Harnessing Scientific AI for Knowledge Discovery: Open Research Knowledge Graph (ORKG)

Allard Oelen

TIB Leibniz Information Centre for Science and Technology, Hannover, Germany

Co-funded by the Horizon 2020 programme
of the European Union

About me

Allard Oelen

- Post-doc at TIB Hannover, Germany
- Frontend lead ORKG Team

Main research topics

- Human-Computer Interaction (**HCI**)
- UIs for **AI** and Knowledge Graphs (**KGs**)
- Scholarly knowledge management

About me

Allard Oelen

- Post-doc at TIB Hannover, Germany
- Frontend lead ORKG Team

Main research topics

- Human-Computer Interaction (**HCI**)
- UIs for **AI** and Knowledge Graphs (**KGs**)
- Scholarly knowledge management

Programming languages and tools:

- **TypeScript**
- **React**
- **Next.js**
- **Tailwind**
- **Python**
- **Backend-as-a-Service**
- **Figma**

Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)

- July 26 (tomorrow): ORKG Ask

Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)
  - Introduction
    - Current issues in scholarly communication
    - ORKG as scholarly knowledge graph
  - Content types
  - UIs for Human-AI collaboration

- July 26 (tomorrow): ORKG Ask

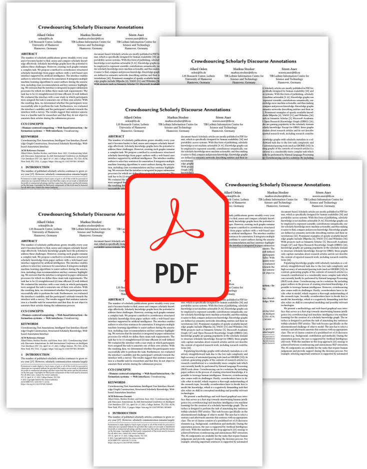
Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)
    - Introduction
        - **Current issues in scholarly communication**
        - ORKG as scholarly knowledge graph
    - Content types
    - UIs for Human-AI collaboration

- July 26 (tomorrow): ORKG Ask

About scholarly communication

- Scholarly knowledge is communicated in **narrative document-based** forms
- Lacking **machine-actionability**: cumbersome for machines to parse the content



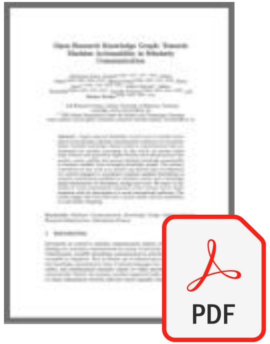
*Information need: **"Author name disambiguation"** approaches leveraring **"graph embeddings"***

About scholarly communication

Scholarly papers have not changed much over time, still the **document-based** format

1620

1883



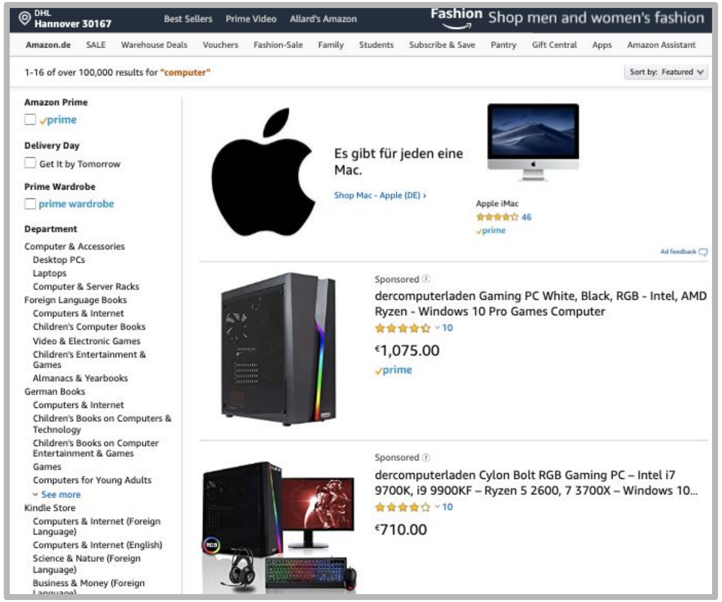
2024

Other domains

Other domains

Other domains

Other domains



Metadata

Well... some things changed

Microsoft Academic Graph (MAG),
Crossref,
Wikidata, WikiCite,
Researchgate,
Semantic Scholar
etc.

Metadata

Well… some things changed

Microsoft Academic Graph (MAG),
Crossref,
Wikidata, WikiCite,
Researchgate,
Semantic Scholar
etc.

== "Metadata"

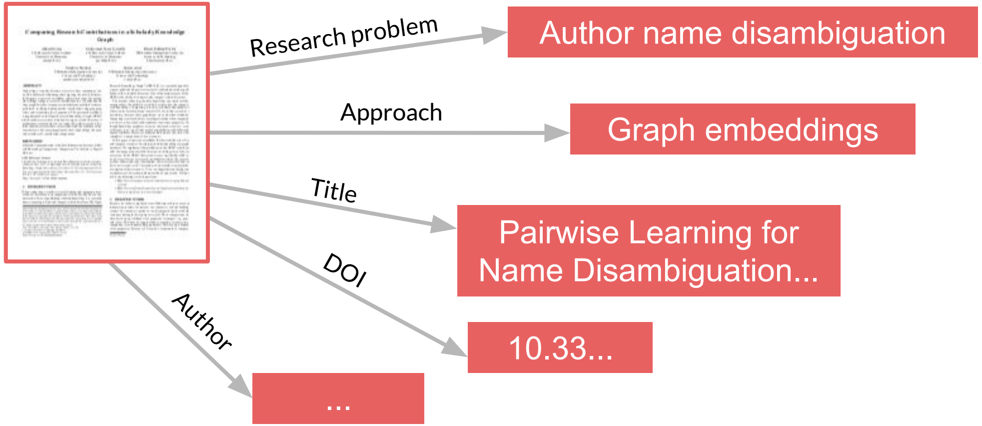
Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)
  - Introduction
    - Current issues in scholarly communication
    - **ORKG as scholarly knowledge graph**
  - Content types
  - UIs for Human-AI collaboration

- July 26 (tomorrow): ORKG Ask

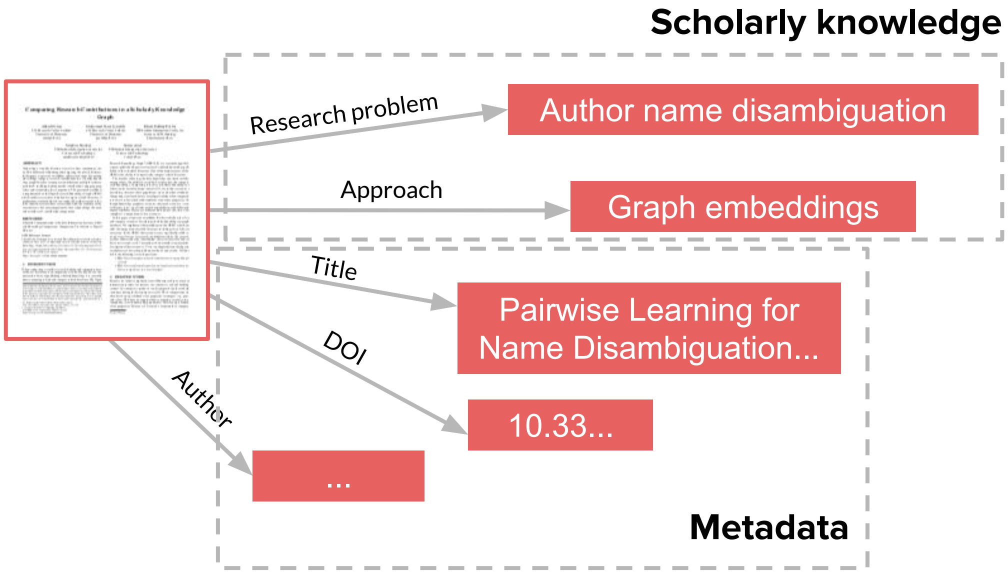
Scholarly knowledge graphs

If **knowledge graphs** are used to represent scholarly instead, retrieving information becomes more effective



Scholarly knowledge graphs

If **knowledge graphs** are used to represent scholarly instead, retrieving information becomes more effective



Scholarly knowledge graphs

If **knowled** ... olarly instead,
retrieving i ...

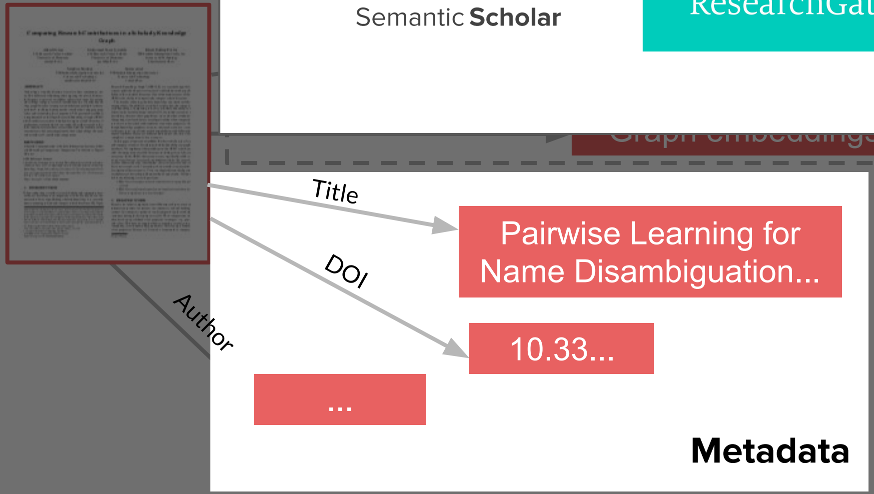
Plenty of existing initiatives

Google Scholar

Microsoft Academic

Semantic Scholar

ResearchGate

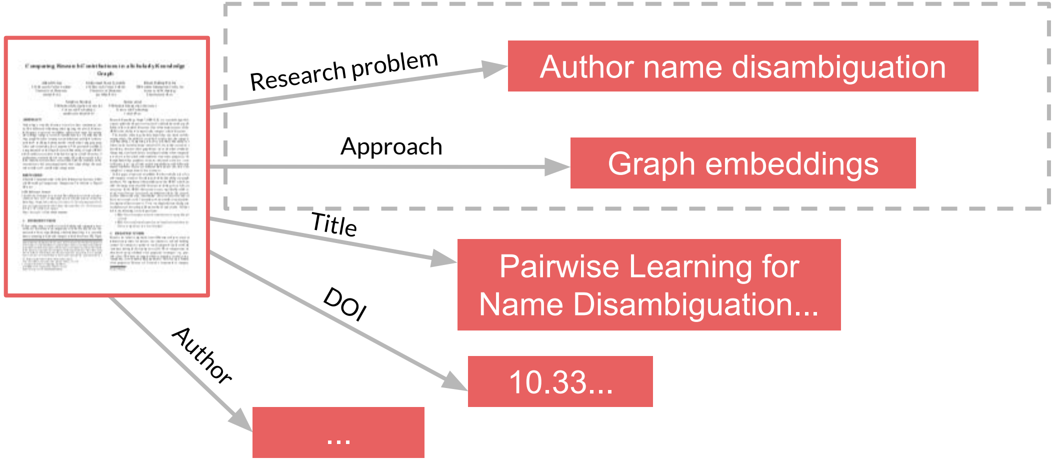


Scholarly knowledge graphs

Jaradeh, Mohamad Yaser, et al. "Open research knowledge graph: next generation infrastructure for semantic scholarly knowledge." *Proceedings of the 10th international conference on knowledge capture*. 2019.

If **knowledge graphs** are used to represent scholarly instead, retrieving information becomes more effective

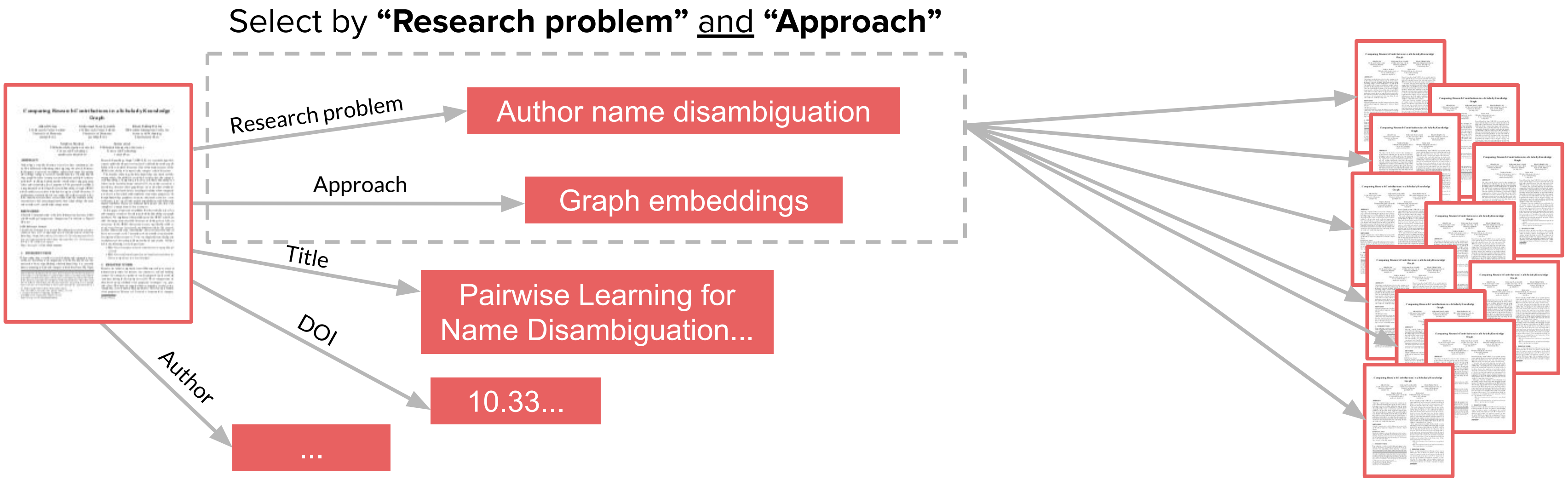
Select by **"Research problem"** and **"Approach"**



Scholarly knowledge graphs

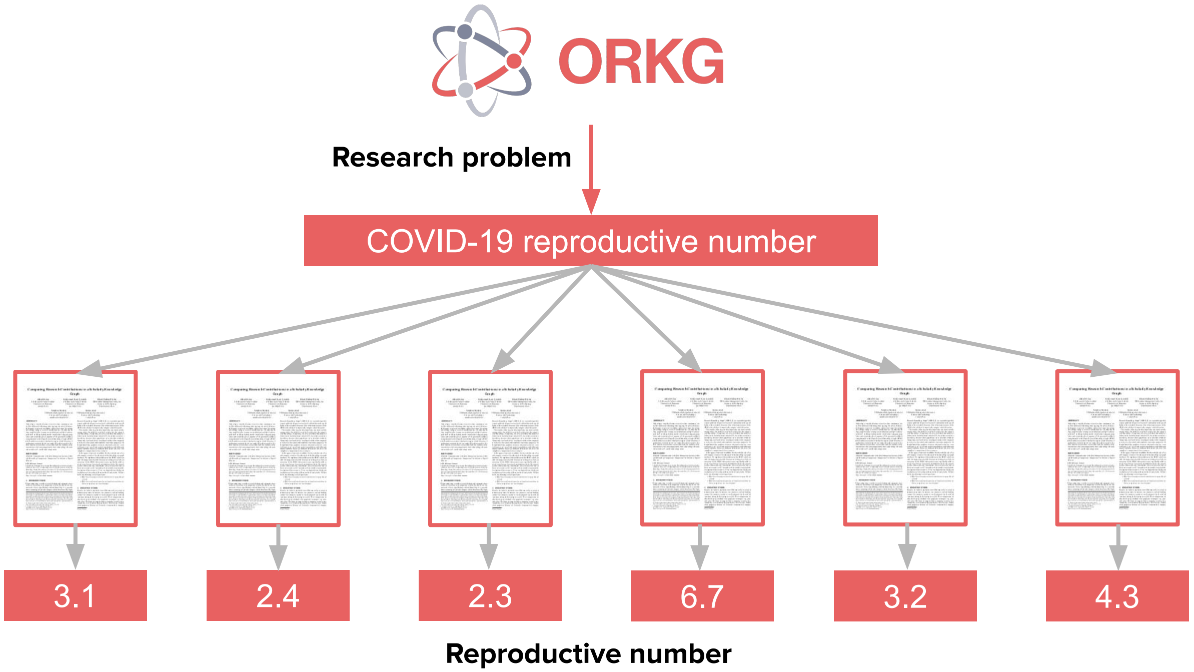
If **knowledge graphs** are used to represent scholarly instead, retrieving information becomes more effective

ORKG
www.orkg.org



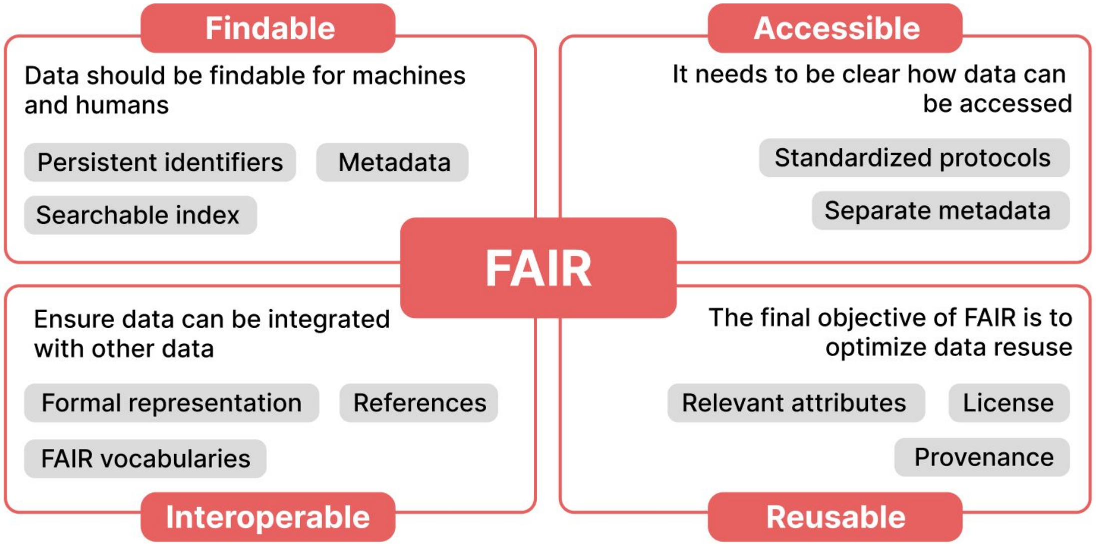
Scholarly knowledge graphs

If such a scholarly knowledge graph exists, the use cases are **virtually limitless**!



FAIR scholarly knowledge

In the end, the knowledge should become Findable, Accessible, Interoperable and Reusable (FAIR)



Wilkinson, Mark D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3.1 (2016): 1-9.

Okay, we need a knowledge graph

But… how?

