framework (RDF)
- Machine-readable and human-understandable
- Supports interoperability.

ANSI/NISO Z39.19 *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*

ISO 25954 *Thesauri and Interoperability with other Vocabularies*

- Comprehensive best practices for thesaurus design and creation
- Supports good user experience and good results

© Semantic Web Company 2020

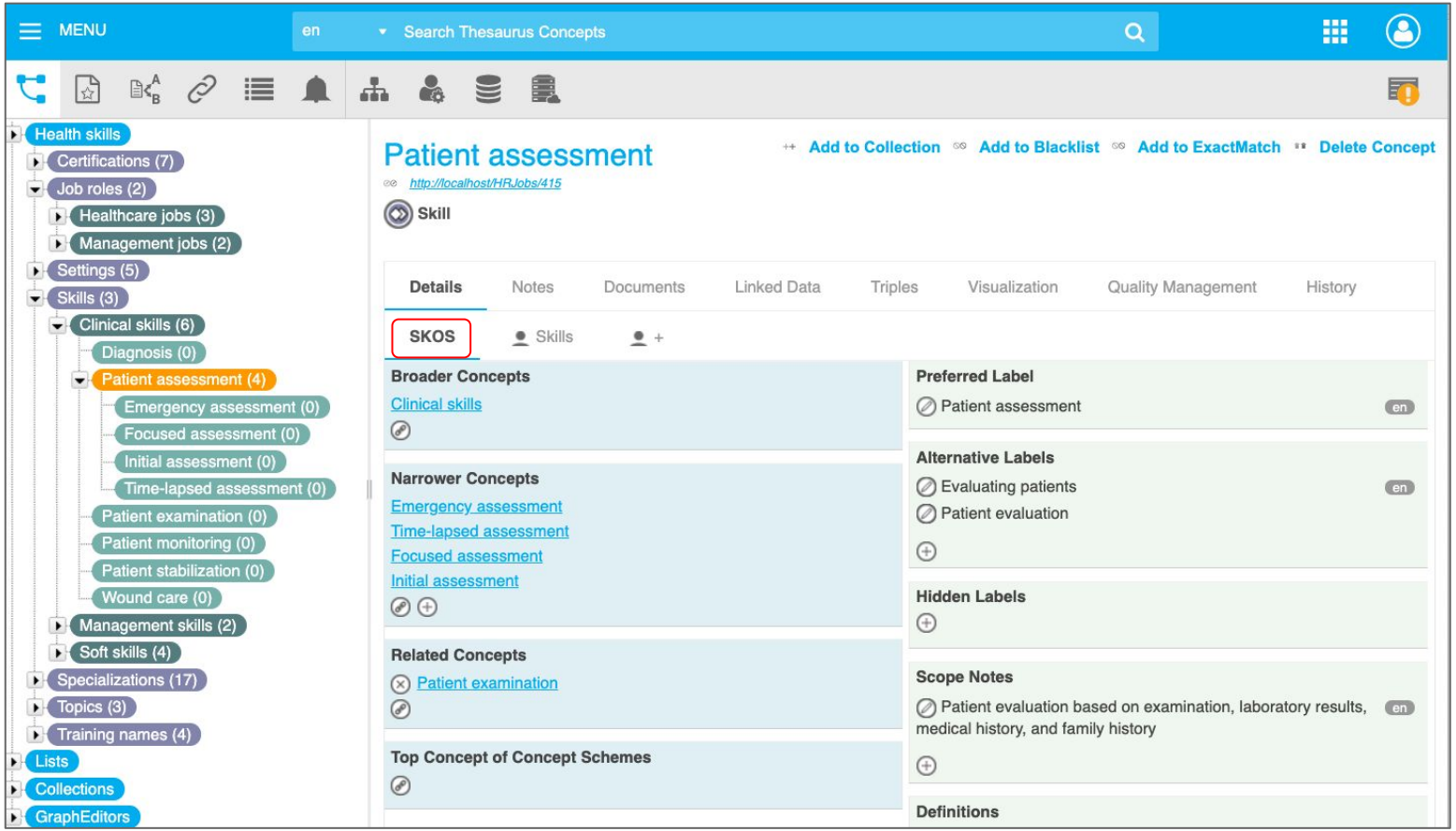
Taxonomy Management

Why use a dedicated taxonomy/thesaurus management tool, such as PoolParty Thesaurus Management?

