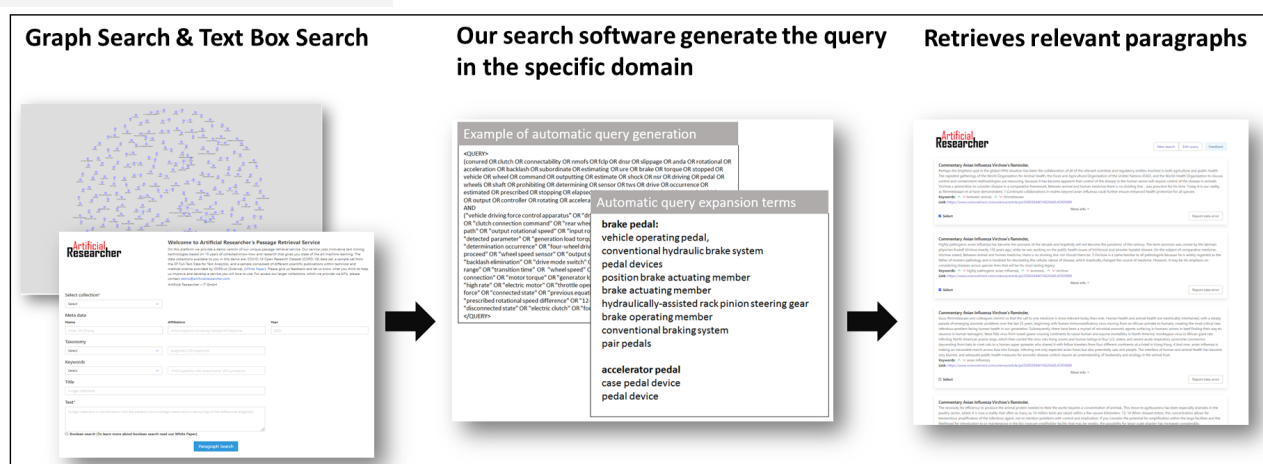


NEXT GENERATION SEARCH SOFTWARE

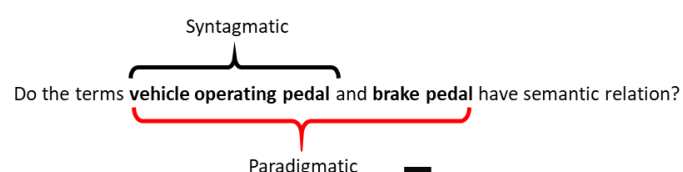
WE GENERATE SUSTAINABLE TEXT MINING SOLUTIONS, WHICH
GIVE USERS THE OPTION TO SEARCH FOR PRECISE
INFORMATION WITH TRANSPARENT AI.

- ⇒ Get ready package solutions for technical documents (client data, Open Access and Patent Data)
- ⇒ XML enhanced data
- ⇒ Streamline dataflow processes generating domain ontologies and indices
- ⇒ Access to Deep Learning tools for inhouse text mining

The Artificial Researcher (AR) search technology is based upon a two-time award-winning PhD research result from TU Wien. The Artificial Researcher Data Pipeline solution is deployed as a SaaS in cloud infrastructure setting. The services is scalable and a complete solution, which can process any type of machine-readable text data, and create sustainable ready to use indices and ontologies, as well as enhanced data formats, which can be integrated into client's own dataflow system. For easy deployment, the indices and ontologies can be packaged together with the **Artificial Researcher Passage Retrieval Service** ([link for developer](#) and [demo](#)) and the **Artificial Researcher Ontology Service** ([link for developers](#) and [demo](#)). By integrating domain ontologies into the information retrieval system, the AR technology provides automatic query expansions with understandable semantic information (Transparent Artificial Intelligence).



We offer Transparent AI by focusing on addressing the limitation and reducing the natural biases in Distributional Semantic models (BERT, w2v)



$$JoinedSimilarity = \sum_{\substack{i,j=1,n \\ i \neq j \\ i < j}}^N \frac{\cos(\vec{w_i}, \vec{w_j})}{N}$$

- w_i, w_j represent each word vector pair cosine similarity of an MWT
- N is the number of words for an MWT

The AR technology extracts single words and phrases by combining NLP, with domain specific modules for automatic domain term recognition, part of the **Artificial Researcher NLP-toolkit Services** (developer).