Knowledge transformation

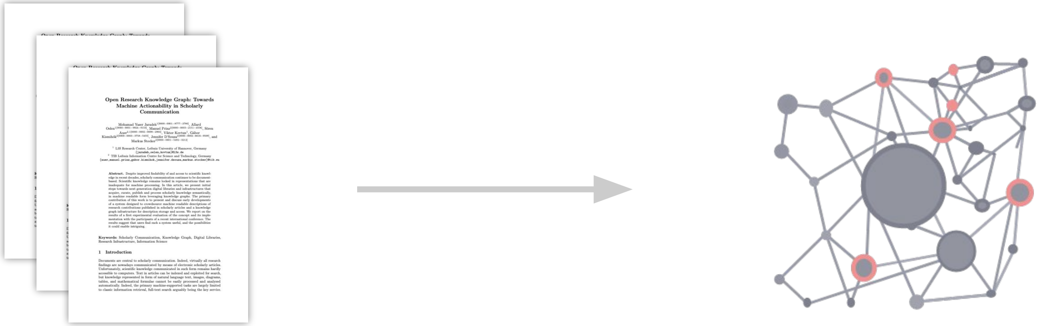
To create a scholarly knowledge graph, a **transformation** from unstructured to structured knowledge should happen

Unstructured knowledge

Structured knowledge

Knowledge transformation

To create a scholarly knowledge graph, a **transformation** from unstructured to structured knowledge should happen



Unstructured knowledge **Structured knowledge**

Can we use AI for the transformation process?

Knowledge transformation

- NLP techniques are **not sufficiently accurate** to perform this task autonomously

Can we use AI for the transformation process?

74%	x	**84%**	x	**78%**	=	**48%**		**Error propagation**

- But we can **intertwine machine intelligence with human intelligence** to get a synergy ➔ the best of both worlds!

Unstructured to structured knowledge

Automatic transformation

AI

- \+ Scales well
- – Not accurate

Manual transformation

Crowdsourcing

- – Does not scale well
- \+ Accurate

Unstructured to structured knowledge

Automatic transformation

Natural Language Processing (NLP)

+ Scales well
– Not accurate

Manual transformation

Crowdsourcing

– Does not scale well
+ Accurate

Intertwining artificial intelligence with human intelligence: best of both worlds

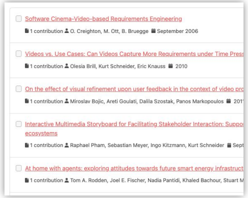
Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)
    - Introduction
        - Current issues in scholarly communication
        - ORKG as scholarly knowledge graph
    - **Content types**
    - UIs for Human-AI collaboration

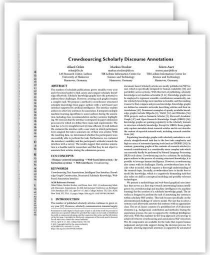
- July 26 (tomorrow): ORKG Ask

ORKG content types

Lists

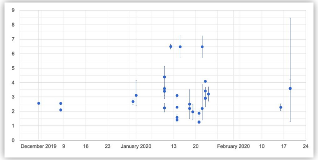


Papers

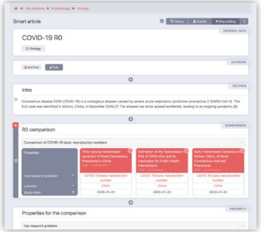


Comparisons

Visualizations

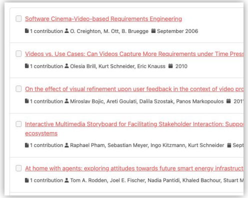


Reviews

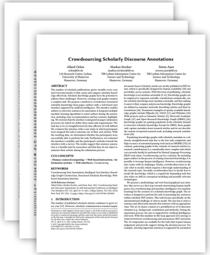


ORKG content types

Lists

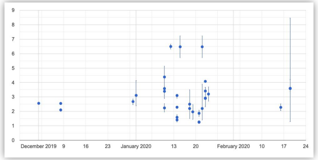


Papers

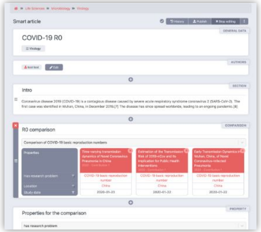


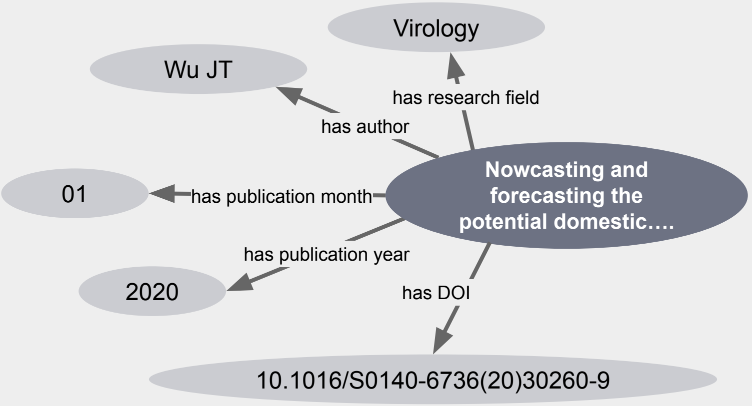
Comparisons

Visualizations



Reviews



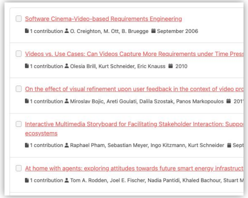


ORKG lists

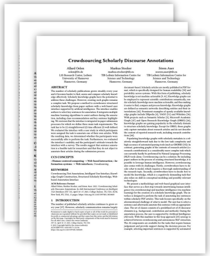
- Form the starting point for structured knowledge descriptions
- Mainly focus on metadata organization
- Consists of papers, software, or datasets

ORKG content types

Lists

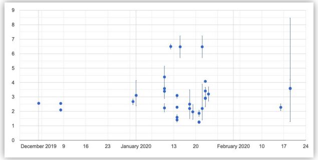


Papers

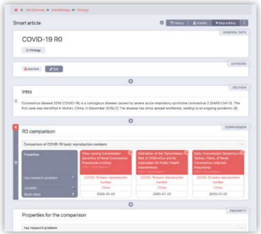


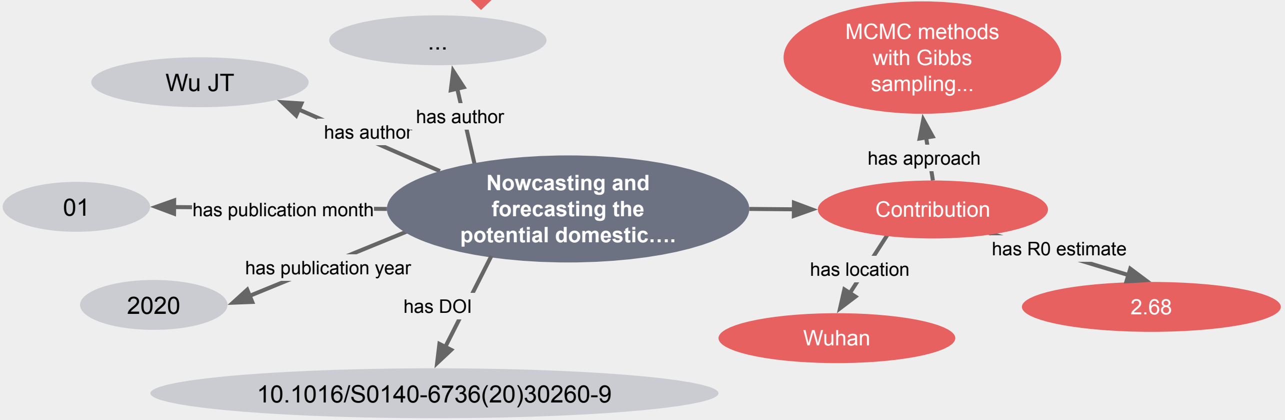
Comparisons

Visualizations



Reviews

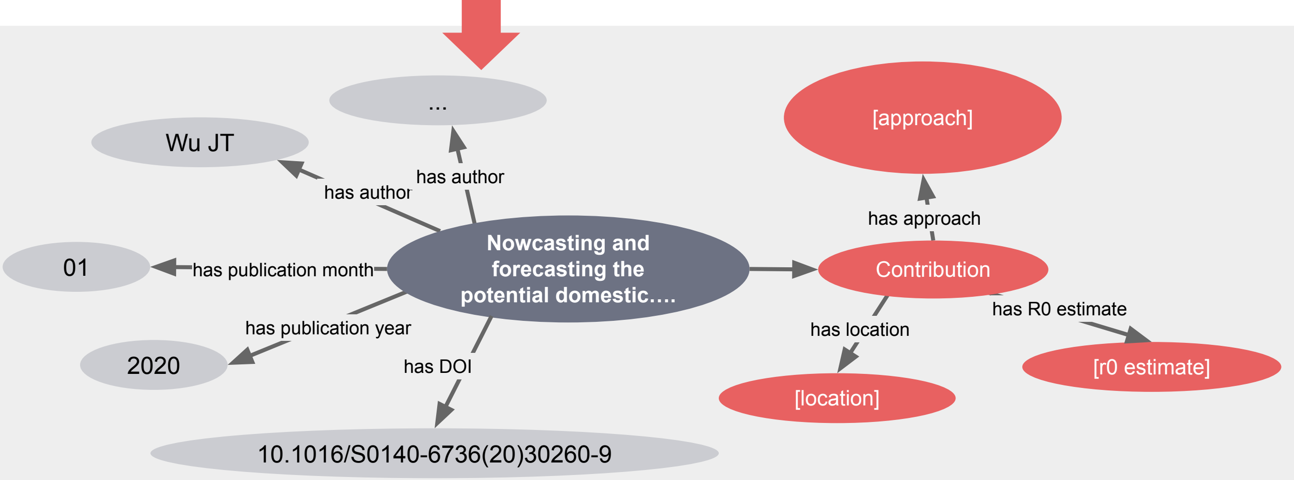




ORKG papers

- Structured scholarly knowledge
- Simple or more complex graph structures
- Possibly using templates

ORKG content types



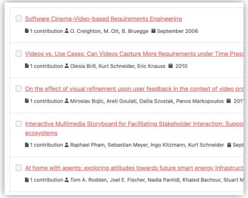
ORKG papers

- Structured scholarly knowledge
- Simple or more complex graph structures
- **Possibly using templates**

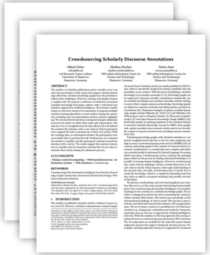
ORKG content types

Oelen, Allard, et al. "Generate FAIR literature surveys with scholarly knowledge graphs." Proceedings of the ACM/IEEE joint conference on digital libraries in 2020. 2020.

Lists

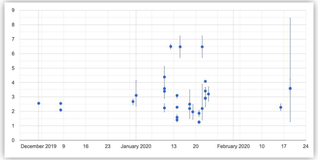


Papers

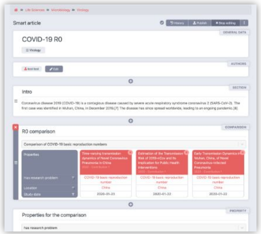


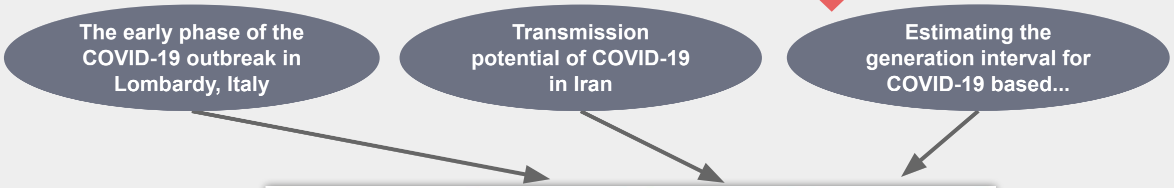
Comparisons

Visualizations



Reviews





Properties	The early phase of the COVID-19 outbreak in Lombardy, Italy Contribution 1 - 2020	Transmission potential of COVID-19 in Iran Contribution 1 - 2020	Estimating the generation interval for COVID-19 based on symptom onset data Contribution 1 - 2020
Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)
Location	Lombardy, Italy	Iran	Singapore
Has value	3.1	3.6	1.27
Lower confidence limit	2.9	3.4	1.19
Upper confidence limit	3.2	4.2	1.36

ORKG comparisons

- Tabular view of ORKG papers addressing the same research problems
- One of the key content types in the ORKG

ORKG content types

Oelen, Allard, et al. "Generate FAIR literature surveys with scholarly knowledge graphs." Proceedings of the ACM/IEEE joint conference on digital libraries in 2020. 2020.

Lists | Papers | **Comparisons** | Visualizations | Reviews

Visualize in **tabular form**, with focus on customizability and reusability

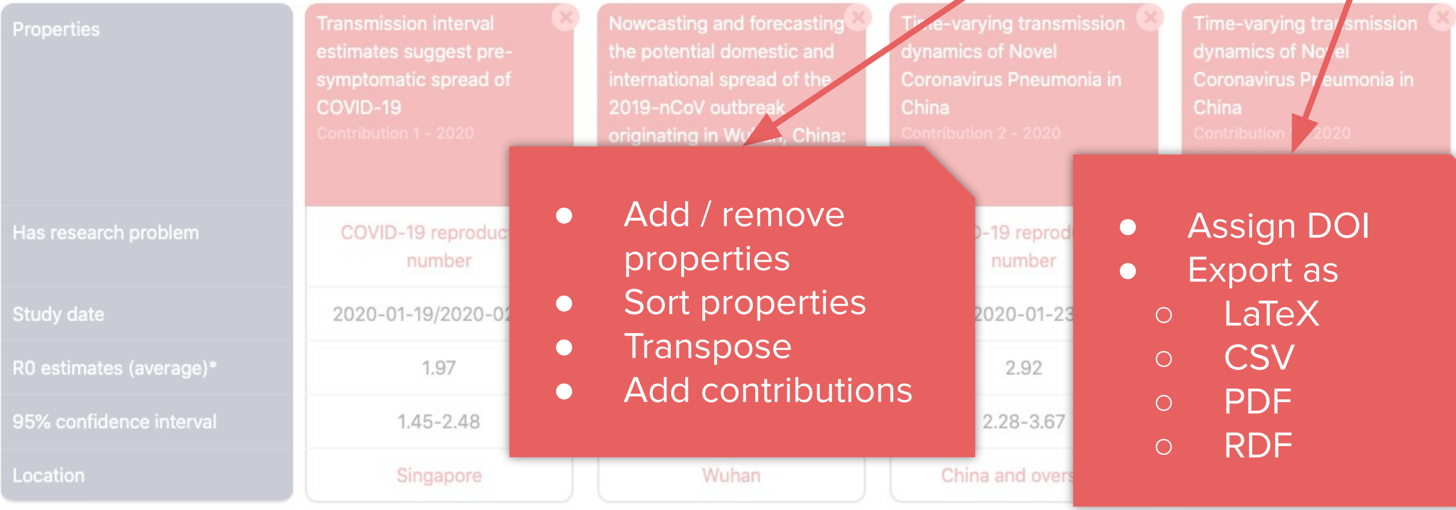
Properties	Transmission interval estimates suggest pre-symptomatic spread of COVID-19 Contribution 1 - 2020	Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study Contribution 1 - 2020	Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China Contribution 2 - 2020	Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China Contribution 1 - 2020
Has research problem	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number	COVID-19 reproductive number
Study date	2020-01-19/2020-02-26	2019-12-31/2020-01-28	2020-01-23	2020-01-23
R0 estimates (average)*	1.97	2.68	2.92	2.9
95% confidence interval	1.45-2.48	2.47-2.86	2.28-3.67	2.32-3.63
Location	Singapore	Wuhan	China and overseas	China and overseas

ORKG content types

Oelen, Allard, et al. "Generate FAIR literature surveys with scholarly knowledge graphs." Proceedings of the ACM/IEEE joint conference on digital libraries in 2020. 2020.

Lists | **Papers** | **Comparisons** | **Visualizations** | **Reviews**

Visualize in **tabular form**, with focus on customizability and reusability

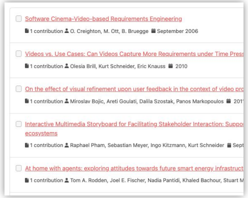


- Add / remove properties
- Sort properties
- Transpose
- Add contributions

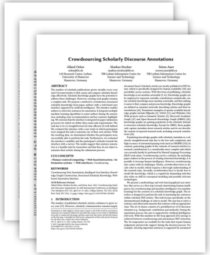
- Assign DOI
- Export as
  - LaTeX
  - CSV
  - PDF
  - RDF

ORKG content types

Lists

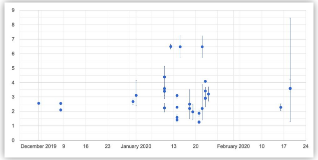


Papers

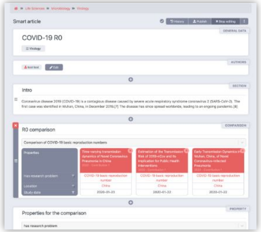


Comparisons

Visualizations



Reviews



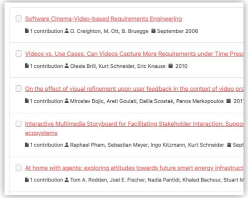
ORKG visualizations

- Visualize numeric data from ORKG comparisons
- Various visualization types available

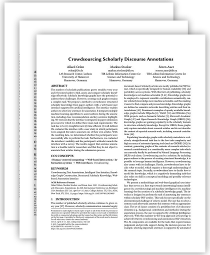
ORKG content types

Oelen, Allard, Markus Stocker, and Sören Auer. "SmartReviews: towards human-and machine-actionable reviews." *Linking Theory and Practice of Digital Libraries: 25th International Conference on Theory and Practice of Digital Libraries, TPDL 2021,*