- Compliant with W3C interoperability standards, **SKOS**, **RDF**, etc.
- Enforces quality standards, including **ANSI/NISO** and **ISO** thesaurus standards and custom rules:
  - No circular relations, no conflicting relations, no reused labels, etc.
- Supports multiple simultaneous users, user privileges, and workflows.
- Exports/imports spreadsheets and various RDF XML formats, and has APIs for additional integrations.
- Includes taxonomy enrichment tools:
  - Cardsorting, linking to public vocabularies, corpus term extraction
- Can be semantically enriched in the same software with an ontology.

© Semantic Web Company 2020

Taxonomy Management in PoolParty

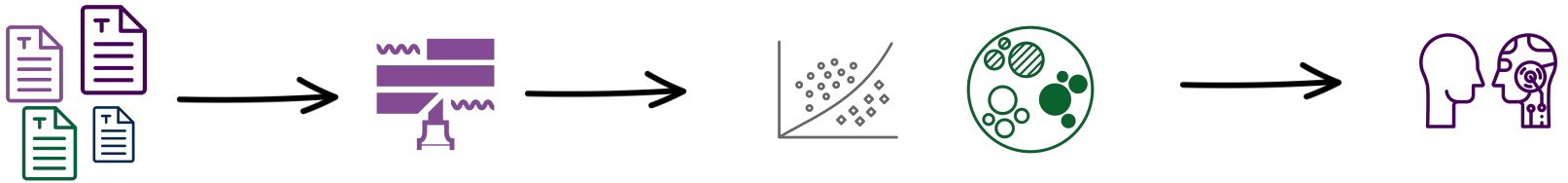


© Semantic Web Company 2020

The Taxonomist's NLP Toolkit

Starting from 10,000+ of documents, we can use machine learning to automate the generation of a robust set of hierarchically clustered terms that the taxonomist can organize into the final design.

Input Content → **NER** → **Clustering + Statistical Classifier** → **Human Augmented AI**



Nurse Sara spends most of her time in the operating room. She administers chemotherapy and radiation therapy. She has a Basic Life Support (BLS) certification.

Key Terms
Operating Room
Chemotherapy
Radiation Therapy
Basic Life Support (BLS)

Cluster (Topic)	Terms
1	Operating room
2	Chemotherapy Radiation Therapy
3	Basic Life Support

Cluster (Topic)	Terms
Clinical Setting	Operating room
Oncology Specialization	Chemotherapy Radiation Therapy
Certification	Basic Life Support (BLS)

© Semantic Web Company 2020

Text Mining

Term & Concept Extraction

Text Mining

What is text mining?

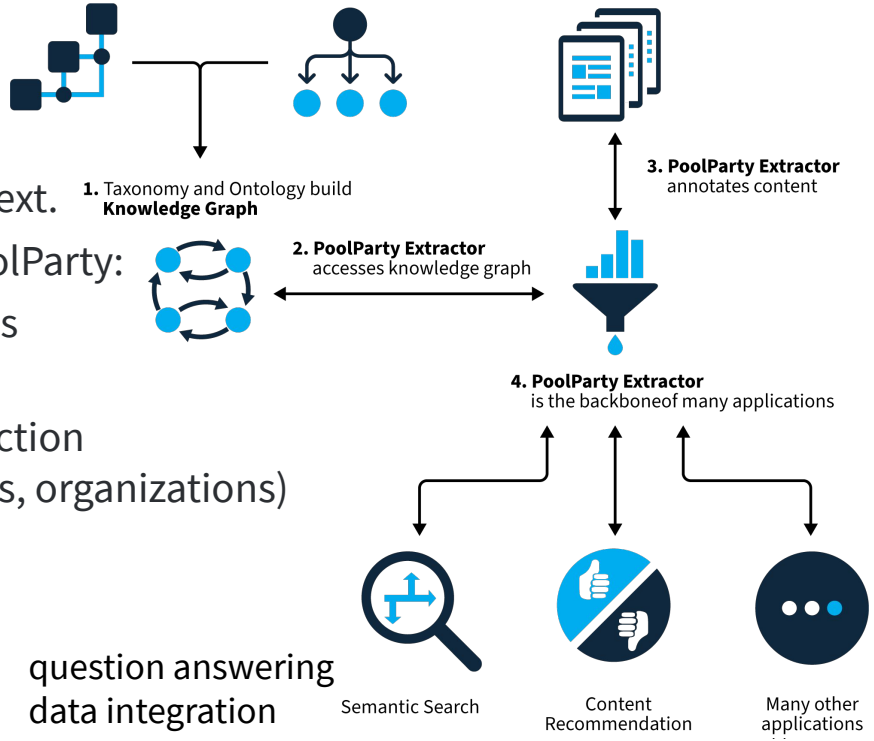
- An application of text analytics, utilizing AI technologies of Natural Language Processing (NLP).
- Extracting passages from text that are relevant in a particular business context.
- Automatically deriving information, and not merely strings of words.
- Transforming unstructured text into meaningful information.