AR technology makes use of a domain modify NLP module, as well as an assembly of similarity values targeting the semantic functions:

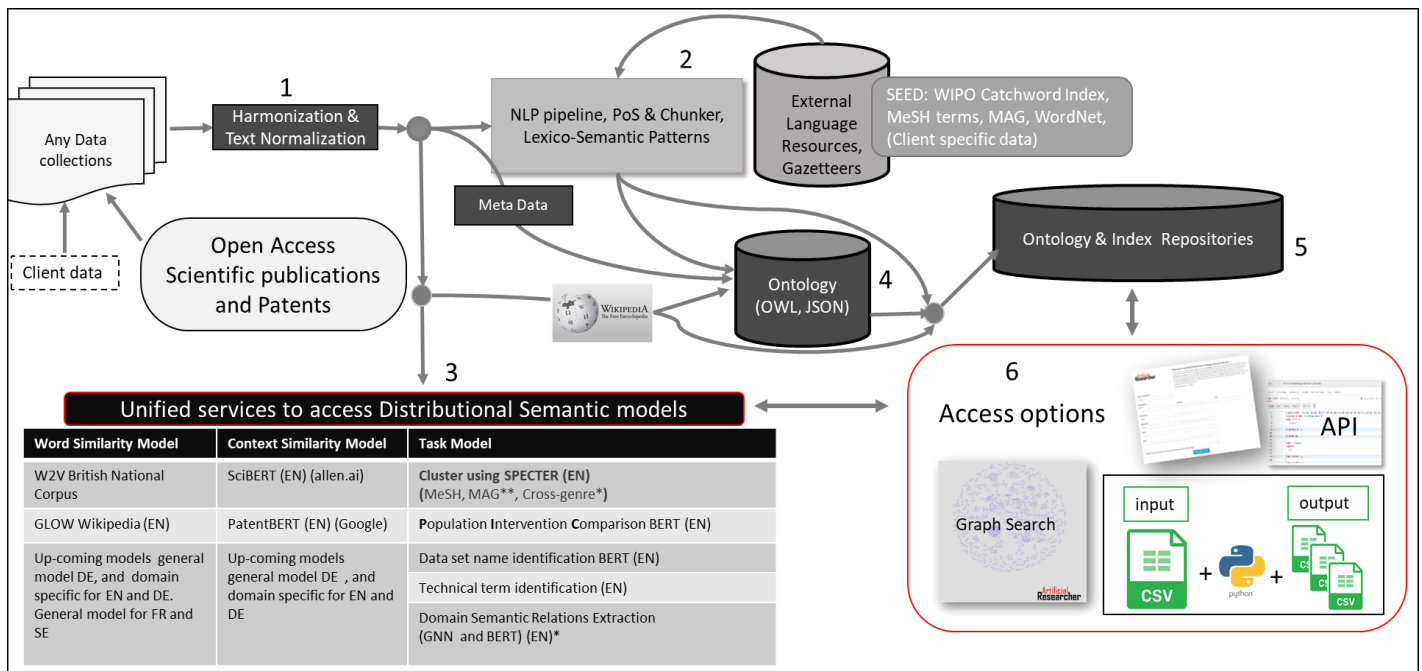
- ⇒ syntagmatic (phrases)
- ⇒ paradigmatic (lexico-semantic relations)

- brake pedal:
- vehicle operating pedal
- conventional hydraulic brake system
- pedal devices
- position brake actuating member
- brake actuating member
- hydraulically-assisted rack pinion steering gear
- brake operating member
- conventional braking system
- pair pedals

- accelerator pedal (antonym)
- case pedal device
- pedal device

ARTIFICIAL RESEARCHER DATA PIPELINE SOLUTION

The Artificial Researcher services and software, aim to provide a sustainable holistic approach to AI-based text mining solutions for industry as well as academy. In the AR Data Pipeline, documents are processed through several text analysis modules: harmonization (1), NLP and relation identification (2), labelling (3), to thereafter be packaged as enhanced XML/JSON (4), software AR Ontology Services and AR Passage Retrieval (5). We also provide end user applications (6) rest APIs, desktop script, UX Graph Search Service and Passage Retrieval.



Data Requirement	Process Time	Deliverables
Any machine-readable format (e.g. PDF or XML, JSON)	1 node process approximately 1000 document per seconds	What? Package Markup data
200K documents gives ca 100K semantic candidate relations	10 nodes process up 1M documents between 10 to 15 days depending on the length of the documents	How? Software package of AR Ontology Services and AR Passage Retrieval
Task Relation Extraction (training or we create training data)		Where? SaaS and on-premises

VERIFIED USE CASES WITH OUR TECHNOLOGY

Artificial Researcher in Open Access Intellectual property Industry Cloud solution <ul style="list-style-type: none"> Scientific publication Patent data 	Artificial Researcher in Science Next generation scientific search tool Onsite solution <ul style="list-style-type: none"> Open access Internal resources 	 iFAIR Identifying datasets in scientific publications (developer) AR-ONTO-COVID A Knowledge-Based Resource for Covid-19 (developer)
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Developing search technology is a skill, but designing search technology is an art.
(Andersson 2021)

CONTACT OUR CEO LINDA ANDERSSON

Linda.Andersson@artificialresearcher.com

Find us on

www.artificialresearcher.com

www.ml4patents.com

<https://www.linkedin.com/company/artificialresearcher/>

Artificial Researcher IT GmbH

Taubstummengasse 11 (i2c),

1040, Wien Austria

Artificial Researcher