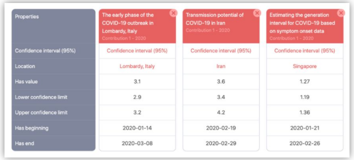
Lists



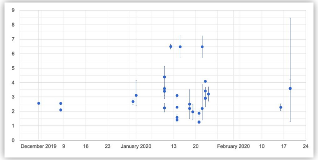
Papers



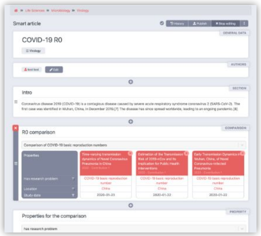
Comparisons



Visualizations

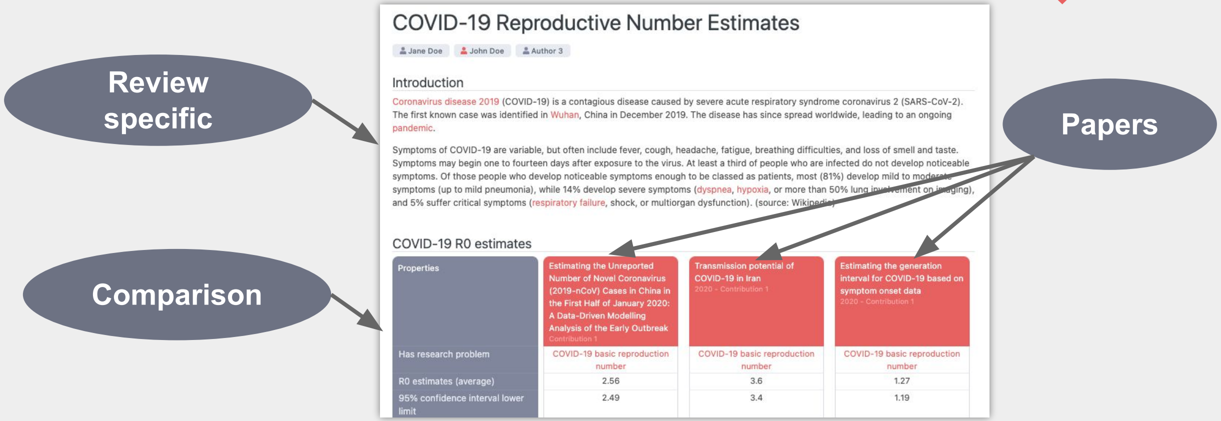


Reviews



ORKG reviews

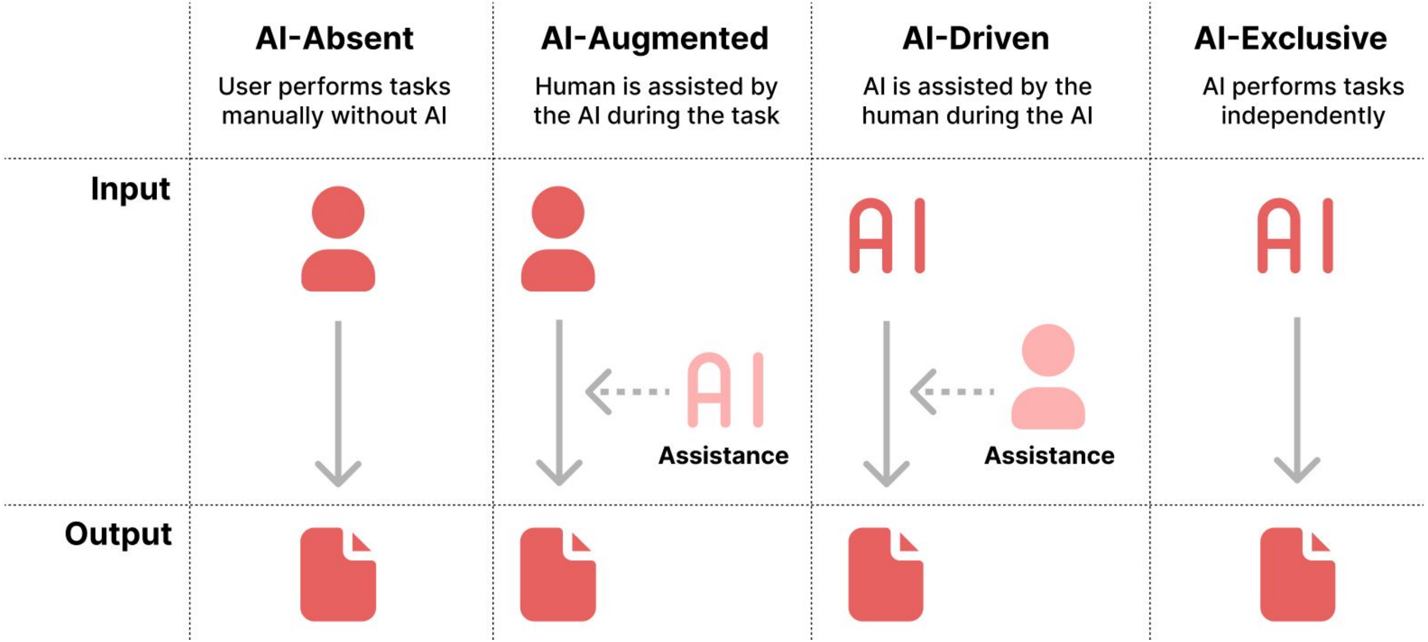
- A living-document reviewing related literature
- Consists of all previous ORKG content types



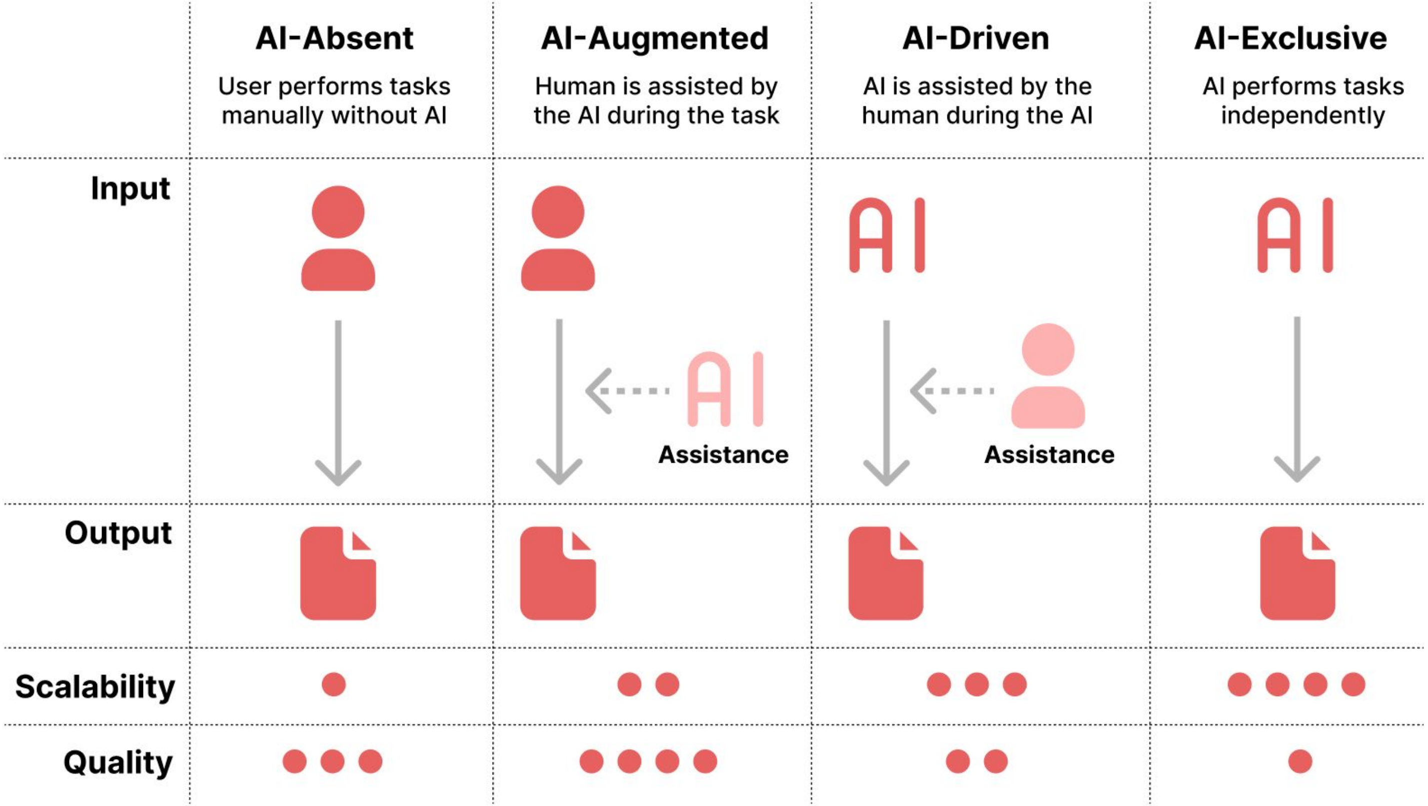
Outline

- July 25 (today): Open Research Knowledge Graph (ORKG)
  - Introduction
    - Current issues in scholarly communication
    - ORKG as scholarly knowledge graph
  - Content types
  - **UIs for Human-AI collaboration**
- July 26 (tomorrow): ORKG Ask

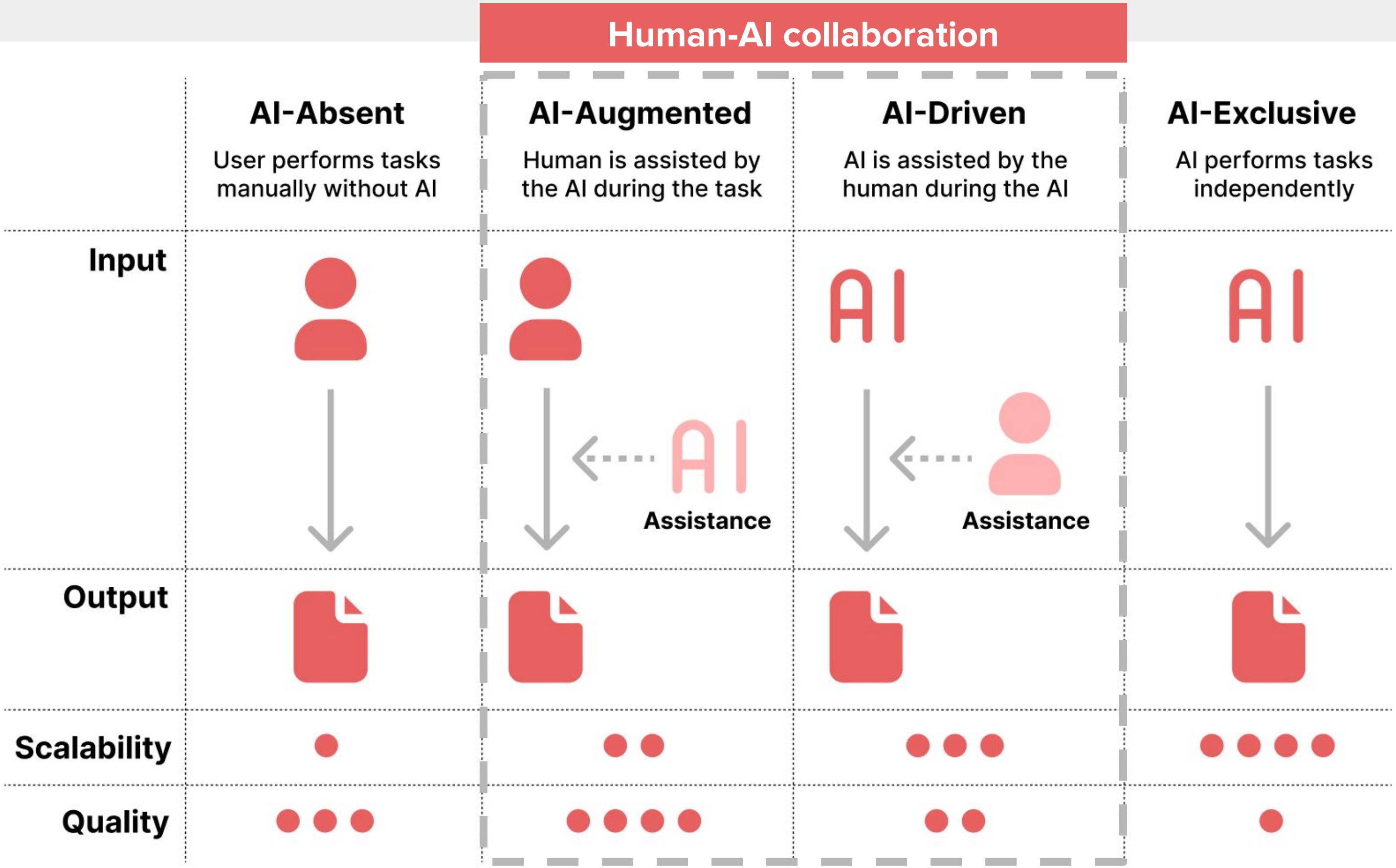
Human-AI collaboration



Human-AI collaboration



Human-AI collaboration



Human-AI collaboration in the ORKG

AI-Augmented

1. Smart suggestions

AI-supported tooltips helping users accomplish their tasks

2. Paper annotator

Annotation of key sentences in scholarly PDF articles

3. Survey extractor

Extract survey tables from existing papers

AI-Driven

4. TinyGenius

Microtasks to validate NLP generated statements

5. ORKG Ask

Tomorrow's topic

Human-AI collaboration in the ORKG

AI-Augmented

1. Smart suggestions

AI-supported tooltips helping users accomplish their tasks

2. Paper annotator

Annotation of key sentences in scholarly PDF articles

3. Survey extractor

Extract survey tables from existing papers

AI-Driven

4. TinyGenius

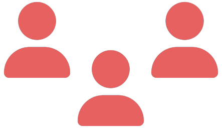
Microtasks to validate NLP generated statements

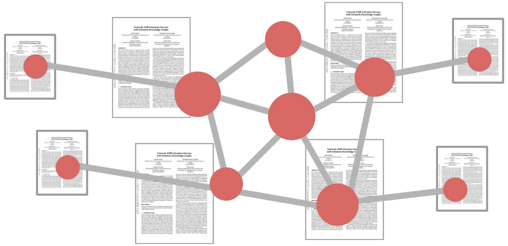
5. ORKG Ask

Tomorrow's topic

Transform unstructured into structured knowledge

Crowdsourcing





Unstructured

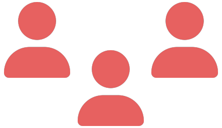
Structured

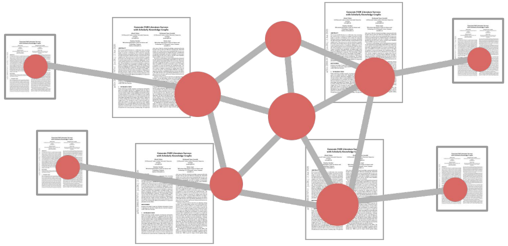
Oelen, Allard, and Sören Auer. "Leveraging Large Language Models for Realizing Truly Intelligent User Interfaces." *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 2024.

Transform unstructured into structured knowledge

Unstructured

Crowdsourcing





LLM support

Structured

Extracted research topics:
Cost-Effective Sensors
Accessible Geoscientific Sensors

Oelen, Allard, and Sören Auer. "Leveraging Large Language Models for Realizing Truly Intelligent User Interfaces." *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 2024.

Guidelines for integrating LLMs into existing UIs

1. Transparency
The integration is clearly distinguishable within the UI

2. Control
The integration is non-intrusive and can be hidden by the UI

3. Usability
The UI integration makes the integration blend seamlessly in the UI

4. Error management
Graceful degradation: the UI does not break in case of errors

5. Feedback and statistics
Users are able to give feedback

6. System performance
The integration has minimal response time, and works within seconds

Based on Nielsen, Jakob. "Enhancing the explanatory power of usability heuristics." Proceedings of the SIGCHI conference on Human Factors in Computing Systems. 1994.

Guidelines for integrating LLMs into existing UIs

Task	Description	Implementation directions
1. Transparency		
1.1. Distinguishable	The system shall be clearly distinguishable in the UI.	Using distinctive color scheme and recognizable icons.
1.2. Suggestions	The system shall be displayed explicitly as suggestive.	Informing users that the suggestions can be wrong or misleading.
1.3. Transparency	The system shall make it clear how suggestions are generated.	Mention the model (e.g., ChatGPT), model input, and prompt.
1.4. Multiple variants*	The system shall provide multiple values when appropriate to stress uncertainty.	Provide a list of different options from which users have to select the desired option.
1.5. Language	The system shall use appropriate language to express uncertainty.	Use words such as might, could, possibly, seems to be, etc.
2. Control		
2.1. Non-intrusive*	The system shall have the option to hide it.	Use collapsible UI components.
2.2. On demand*	The system shall be displayed on demand.	Do not open the suggestions by default.
2.3. Deactivation*	The system shall provide an option to be deactivated.	Provide a setting on user level to hide the suggestions in the entire UI.

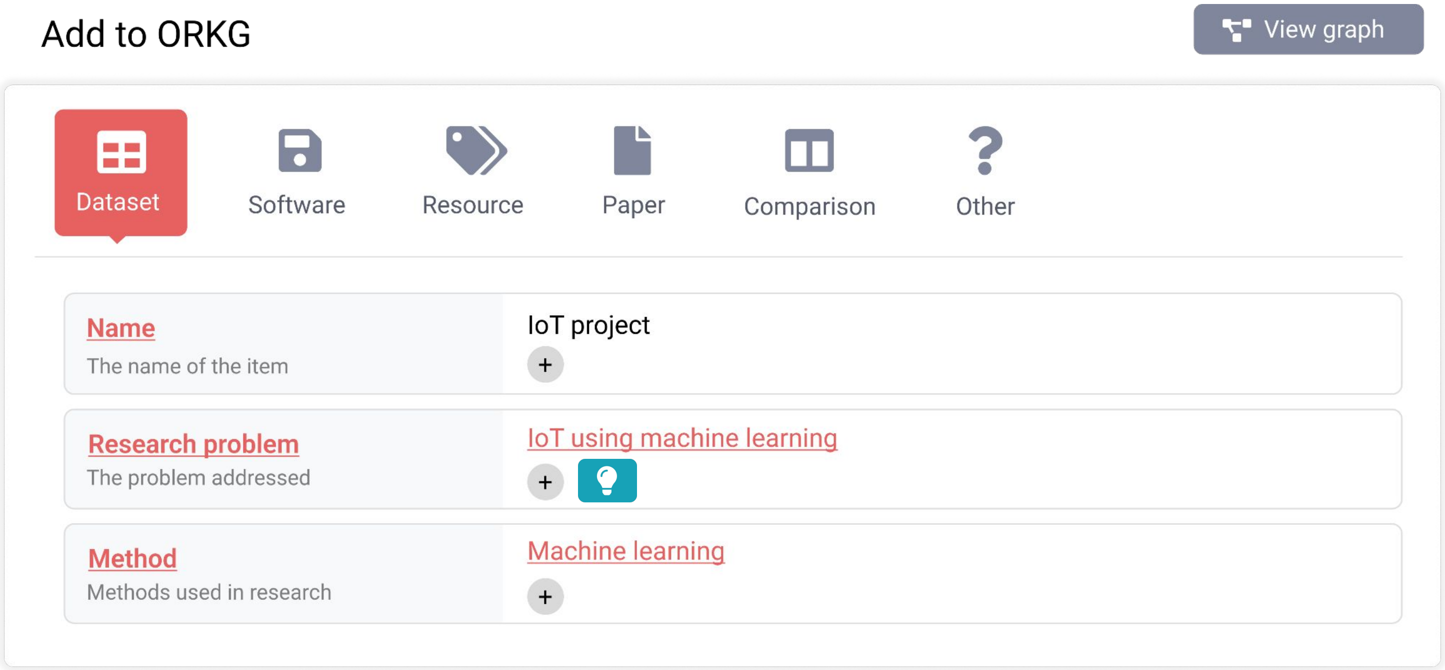
Guidelines for integrating LLMs into existing UIs

3. Usability		
3.1. UI integration	The system shall seamlessly blend into the UI.	Instead of a separate UI, integrate the LLMs into the existing UIs, ensuring the users' attention is focused towards the task.
3.2. Consistent availability*	The system shall be available when expected by users.	Smart Suggestions should be available both when adding and editing data.
3.3. Optional usage	The system shall not be required to fulfill the task.	Users can still perform the task manually.
3.4. Modifiable*	The system shall provide the option to modify suggestions.	After selecting a recommended value, allow the possibility to edit the value.
3.5. Regenerating*	The system shall provide an option to regenerate the response.	Using a reload button to get additional LLM responses.
4. Error Management		
4.1. Graceful degradation	The system shall not break the UI when it is failing.	In case the LLM is not available or not returning the response as expected, ensure the UI remains operable and do not present them as critical errors.
4.2. Error recovery*	The system shall provide a possibility to recover from errors.	Add a reload button when errors appear and explain how to present errors.
4.3. Error prevention	The system shall minimize the user input to mitigate potential errors.	Prevent errors by built-in prompts with placeholders that contain user input.