© Semantic Web Company 2020

Text Mining in PoolParty

Text Mining Tool: **PoolParty Extractor**

- A software component of PoolParty that extracts meaningful phrases, named entities, and taxonomy concepts from documents and text.
- Leverages the taxonomies you have built in PoolParty:
  - Considers taxonomy concept relationships and alternative labels
- Performs named entity recognition (NER) extraction using pre-trained libraries (of persons, locations, organizations) and custom entities.
- Applications include:



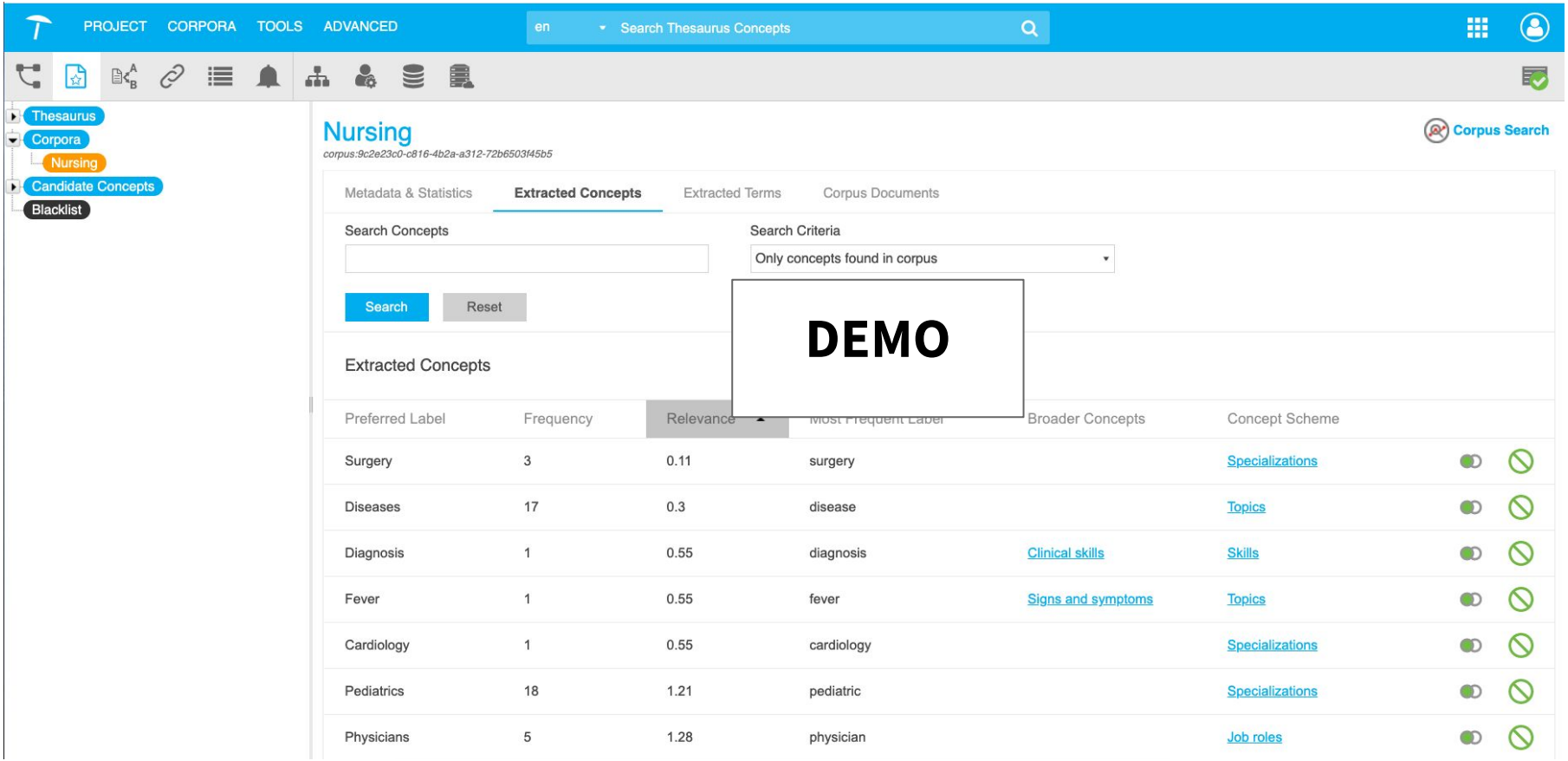
intelligent auto-tagging
categorization/classification

smart semantic search
content recommendation

question answering
data integration

© Semantic Web Company 2020

Text Mining in PoolParty



© Semantic Web Company 2020

Ontologies

Knowledge Graph Structure

Ontologies

Why ontologies?

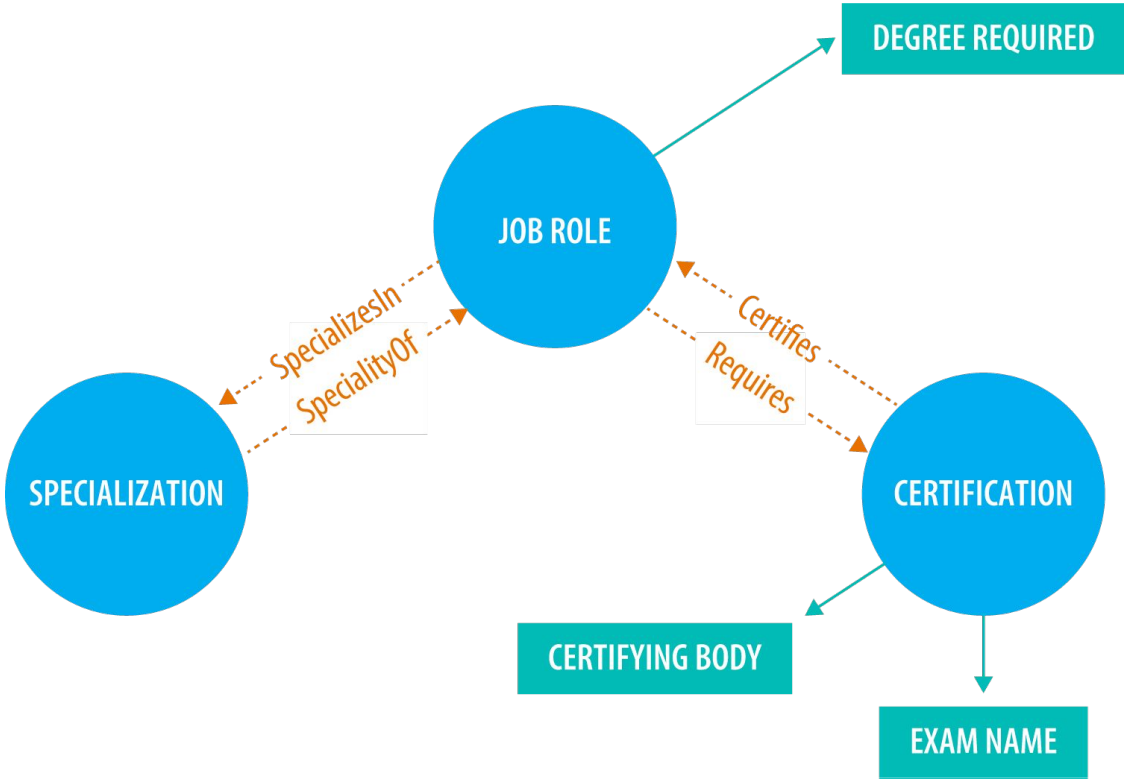
- Ontologies express the conceptual framework of knowledge graphs.
- Ontologies have semantic enrichment, which is lacking in mere taxonomies and thesauri:
  - Customized semantic relationships that explain *in what way* classes of concepts are related; and
  - Customized attributes (properties, metadata) for concepts.