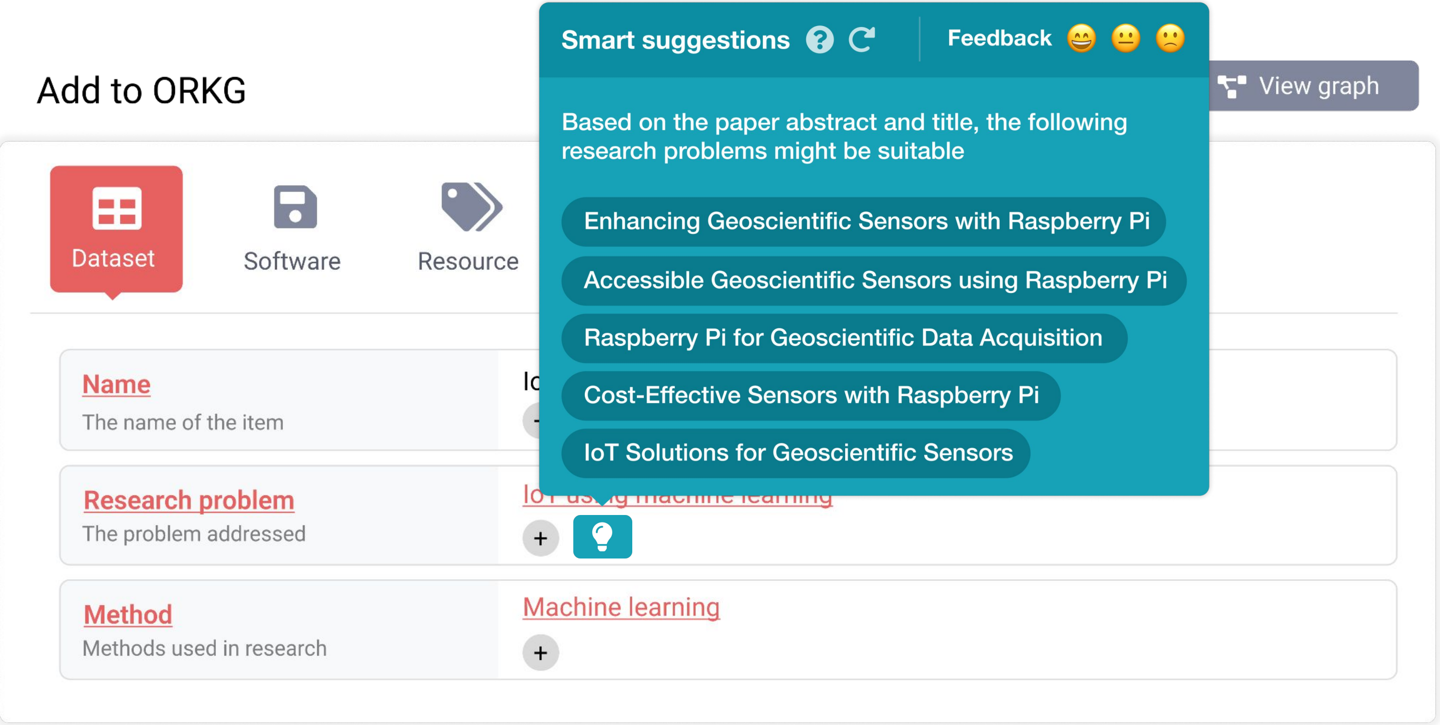
Guidelines for integrating LLMs into existing UIs

5. Feedback and Statistics		
5.1. High-level feedback*	The system shall facilitate the process of providing feedback with minimal effort.	A three-level scale: positive, neutral, negative to determine whether tasks are performing well, need to be improved, or need to be removed.
5.2. Detailed feedback*	The system shall facilitate the process of providing more detailed feedback.	Standardized answers to indicate usefulness and correctness. Optionally provide additional input.
5.3. Usage statistics	The system shall be recording usage statistics without explicit efforts from users.	Record clicks when LLM suggestions are being used.
6. System Performance		
6.1. Response time	The system shall respond within seconds.	Ensure prompts and answers are short to ensure users can access the LLM tool to get quick access.
6.2. Minimize requests	The system shall debounce function calls to minimize requests for environmental and monetary reasons.	Activate LLM support on demand when a button is clicked.
6.3. Prevent misuse	The system shall use a backend service to generate the prompts being sent to the LLM.	Prompts are stored in the service and the LLM interface is not exposed to the client, but made available through middleware.

Smart Suggestions implementation

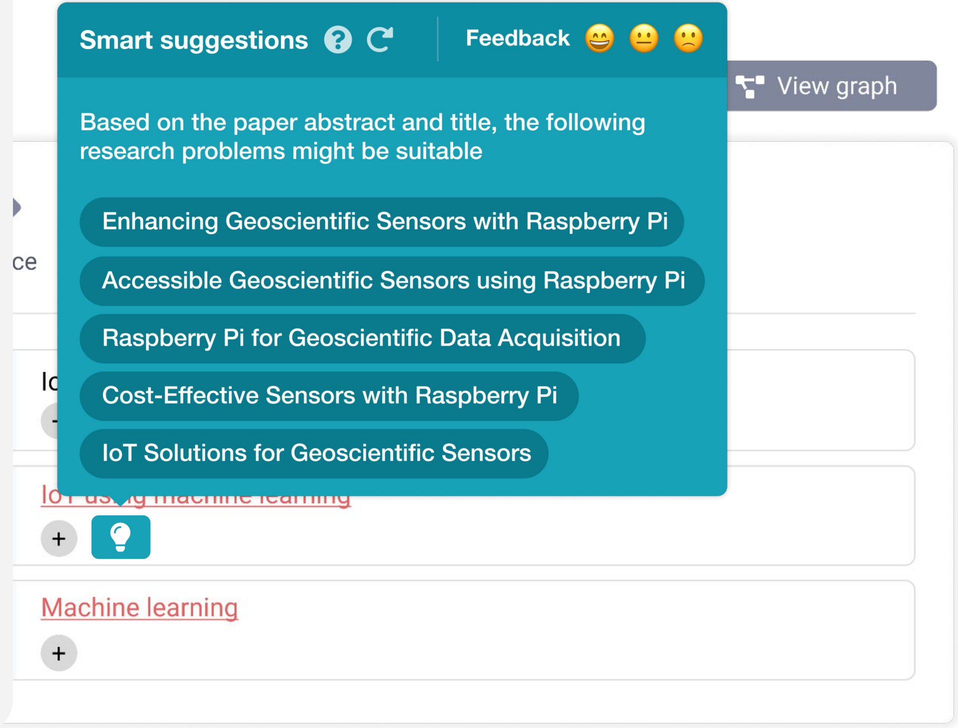


Smart Suggestions implementation



Smart Suggestions implementation

- Implemented for **6 use cases** in the UI
- Recognizable icon
- Distinctive color palette



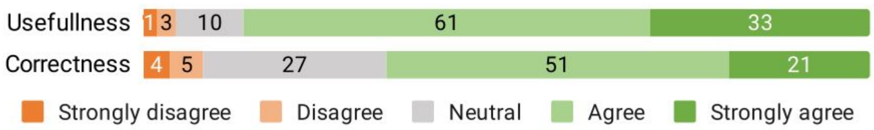
Smart suggestions prompts

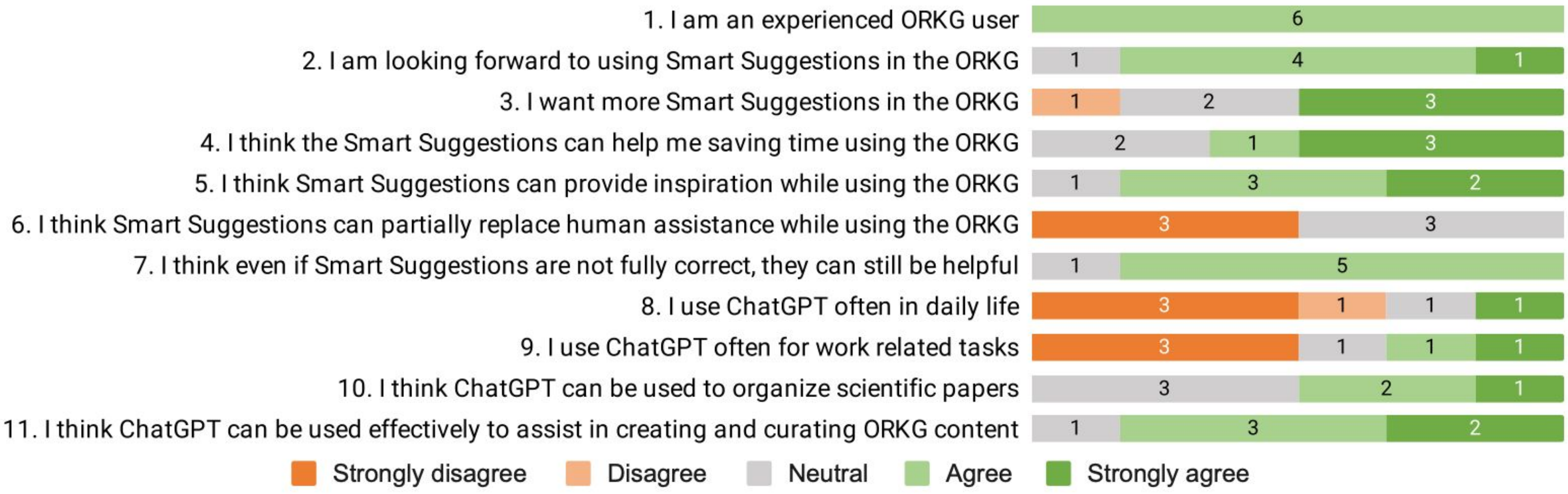
Use case	Description	Prompt
1. Related Predicates	When making statements in an RDF knowledge graph, a subject, predicate, and object are required. The object can either be a resource (a piece of information with an identifier that can be linked to) or a literal (information that cannot be linked to, such as a string, numbers, dates, etc.). This type of Smart Suggestion recommends predicates to users based on a set of predicates coming from the existing paper description.	**System prompt:** You are an assistant for building a knowledge graph for science. Your task is to recommend additional related predicates based on the set of existing predicates. Recommend a list maximum 5 additional predicates. **User prompt:** The existing predicates are: [list of predicates]
2. Related Objects	This relates to the previous task but aims to find a set of related objects instead. Since it requires a prompt that provides the LLM with the necessary context, this is only activated for a selected set of predicates, namely: research problem, method, and approach. Thus, each of these predicates has its own prompt.	**System prompt:** A [research problem] contains a maximum of approximately 4 words to explain the research task or topic of a paper. Provide a list of maximum 5 research problems based on the title and optionally abstract provided by the user. **User prompt:** [paper title] [abstract]

Smart suggestions prompts

Open Feedback		
3. Literal Applicability	In addition to creating resources at the object position of a statement, RDF also allows creating literals, which resemble a piece of textual information that cannot be linked to. Based on our previous experiences with ORKG users, we learned that it can be difficult for users to decide whether an object should be a resource or a literal. This Smart Suggestion helps to determine the most appropriate type when creating an object. It is evaluating if a piece of text should indeed be a literal, or if it is more appropriate as a resource.	**System prompt:** You are an assistant in building a knowledge graph for science. You task is to advice users whether they should use a RDF resource or RDF literal. Based on a user-provided label, advice whether the type should be 'literal' or 'resource'. Literals are generally larger pieces of text and are not reusable, resource are atomic and can be reused. **User prompt:** [label]
4. Decomposable Resources	If resources are represented in an atomic fashion, they contain general information and can thus be reused more easily. This facilitates interconnections which enhances the graph quality. This use case evaluates if a resource label can be decomposed into multiple labels, or if the content is already sufficiently atomic.	**System prompt:** You are an assistant for building a knowledge graph for science. Provide advice on if and how to decompose a provided resource label into separate resources. Only provide feedback is decomposing makes sense. **User prompt:** [label]

Smart suggestions - Preliminary evaluation results





Human-AI collaboration in the ORKG

AI-Augmented

1. Smart suggestions

AI-supported tooltips helping users accomplish their tasks

2. Paper annotator

Annotation of key sentences in scholarly PDF articles

3. Survey extractor

Extract survey tables from existing papers

AI-Driven

4. TinyGenius

Microtasks to validate NLP generated statements

5. ORKG Ask

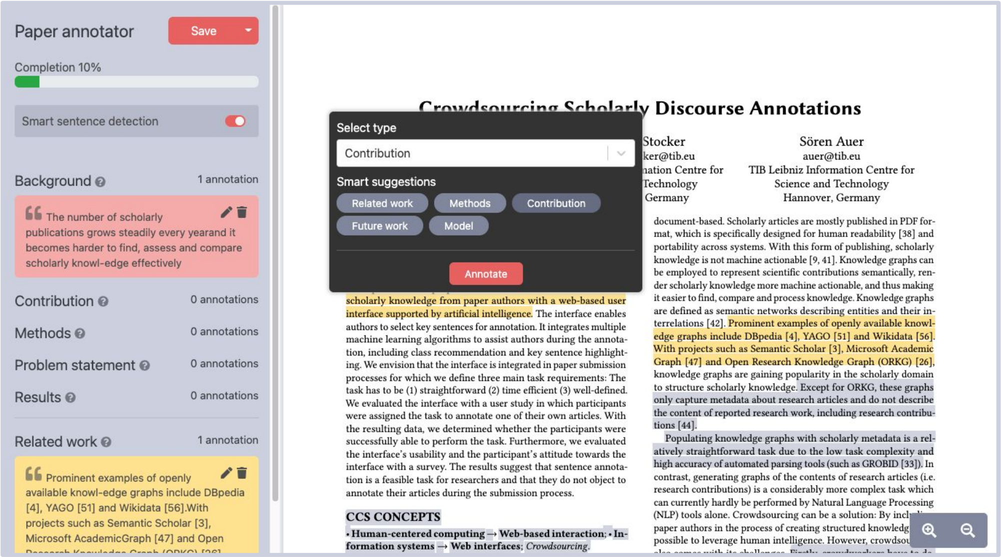
Tomorrow's topic

Paper annotator

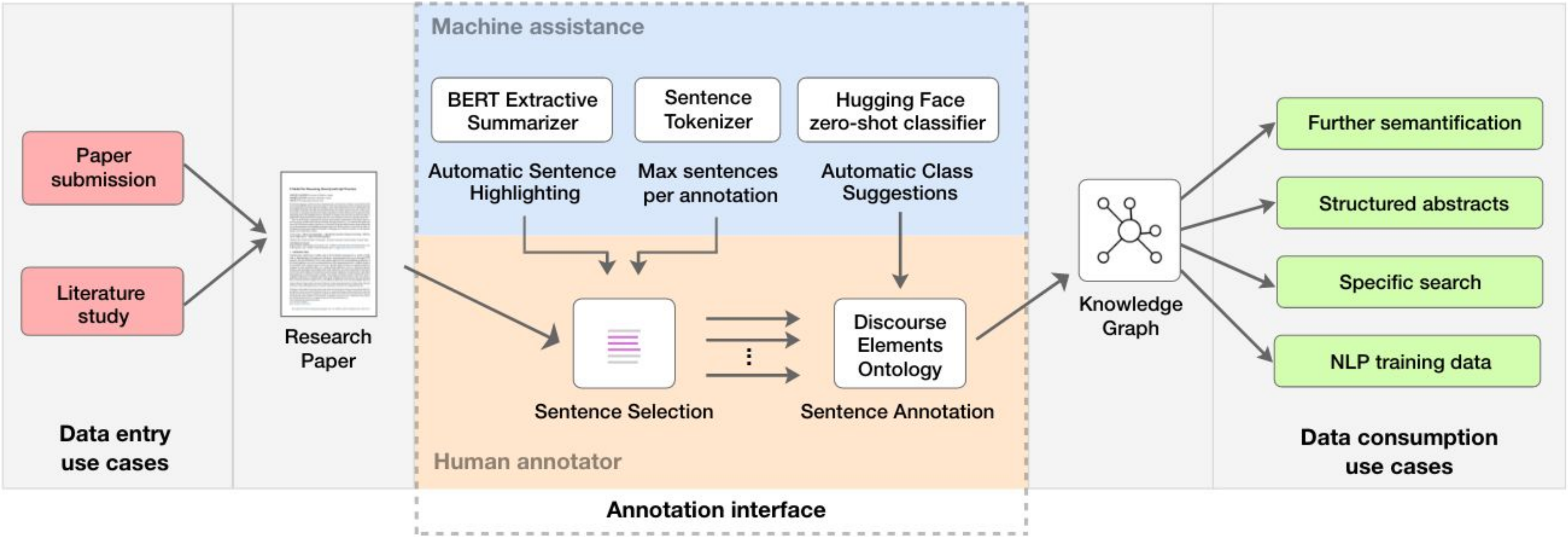
Oelen, Allard, Markus Stocker, and Sören Auer. "Crowdsourcing scholarly discourse annotations." *Proceedings of the 26th International Conference on Intelligent User Interfaces*. 2021.

- Goal: **annotate key sentences** in scholarly articles with discourse classes

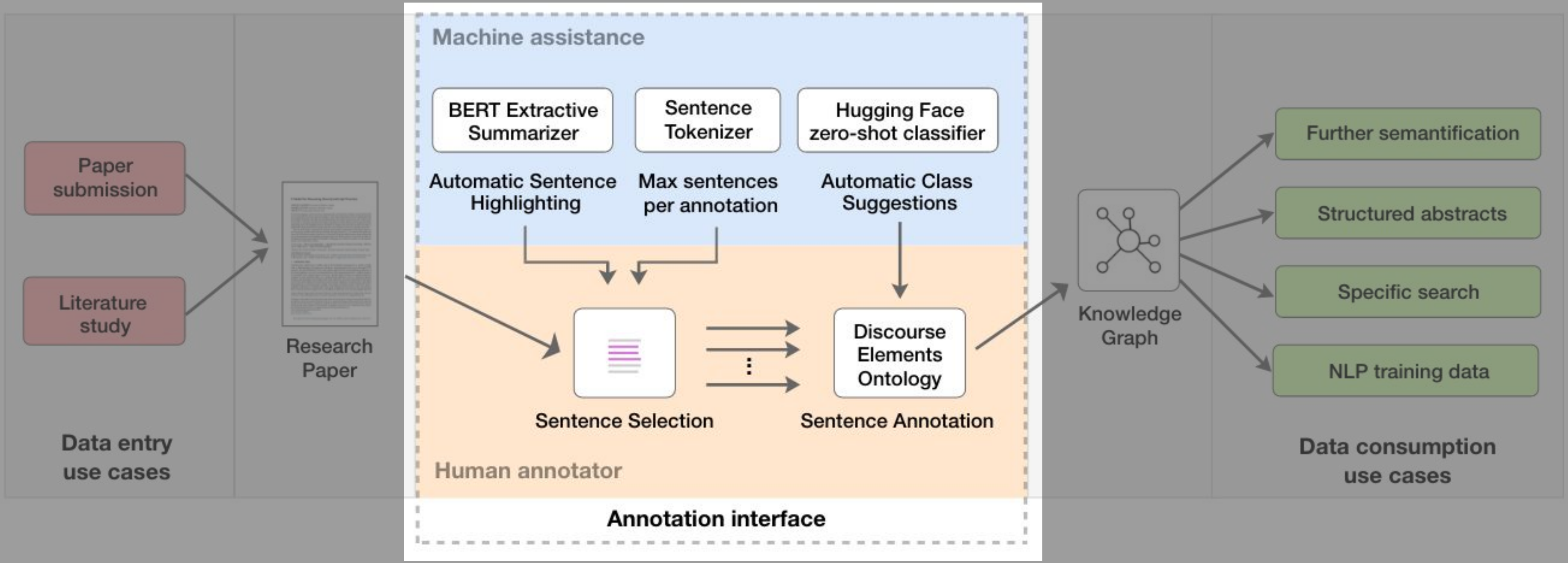
- Two AI-augmented approaches: **sentence highlighting** and **class recommendations**



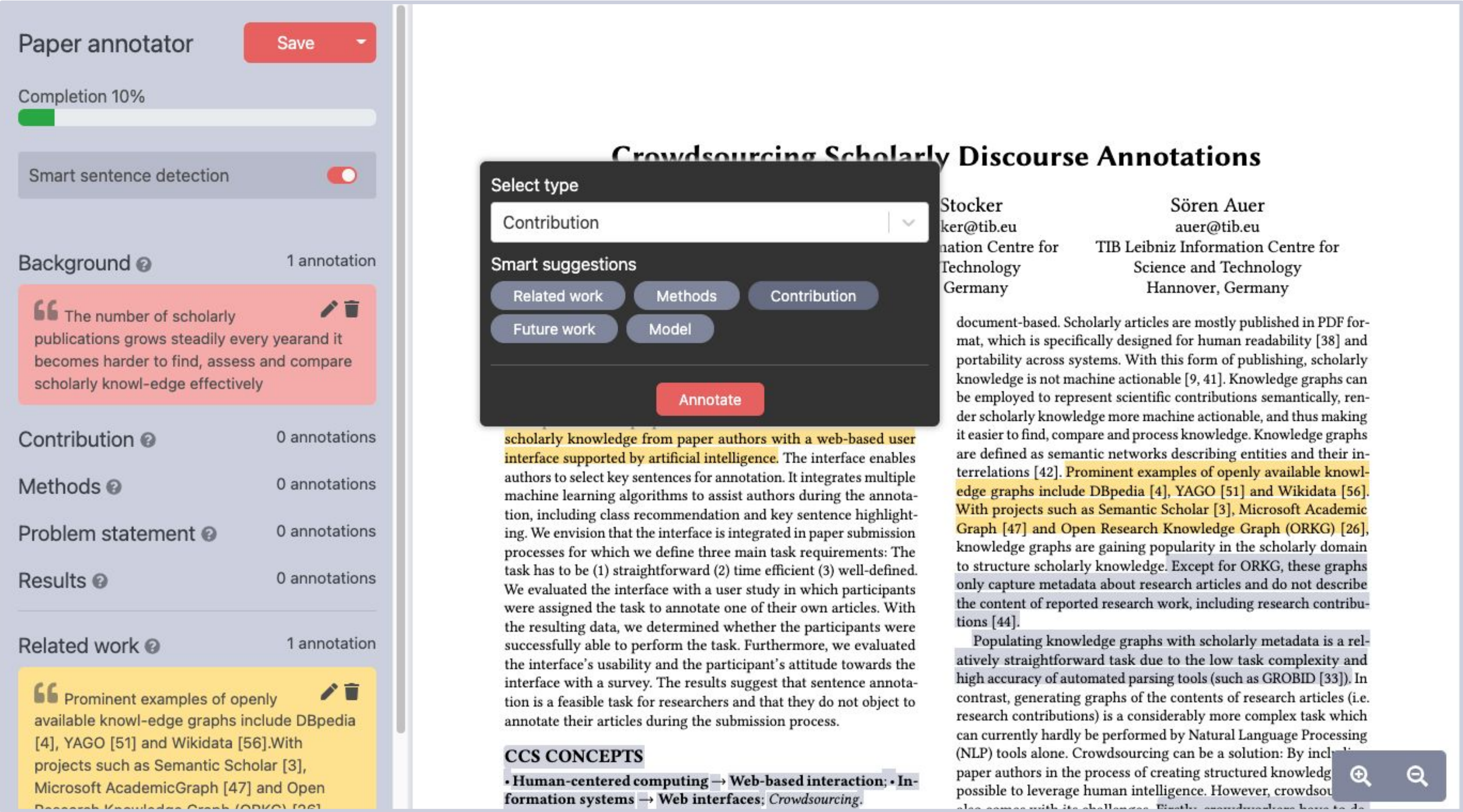
PDF sentence annotation



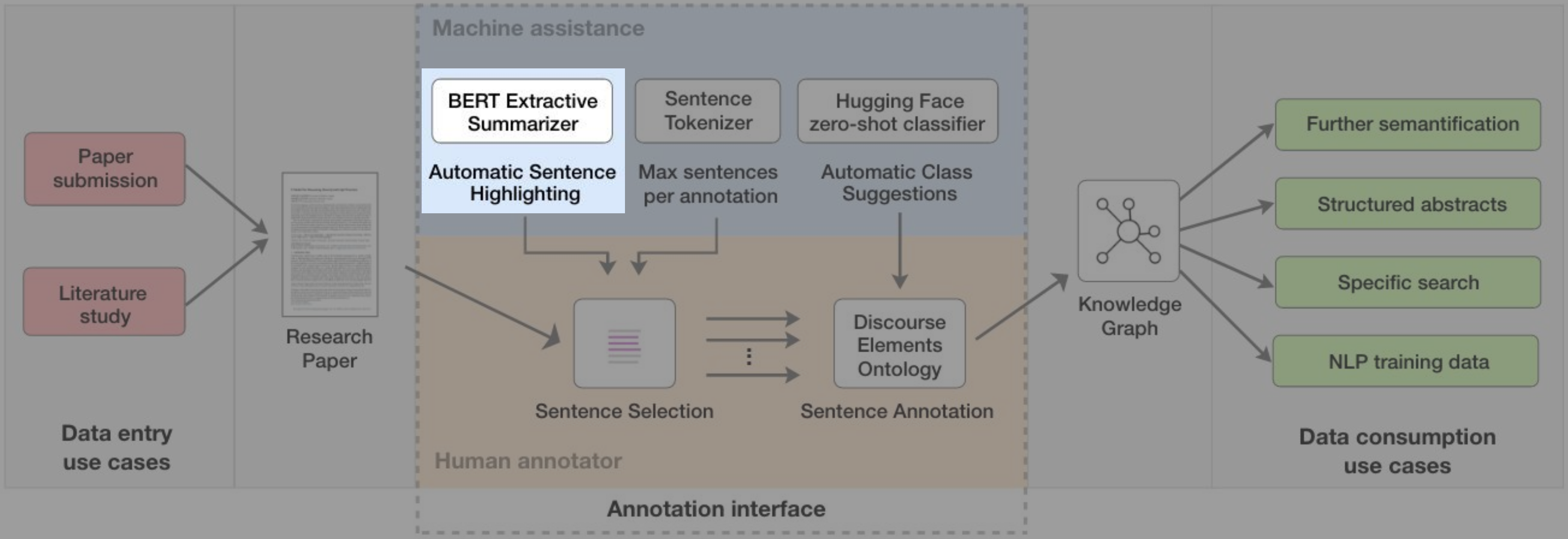
PDF sentence annotation



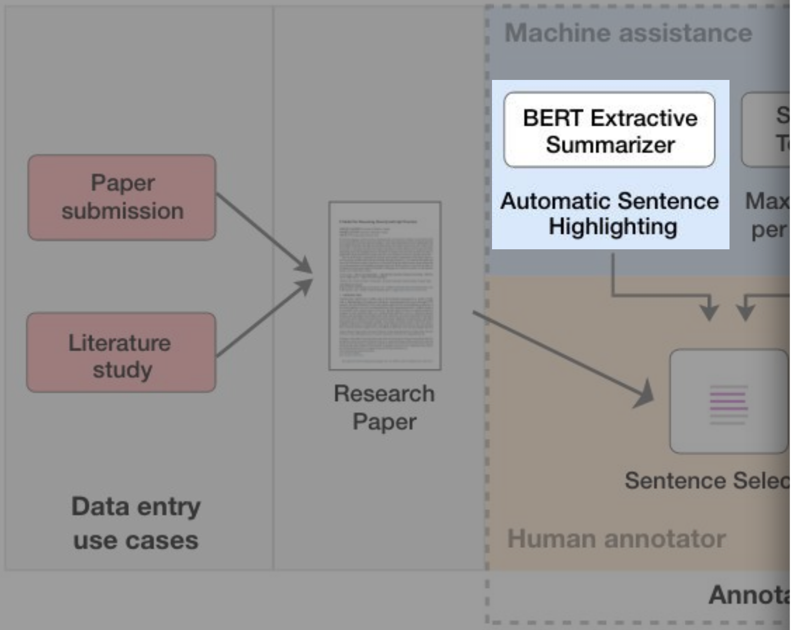
Annotation interface



PDF sentence annotation



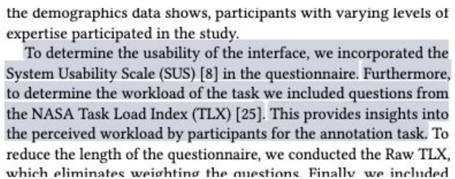
PDF sentence annotation



Automatic sentence highlighting

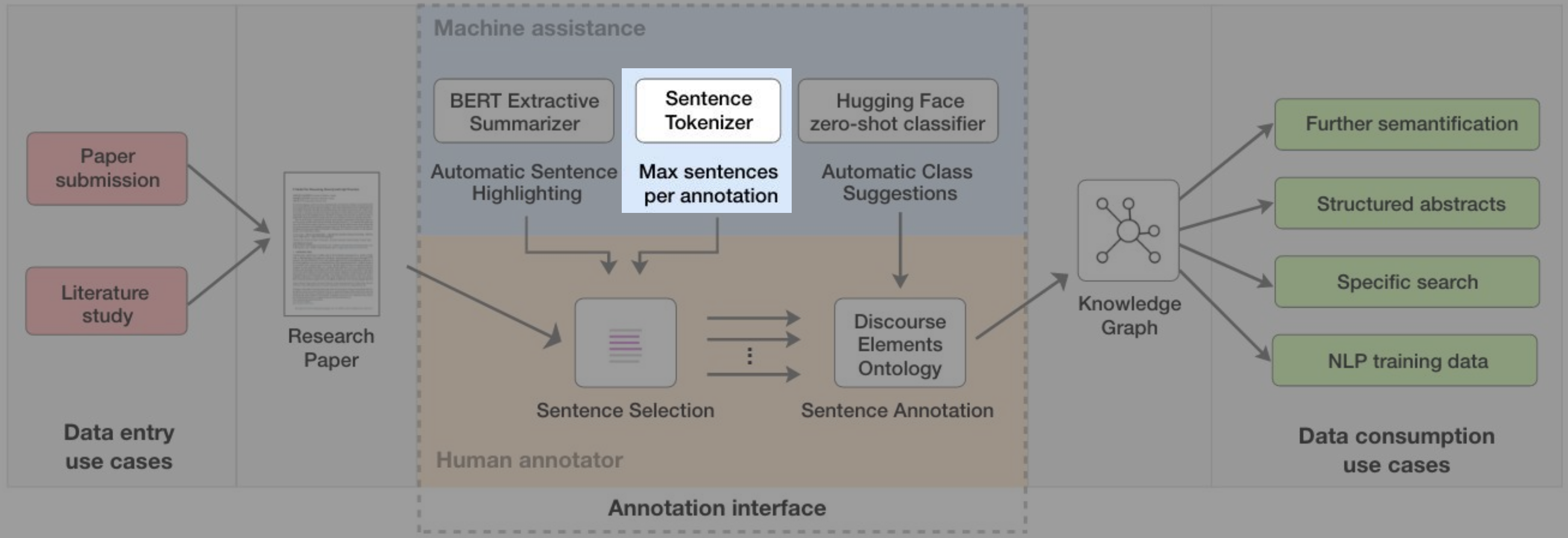
- Extractive summarization using BERT embeddings
- Summary is split by sentence endings and highlighted in the original PDF article

Highlighted sentence

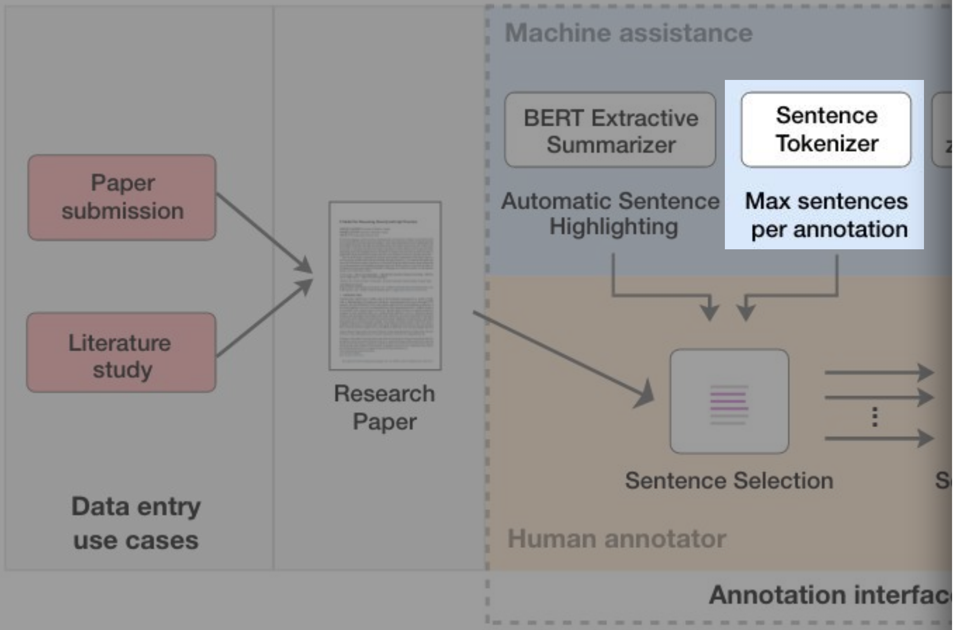

the demographics data shows, participants with varying levels of expertise participated in the study.
To determine the usability of the interface, we incorporated the System Usability Scale (SUS) [8] in the questionnaire. Furthermore, to determine the workload of the task we included questions from the NASA Task Load Index (TLX) [25]. This provides insights into the perceived workload by participants for the annotation task. To reduce the length of the questionnaire, we conducted the Raw TLX, which eliminates weighting the questions. Finally, we included

Activate highlighting

PDF sentence annotation

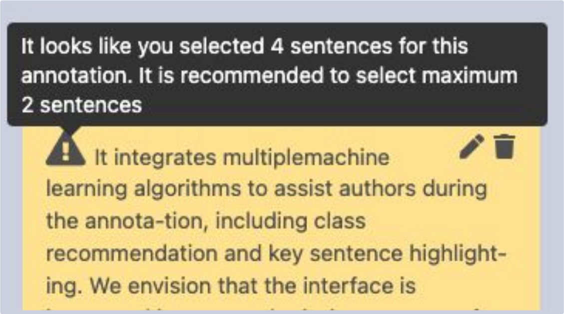


PDF sentence annotation

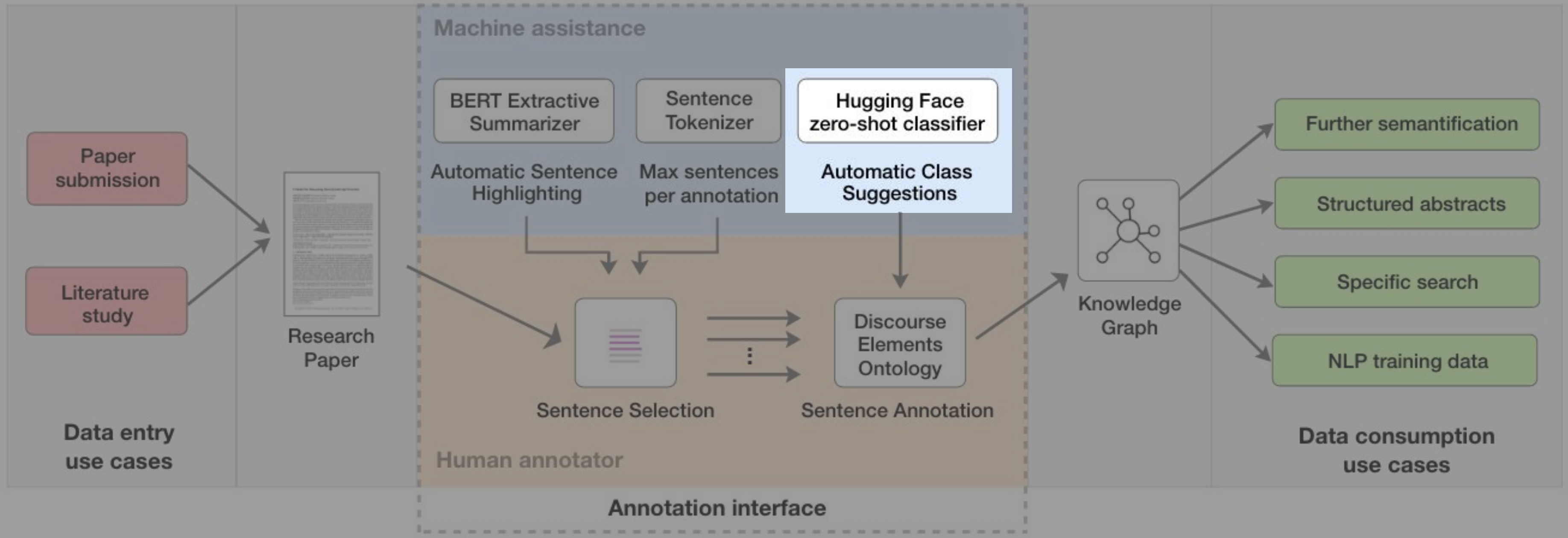


Maximum sentences per annotation

- Selected text is split per sentence
- Warning is displayed if more than two sentences are selected