Ontology standards

- **RDF** and **RDFS** (RDF-Schema)
- **OWL** (Web Ontology Language)

© Semantic Web Company 2020

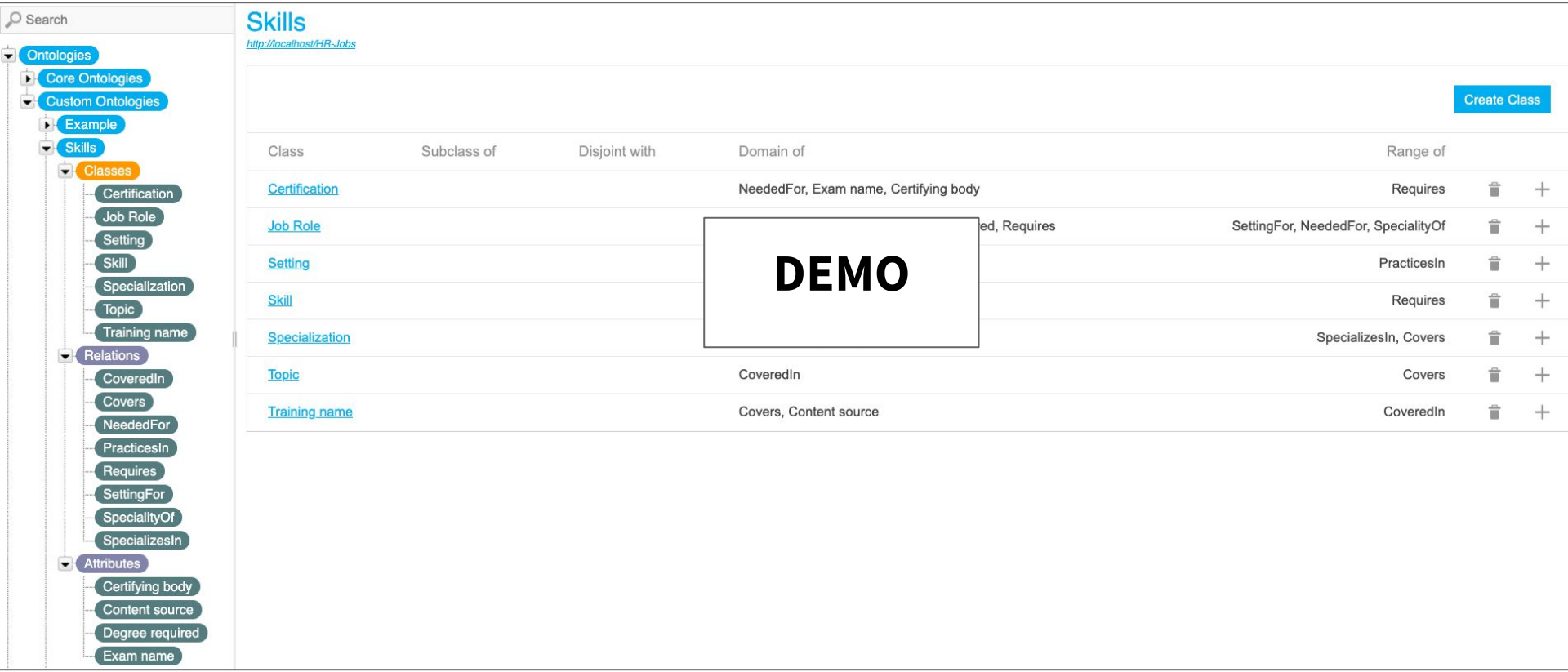
Ontology Features



- **Classes** – groups of concepts.
- **Relations** – semantic, customized relations between classes.
- **Attributes** – semantic, customized properties for classes.

© Semantic Web Company 2020

Ontology Management in PoolParty



Contact

Heather Hedden
heather.hedden@semantic-web.com
www.poolparty.biz

© Semantic Web Company 2020