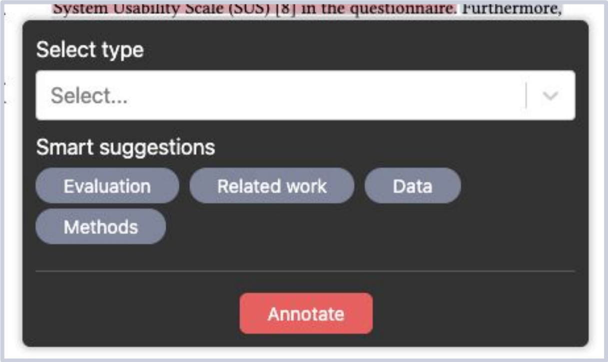
PDF sentence annotation

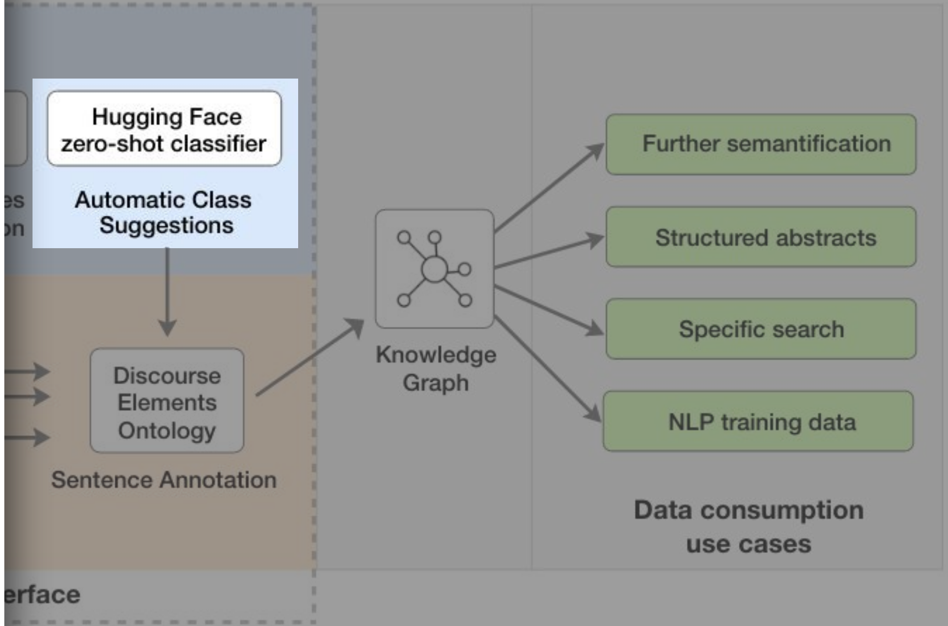


PDF sentence annotation

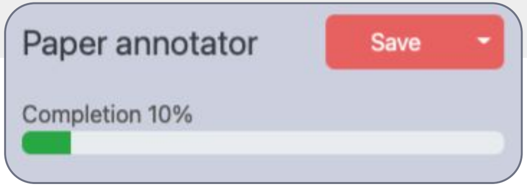
Automatic class suggestions

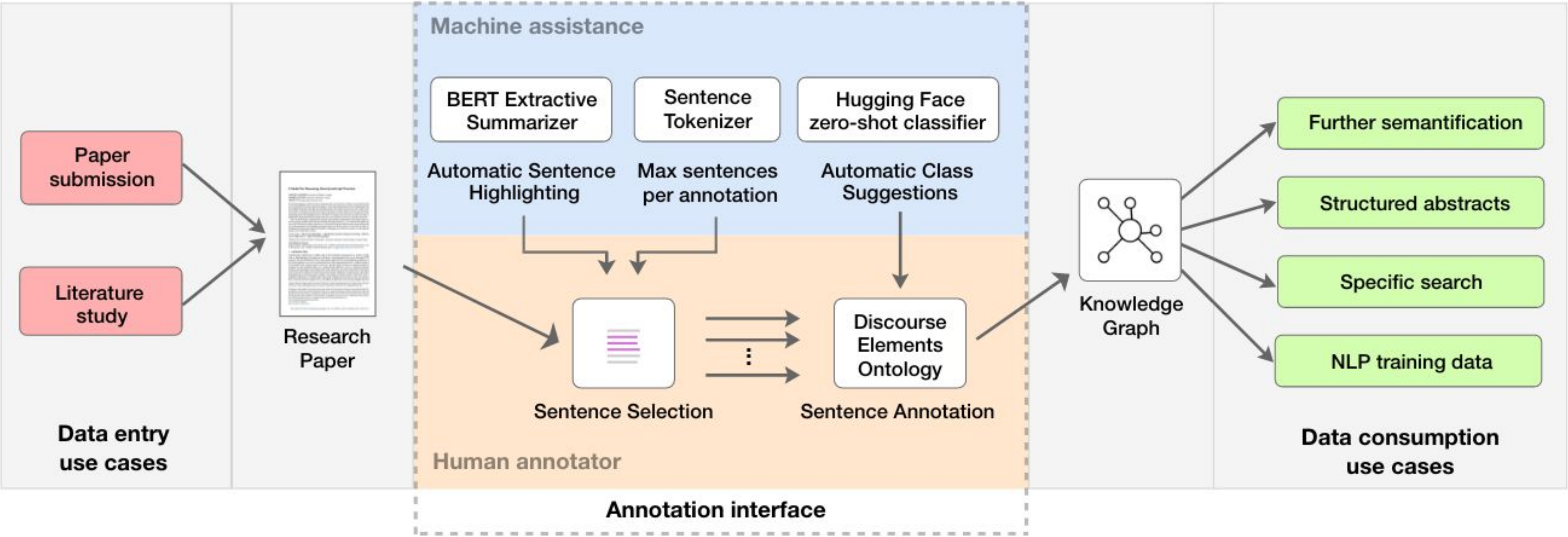
A zero-shot classifier is used (from Hugging Face) to provide annotation class suggestions based on the selected sentence



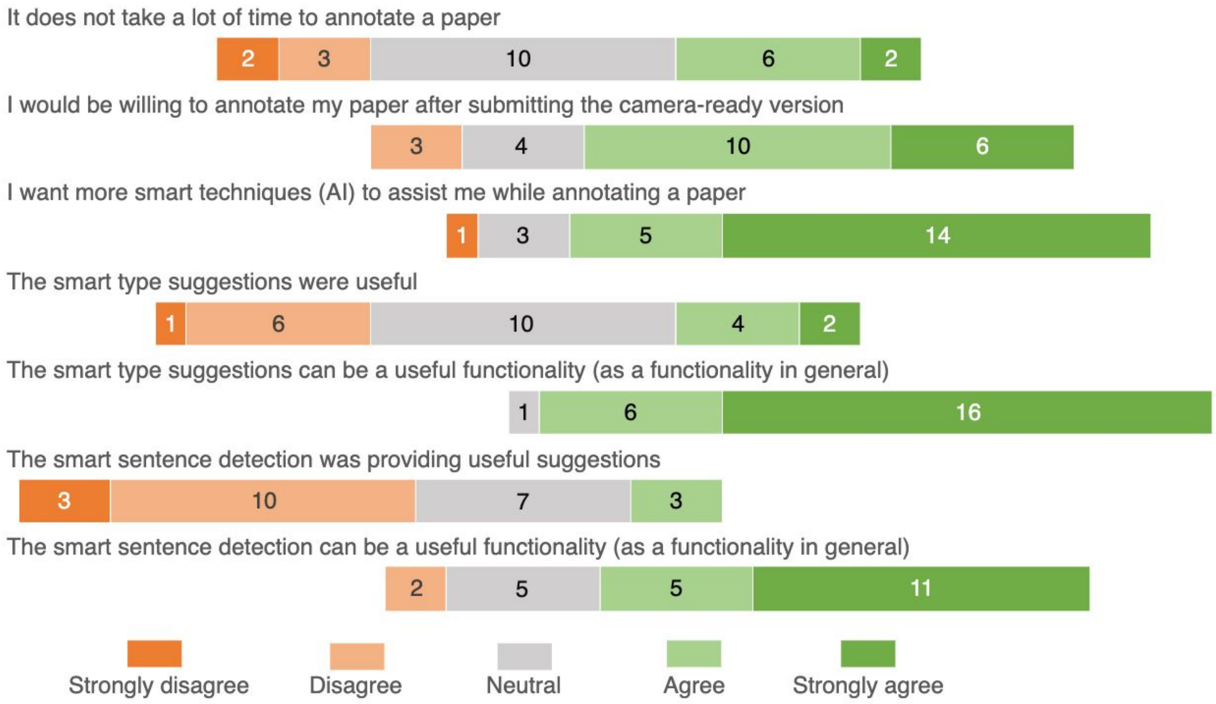


PDF sentence annotation





Paper annotator - Evaluation results



Human-AI collaboration in the ORKG

AI-Augmented

1. Smart suggestions

AI-supported tooltips helping users accomplish their tasks

2. Paper annotator

Annotation of key sentences in scholarly PDF articles

3. Survey extractor

Extract survey tables from existing papers

AI-Driven

4. TinyGenius

Microtasks to validate NLP generated statements

5. ORKG Ask

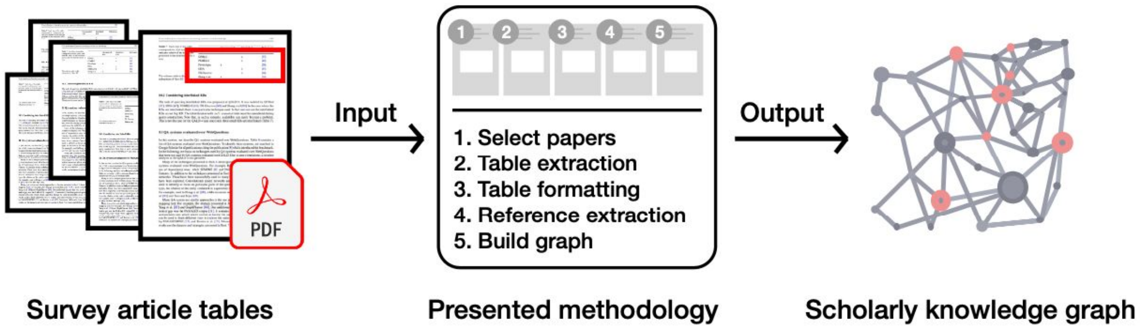
Tomorrow's topic

Literature surveys

Oelen, Allard, Markus Stocker, and Sören Auer. "Creating a scholarly knowledge graph from survey article tables." *International Conference on Asian Digital Libraries*. Cham: Springer International Publishing, 2020.

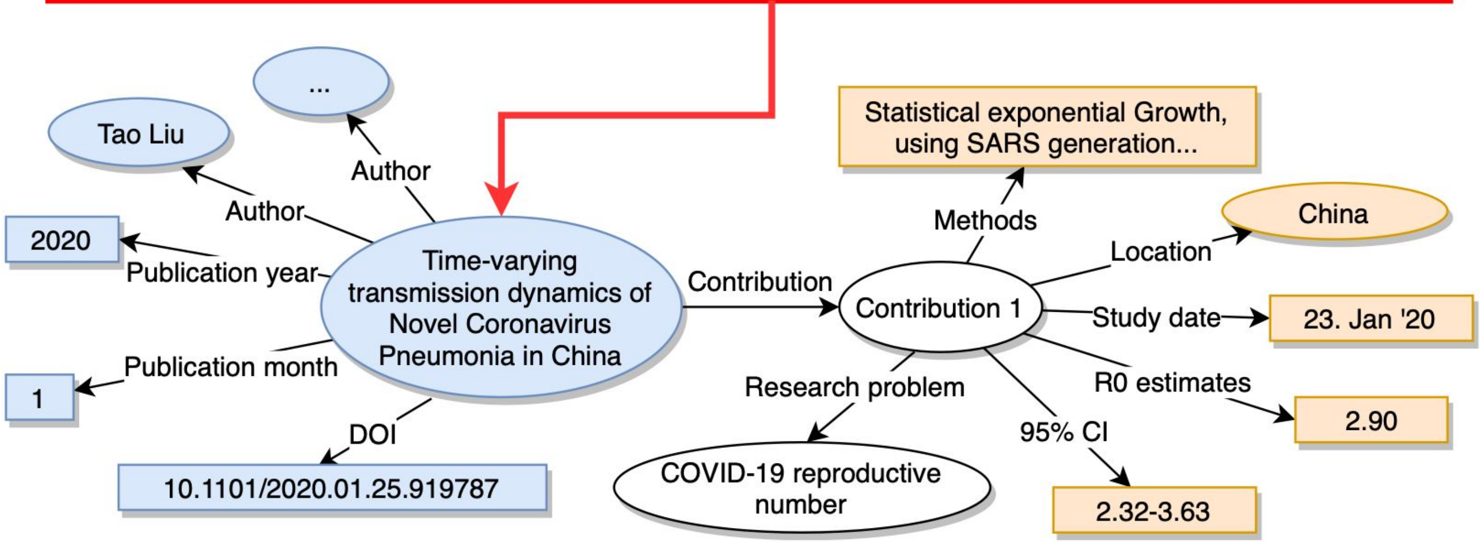
- Objective: we **leverage survey tables** to create a scholarly knowledge graph

- Literature surveys (or reviews): consist of **relevant** and **high-quality** research data that has been **manually curated** by domain experts



Example of survey table import

Study	Location	Study date	Methods	R_0 estimates	95% CI
Joseph et al.[1]	Wuhan	31 Dec '19 - 28 Jan '20	Stochastic Markov Chain...	2.68	2.47-2.86
Shen et al.[2]	Hubei province	12-22 Jan. '20	Mathematical model, dynamic...	6.49	6.31-6.66
Liu et al.[3]	China and overseas	23. Jan '20	Statistical exponential Growth...	2.90	2.32-3.63



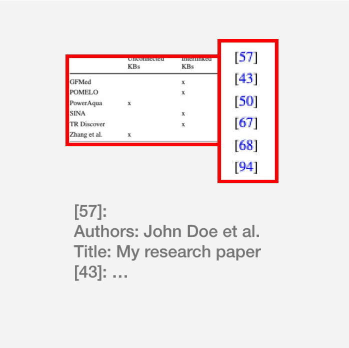
Methodology

Extract tables from survey papers to create a scholarly knowledge graph

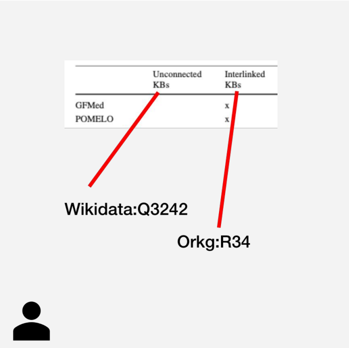
Paper selection

Table extraction

Reference extraction



Ontology mapping



Import papers

Combine collected data from:

2. Table extraction
3. Reference extraction
4. Ontology mapping

↓

Import in ORKG via Python API

= Human assisted by AI

User interface

- **Select survey table in PDF article**
- Fix formatting with spreadsheet editor
- Reference extraction
- Ontology mapping

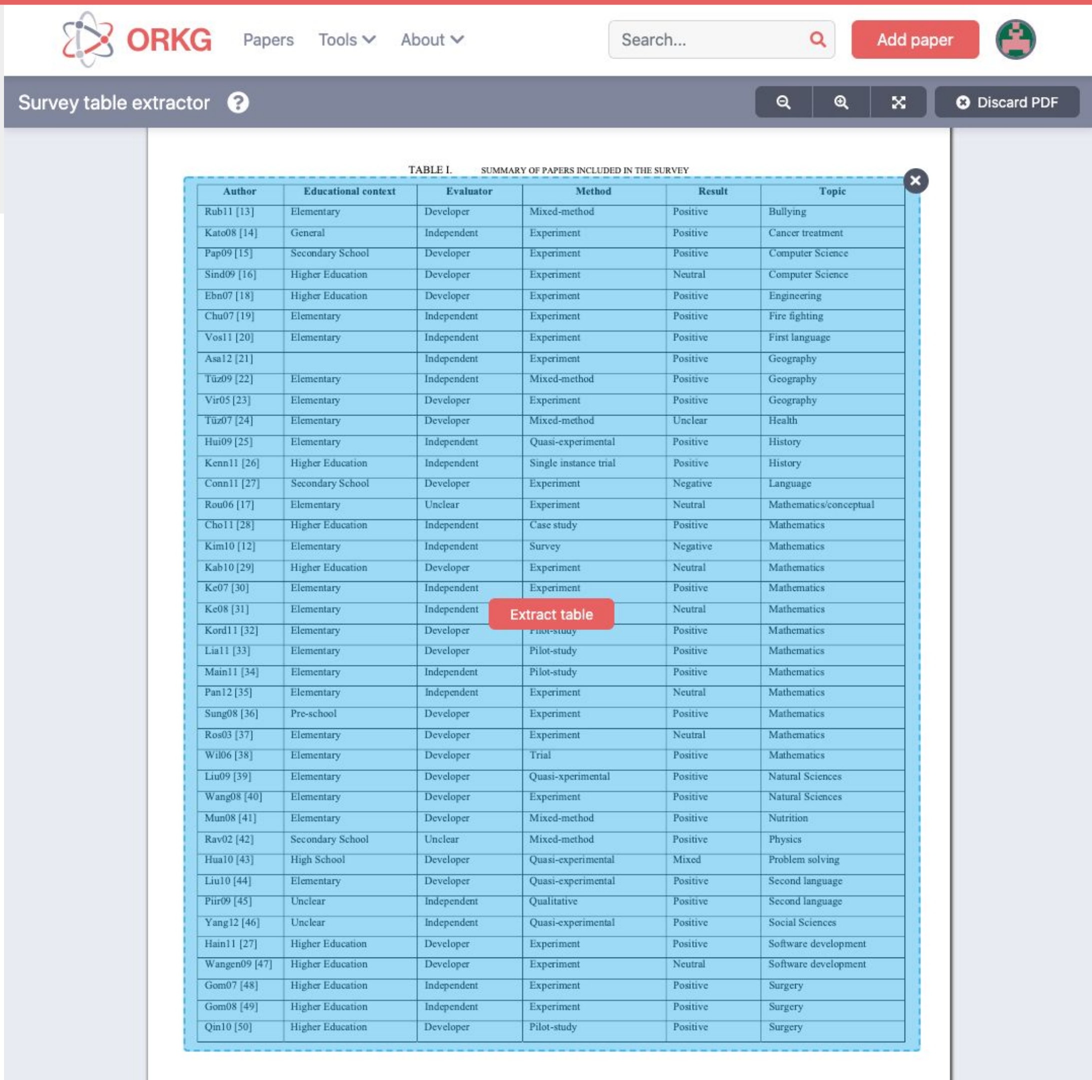


TABLE I. SUMMARY OF PAPERS INCLUDED IN THE SURVEY

Author	Educational context	Evaluator	Method	Result	Topic
Rub11 [13]	Elementary	Developer	Mixed-method	Positive	Bullying
Kato08 [14]	General	Independent	Experiment	Positive	Cancer treatment
Pap09 [15]	Secondary School	Developer	Experiment	Positive	Computer Science
Sind09 [16]	Higher Education	Developer	Experiment	Neutral	Computer Science
Ebn07 [18]	Higher Education	Developer	Experiment	Positive	Engineering
Chu07 [19]	Elementary	Independent	Experiment	Positive	Fire fighting
Vos11 [20]	Elementary	Independent	Experiment	Positive	First language
Asa12 [21]		Independent	Experiment	Positive	Geography
Tüz09 [22]	Elementary	Independent	Mixed-method	Positive	Geography
Vir05 [23]	Elementary	Developer	Experiment	Positive	Geography
Tüz07 [24]	Elementary	Developer	Mixed-method	Unclear	Health
Hui09 [25]	Elementary	Independent	Quasi-experimental	Positive	History
Kenn11 [26]	Higher Education	Independent	Single instance trial	Positive	History
Conn11 [27]	Secondary School	Developer	Experiment	Negative	Language
Rou06 [17]	Elementary	Unclear	Experiment	Neutral	Mathematics/conceptual
Cho11 [28]	Higher Education	Independent	Case study	Positive	Mathematics
Kim10 [12]	Elementary	Independent	Survey	Negative	Mathematics
Kab10 [29]	Higher Education	Developer	Experiment	Neutral	Mathematics
Ke07 [30]	Elementary	Independent	Experiment	Positive	Mathematics
Ke08 [31]	Elementary	Independent	[illegible]	Neutral	Mathematics
Kord11 [32]	Elementary	Developer	Pilot-study	Positive	Mathematics
Lia11 [33]	Elementary	Developer	Pilot-study	Positive	Mathematics
Main11 [34]	Elementary	Independent	Pilot-study	Positive	Mathematics
Pan12 [35]	Elementary	Independent	Experiment	Neutral	Mathematics
Sung08 [36]	Pre-school	Developer	Experiment	Positive	Mathematics
Ros03 [37]	Elementary	Developer	Experiment	Neutral	Mathematics
Wil06 [38]	Elementary	Developer	Trial	Positive	Mathematics
Liu09 [39]	Elementary	Developer	Quasi-xperimental	Positive	Natural Sciences
Wang08 [40]	Elementary	Developer	Experiment	Positive	Natural Sciences
Mun08 [41]	Elementary	Developer	Mixed-method	Positive	Nutrition
Rav02 [42]	Secondary School	Unclear	Mixed-method	Positive	Physics
Hua10 [43]	High School	Developer	Quasi-experimental	Mixed	Problem solving
Liu10 [44]	Elementary	Developer	Quasi-experimental	Positive	Second language
Piir09 [45]	Unclear	Independent	Qualitative	Positive	Second language
Yang12 [46]	Unclear	Independent	Quasi-experimental	Positive	Social Sciences
Hain11 [27]	Higher Education	Developer	Experiment	Positive	Software development
Wangen09 [47]	Higher Education	Developer	Experiment	Neutral	Software development
Gom07 [48]	Higher Education	Independent	Experiment	Positive	Surgery
Gom08 [49]	Higher Education	Independent	Experiment	Positive	Surgery
Qin10 [50]	Higher Education	Developer	Pilot-study	Positive	Surgery

User interface

- Select survey table in PDF article
- **Fix formatting with spreadsheet editor**
- Reference extraction
- Ontology mapping

Table extraction

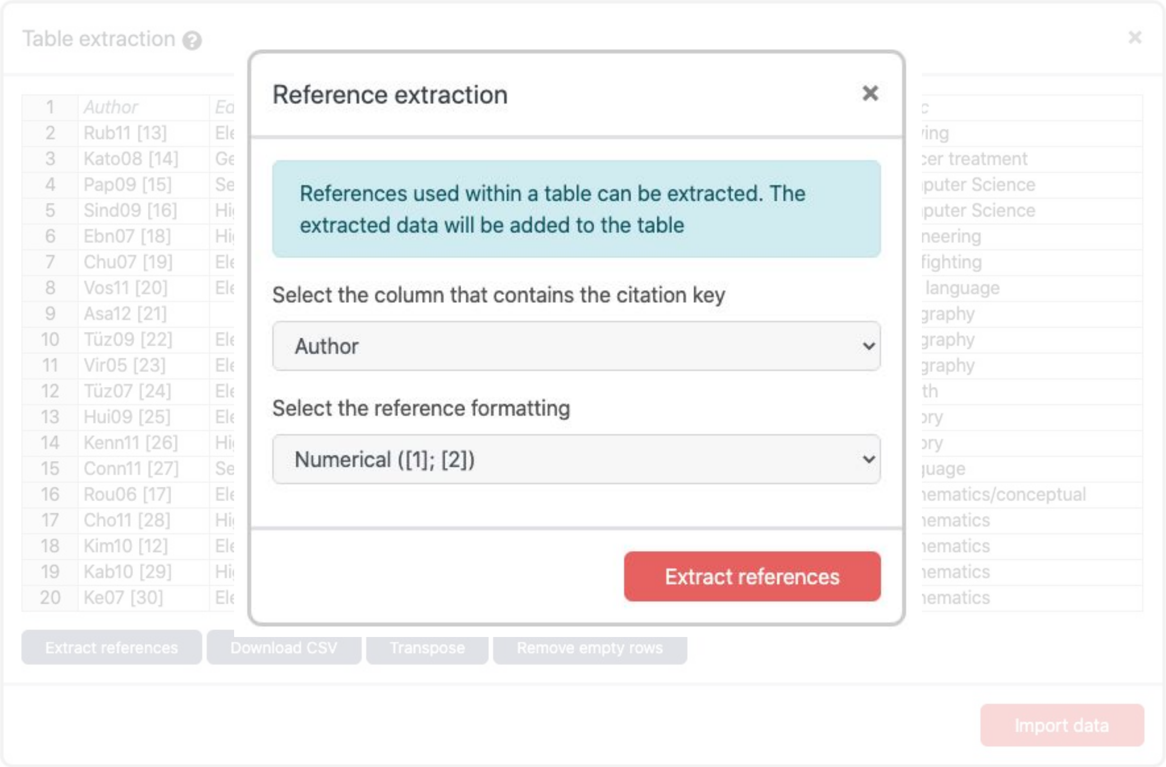
	Author	Educational context	Evaluator	Method	Result	Topic
2	Rub11 [13]	Elementary	Developer	Mixed-method	Positive	Bullying
3	Kato08 [14]	General	Independent	Experiment	Positive	Cancer treatment
4	Pap09 [15]	Secondary School	Developer	Experiment	Positive	Computer Science
5	Sind09 [16]	Higher Education	Developer	Experiment	Neutral	Computer Science
6	Ebn07 [18]	Higher Education	Developer	Experiment	Positive	Engineering
7	Chu07 [19]	Elementary	Independent	Experiment	Positive	Fire fighting
8	Vos11 [20]	Elementary	Independent	Experiment	Positive	First language
9	Asa12 [21]		Independent	Experiment	Positive	Geography
10	Tüz09 [22]	Elementary	Independent	Mixed-method	Positive	Geography
11	Vir05 [23]	Elementary	Developer	Experiment	Positive	Geography
12	Tüz07 [24]	Elementary	Developer	Mixed-method	Unclear	Health
13	Hui09 [25]	Elementary	Independent	Quasi-experimental	Positive	History
14	Kenn11 [26]	Higher Education	Independent	Single instance trial	Positive	History
15	Conn11 [27]	Secondary School	Developer	Experiment	Negative	Language
16	Rou06 [17]	Elementary	Unclear	Experiment	Neutral	Mathematics/conceptual
17	Cho11 [28]	Higher Education	Independent	Case study	Positive	Mathematics
18	Kim10 [12]	Elementary	Independent	Survey	Negative	Mathematics
19	Kab10 [29]	Higher Education	Developer	Experiment	Neutral	Mathematics
20	Ke07 [30]	Elementary	Independent	Experiment	Positive	Mathematics

Extract references | Download CSV | Transpose | Remove empty rows

Import data

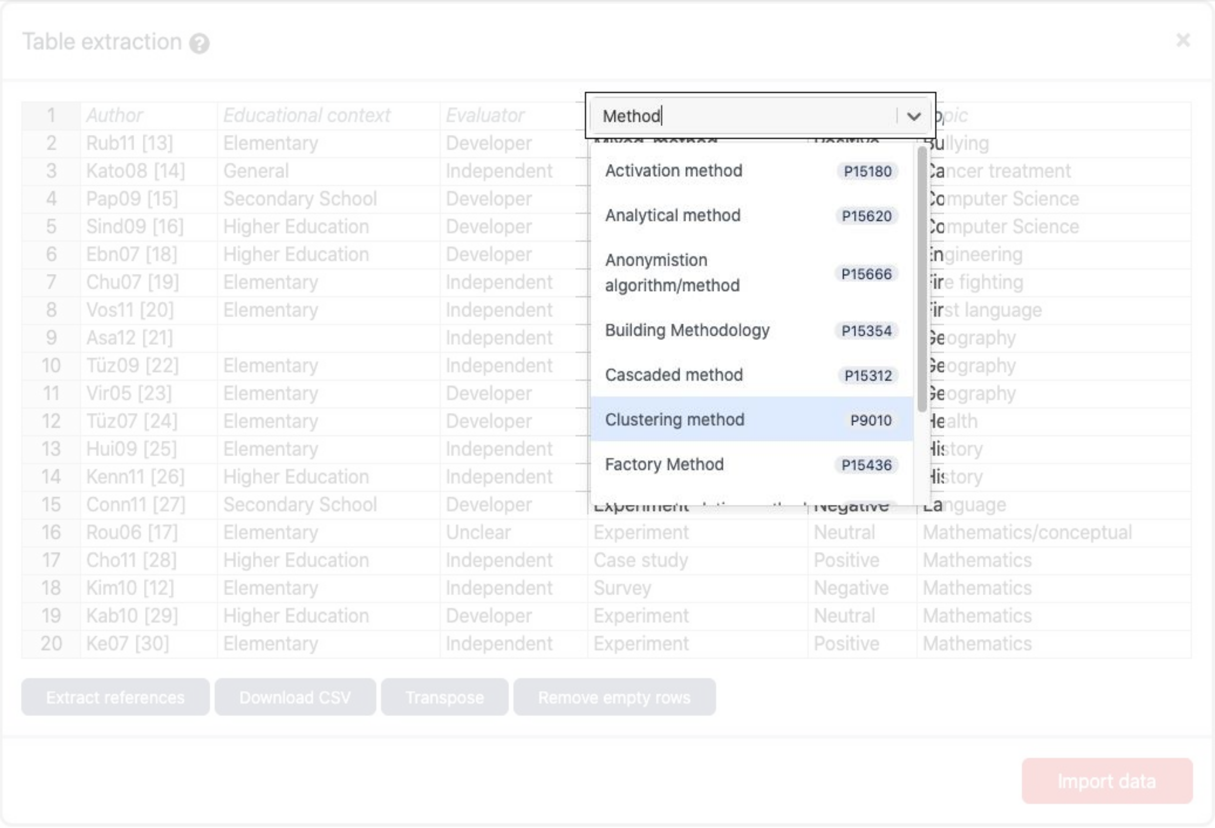
User interface

- Select survey table in PDF article
- Fix formatting with spreadsheet editor
- **Reference extraction**
- Ontology mapping



User interface

- Select survey table in PDF article
- Fix formatting with spreadsheet editor
- Reference extraction
- **Ontology mapping**



Survey extractor tool - Results

Description	Amount
Paper selection	
Amount of evaluated papers	415
Amount of selected papers	92
Table extraction	
Total amount of extractions (partial tables)	265
Amount of extracted complete tables	160
Reference extraction	
Found references	2 069
Not found references	1 137
Build graph	
Individual amount of imported papers	2 626
Imported data cells (with metadata)	40 584
Imported data cells (without metadata)	21 240

#	Issue	Percentage %
1	Columns are not extracted correctly	26
2	Rows are not extracted correctly	14
3	Empty columns in the extracted table	14
4	Text not correctly recognized (e.g., missing letters or formulas)	12
5	Issue with table header text	12
6	Vertical text not imported correctly	4
7	Cell value not supported (e.g., use of image instead of text check marks)	3
8	Table within table not extracted correctly	3

Human-AI collaboration in the ORKG

AI-Augmented

1. Smart suggestions

AI-supported tooltips helping users accomplish their tasks

2. Paper annotator

Annotation of key sentences in scholarly PDF articles

3. Survey extractor

Extract survey tables from existing papers

AI-Driven

4. TinyGenius

Microtasks to validate NLP generated statements

5. ORKG Ask

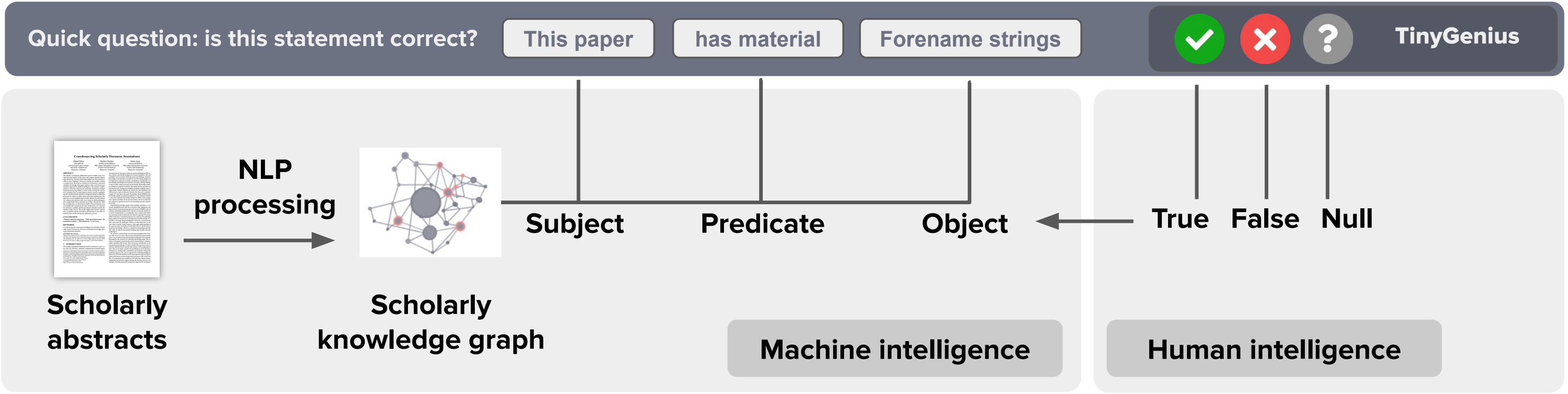
Tomorrow's topic

TinyGenius

Oelen, Allard, Markus Stocker, and Sören Auer. "TinyGenius: intertwining natural language processing with microtask crowdsourcing for scholarly knowledge graph creation." *Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries*. 2022.

- Leverage existing NLP tools to **process large quantities** of scholarly data
- Ask any user/visitor to validate the statements using **simple tasks** (aka microtasks)
- Users that are normally "**content consumers**" can become "**content creators**" as microtasks lower the entrance barrier to contribute significantly

TinyGenius - Validate NLP with microtasks

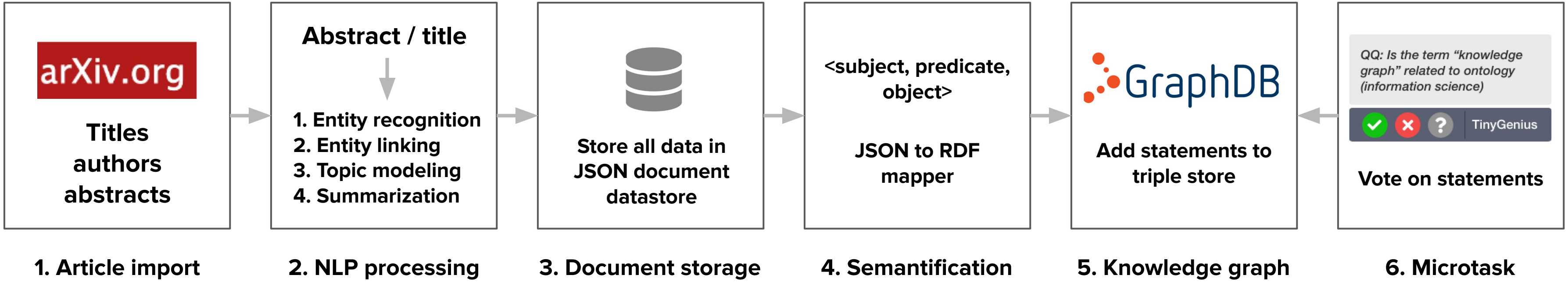


TinyGenius - Approach

The six-step approach **extracts** knowledge from scholarly articles, **creates** a knowledge graph, and let’s humans **validate** the knowledge

TinyGenius - Approach

The six-step approach **extracts** knowledge from scholarly articles, **creates** a knowledge graph, and let’s humans **validate** the knowledge



TinyGenius - NLP tools and templates

Task-specific **question templates** are used for the microtask generation

1. **Open information extraction *(ORKG abstract annotator & ORKG title parser)***

2. **Entity linking *(Ambiverse NLU)***

3. **Topic modeling *(CSO Classifier)***

4. **Summarization *(Hugging face)***

TinyGenius - NLP tools and templates

Task-specific **question templates** are used for the microtask generation

1. Open information extraction *(ORKG abstract annotator & ORKG title parser)*

QQ: Is this statement correct? | This paper | has material | Forename strings | TinyGenius

2. Entity linking *(Ambiverse NLU)*

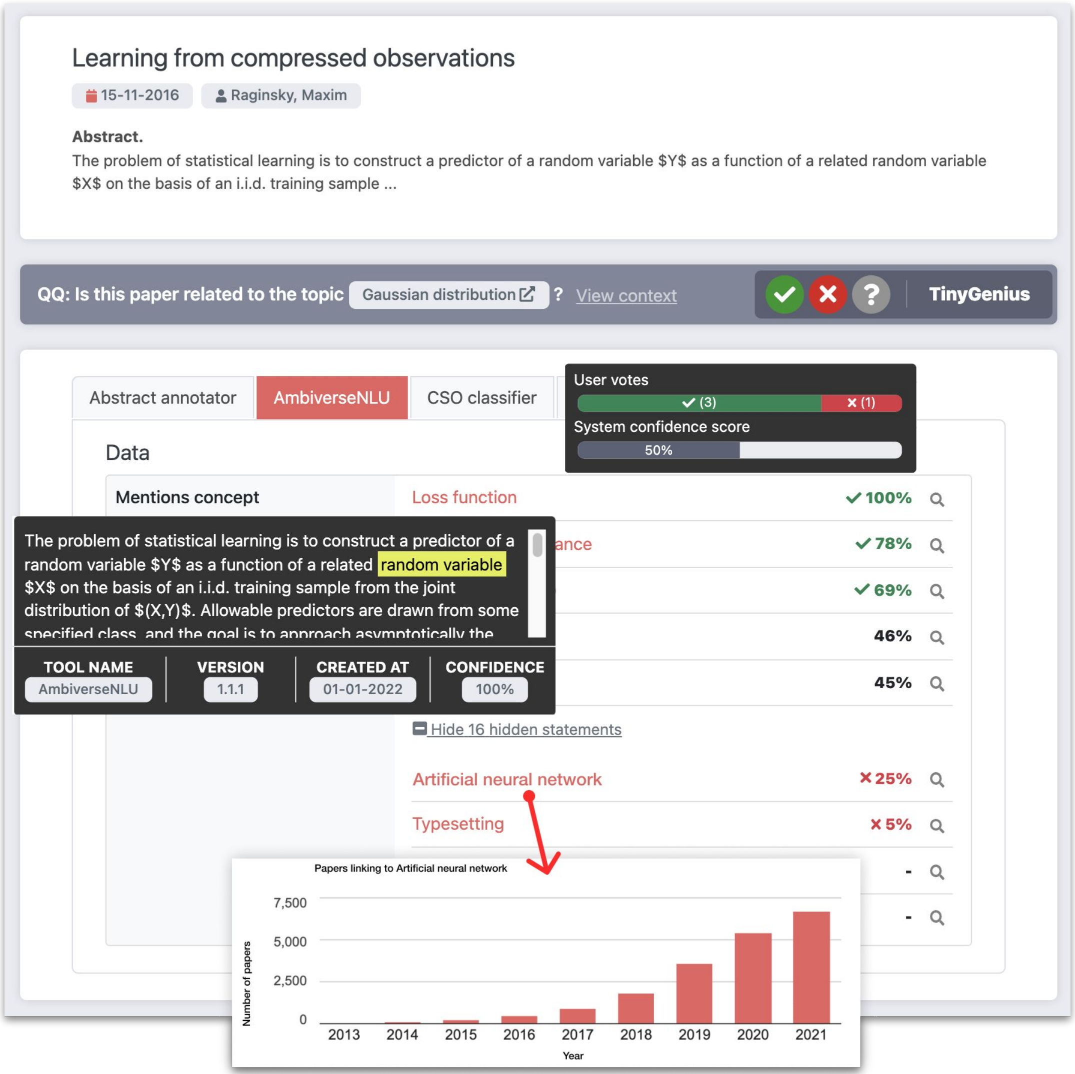
QQ: Is the term “spreading activation” related to | spreading activation (Wikidata) | ? | TinyGenius

3. Topic modeling *(CSO Classifier)*

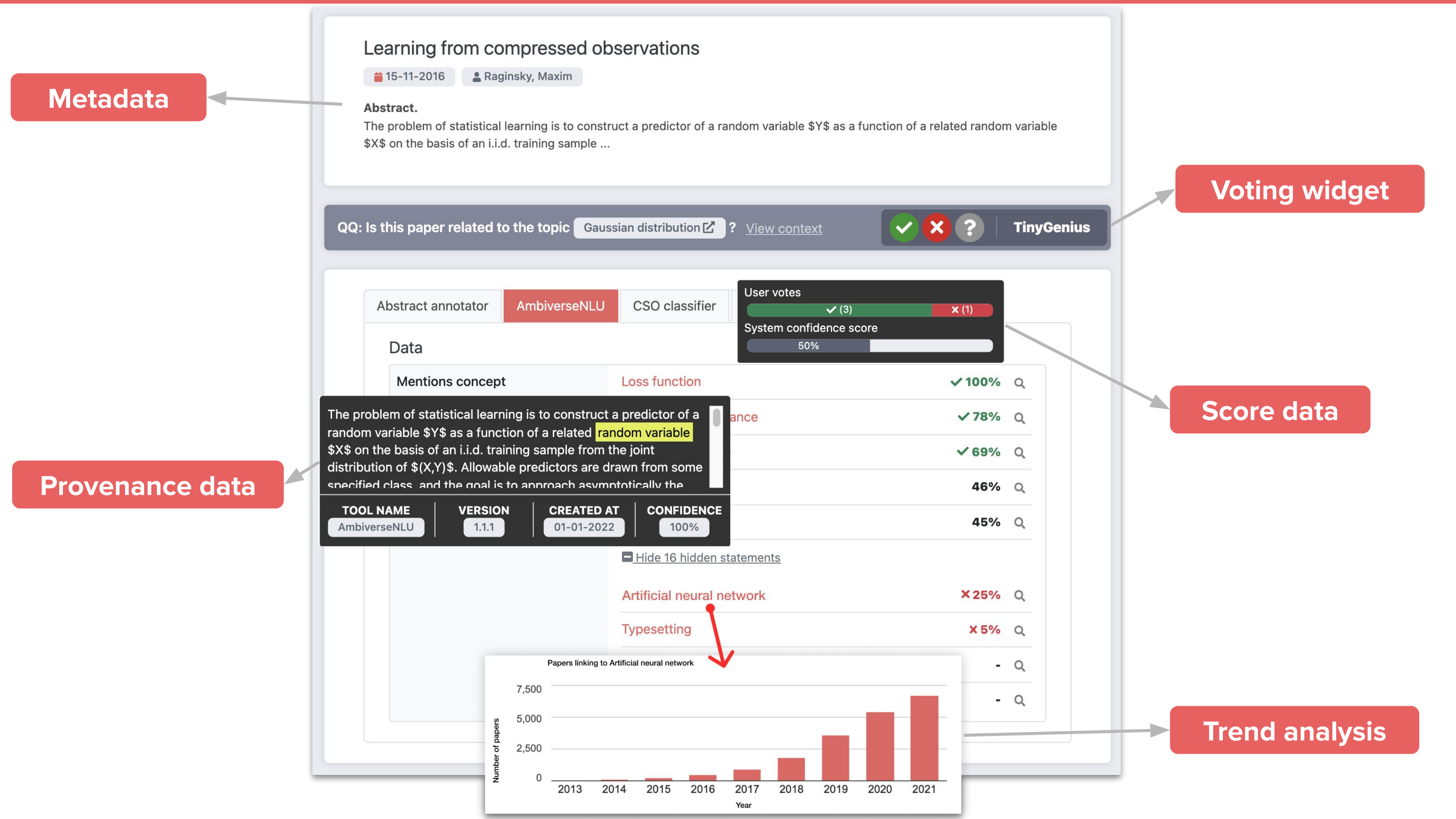
QQ: Is this paper related to the topic | hyperlinks | ? | TinyGenius

4. Summarization *(Hugging face)*

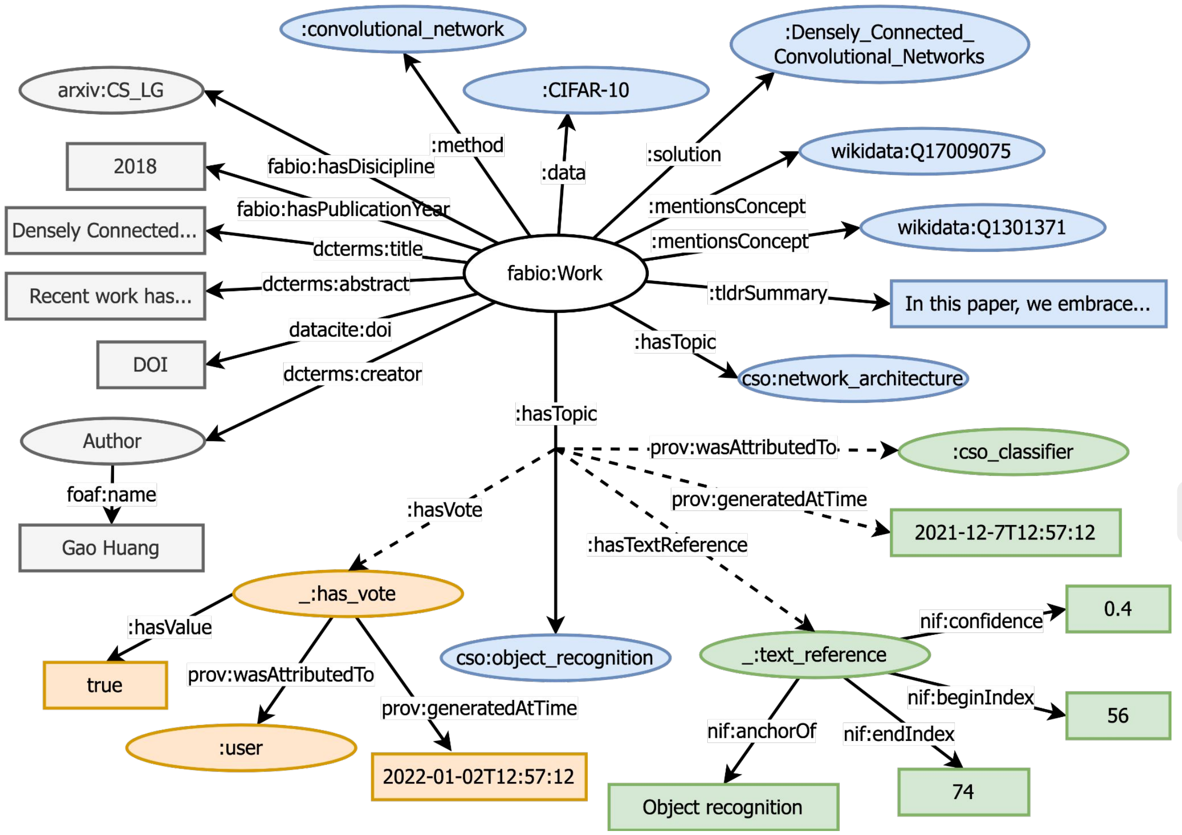
QQ: Does this sentence summarize the paper? *A brain-inspired search engine named DeveloperBot...* | TinyGenius



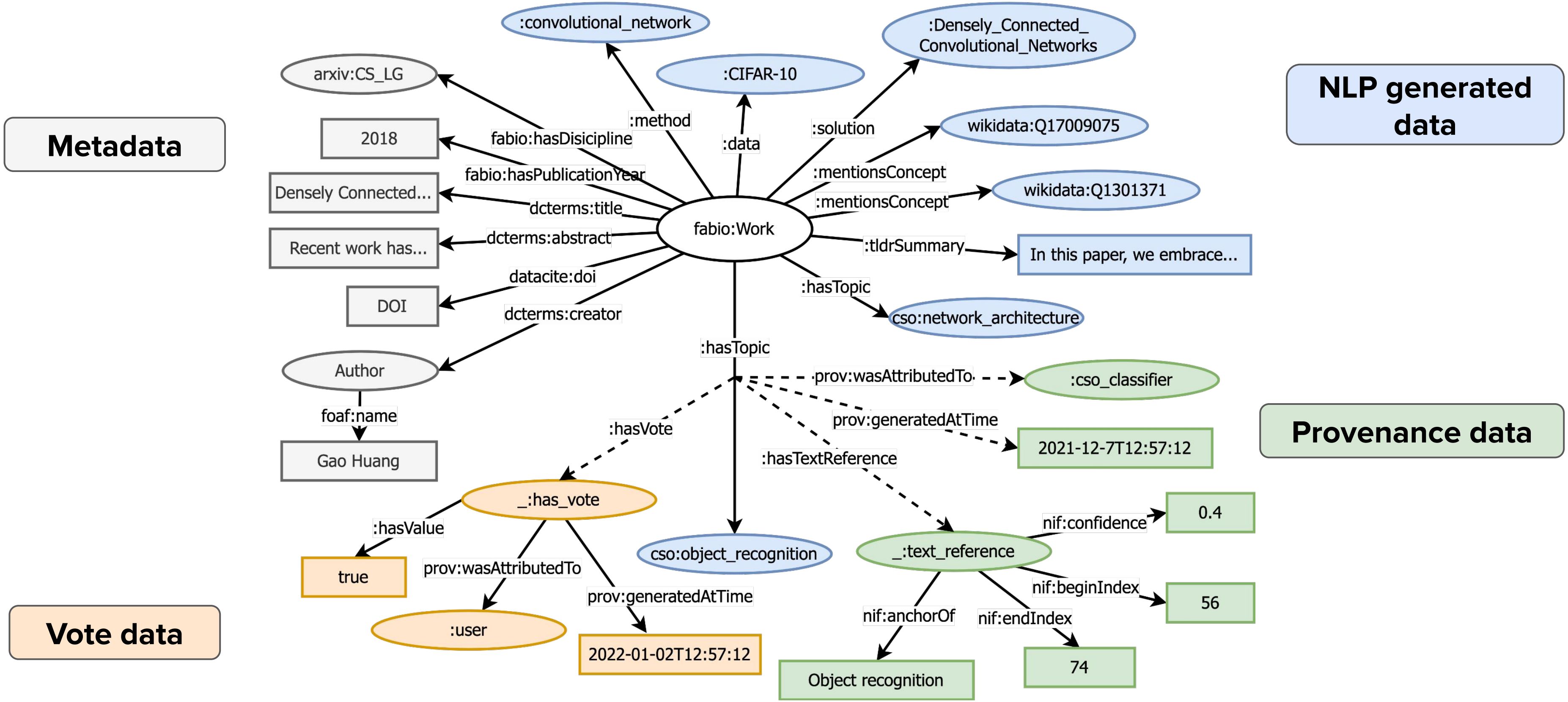
Learning from compressed observations
15-11-2016
Raginsky, Maxim
Abstract.
The problem of statistical learning is to construct a predictor of a random variable Y as a function of a related random variable X on the basis of an i.i.d. training sample ...
QQ: Is this paper related to the topic
Gaussian distribution
?
View context
TinyGenius
Abstract annotator
AmbiverseNLU
CSO classifier
User votes
(3)
(1)
System confidence score
50%
Data
Mentions concept
Loss function
100%
ance
78%
69%
46%
45%
The problem of statistical learning is to construct a predictor of a random variable Y as a function of a related random variable X on the basis of an i.i.d. training sample from the joint distribution of (X,Y). Allowable predictors are drawn from some specified class, and the goal is to approach asymptotically the
TOOL NAME
AmbiverseNLU
VERSION
1.1.1
CREATED AT
01-01-2022
CONFIDENCE
100%
Hide 16 hidden statements
Artificial neural network
25%
Typesetting
5%
-
-
Papers linking to Artificial neural network
7,500
5,000
2,500
0
Number of papers
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


Metadata
Voting widget
Score data
Provenance data
Trend analysis
Learning from compressed observations
15-11-2016
Raginsky, Maxim
Abstract.
The problem of statistical learning is to construct a predictor of a random variable Y as a function of a related random variable X on the basis of an i.i.d. training sample ...
QQ: Is this paper related to the topic Gaussian distribution ? View context
TinyGenius
Abstract annotator
AmbiverseNLU
CSO classifier
User votes
(3)
(1)
System confidence score
50%
Data
Mentions concept
Loss function
100%
ance
78%
69%
46%
45%
The problem of statistical learning is to construct a predictor of a random variable Y as a function of a related random variable X on the basis of an i.i.d. training sample from the joint distribution of (X,Y). Allowable predictors are drawn from some specified class, and the goal is to approach asymptotically the
TOOL NAME
AmbiverseNLU
VERSION
1.1.1
CREATED AT
01-01-2022
CONFIDENCE
100%
Hide 16 hidden statements
Artificial neural network
25%
Typesetting
5%
-
-
Papers linking to Artificial neural network
7,500
5,000
2,500
0
Number of papers
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year

Data model | Provenance



Data model | Provenance



TinyGenius - Queries

```
SELECT DISTINCT * WHERE {
    <<tinygenius:1802.01528  ?pred  ?obj>>
                dcterms:creator  tinygenius:ambiverse_nlu .
    tinygenius:ambiverse_nlu  dcterms:hasVersion  "1.1.1" .
}
```

```
SELECT DISTINCT * WHERE {
    <<tinygenius:1608.06993  ?pred  ?obj>>  ?provPred  ?provObj .
    OPTIONAL {
        ?provObj  ?provPred2  ?provObj2 .
    }
}
```

```
SELECT ?year (COUNT(DISTINCT ?paper) AS ?count) WHERE {
    ?paper  a                          fabio:Work ;
            fabio:hasPublicationYear  ?year ;
            ?predicate                 tinygenius:artificial_neural_network .
} GROUP BY ?year
```

Results - Statistics

Triple related statistics

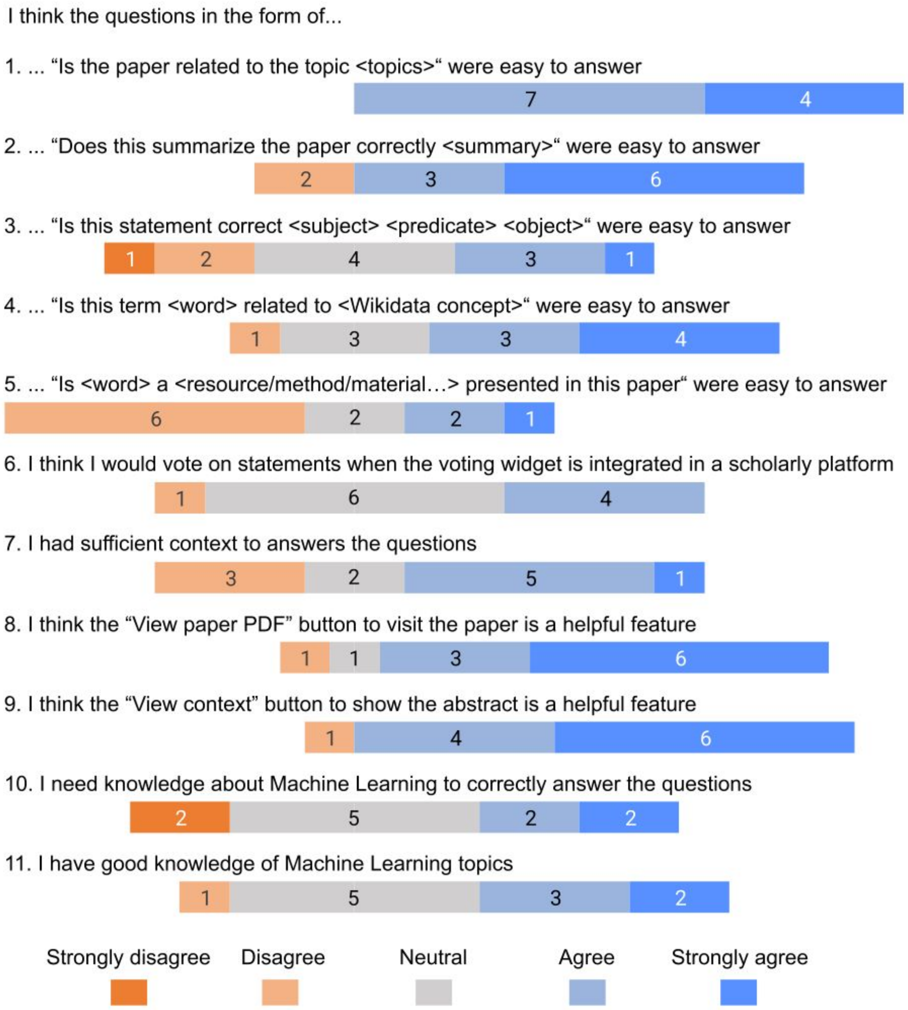
Processed articles	95,376*
Triples metadata	1,521,492
Triples provenance	47,595,706
Triples total	65,608,902
Average number of triples per article	688

** Approximately 5% of the complete arXiv corpus. Includes all papers classified as "Machine Learning"*

Processing time per NLP tool

Abstract annotator	62,056s (≈ 17 hour)
Title parser	87s
Ambiverse NLU	137,060s (≈ 38 hour)
CSO classifier	27,803s (≈ 8 hour)
Summarizer	N/A

TinyGenius - Results



Thank you!
Any Questions?

Allard Oelen
allard.oelen@tib.eu
linkedin.com/in/allard-oelen

Meet the team:
https://orkg.org/about/9/Team

Co-funded by the Horizon 2020 programme
of the European Union