Koala-UI: An Interactive User Interface for Tabular Data Linking

Roberto Avogadro[1,*,†], Iroshani Jayawardene[1,†], Xiang Ma[1,†], Ahmet Soylu[2,†] and Dumitru Roman[1,3,†]

[1] SINTEF AS, POB 124, Blindern, 0314 Oslo, Norway
[2] Kristiania University College, Oslo, Norway
[3] OsloMet—Oslo Metropolitan University, Oslo, Norway

Abstract

This paper introduces Koala-UI – a user interface system aimed at simplifying the entity linking process within data enrichment pipelines. Koala-UI provides an intuitive mechanism for linking entities across datasets, combining automation with human feedback to ensure accurate and consistent data. Koala-UI was successfully applied in use cases such as public procurement, where it enabled enrichment of a tenders dataset by linking entities to external knowledge graphs. Future developments will focus on expanding its backend to support additional models and enhance its human-in-the-loop capabilities.

Keywords

Entity Linking, Data Enrichment, Data Pipelines, User Interfaces

1. Introduction

In today´s rapidly evolving data-driven world, organizations are increasingly relying on large and complex datasets, particularly in tabular form, to make strategic decisions. A critical step in effectively leveraging such data is *tabular data linking*, where entities in different tables are identified and connected to their corresponding records in external reference knowledge bases. This process is essential for ensuring consistency, accuracy, and completeness of enriched datasets, which, in turn, supports tasks such as analytics, compliance monitoring, and artificial intelligence-driven applications [1, 2, 3, 4, 5].

Tabular data linking presents several challenges, particularly when dealing with large-scale or heterogeneous tables. Inconsistent naming conventions, ambiguous references, and evolving organizational structures can make it difficult to reconcile entities accurately [6, 7, 8]. Traditional methods, which often rely on rule-based systems or manual intervention, are inefficient and do not scale well to modern big data environments. Therefore, there is a growing demand for intuitive, user-friendly tools that allow both technical and non-technical users to perform complex data linking tasks with minimal effort while maintaining control over the results. This is where Koala-UI becomes a valuable asset.

Koala-UI is an interactive user interface designed specifically to address the complexities of tabular data linking in large datasets. It provides a streamlined, intuitive interface that simplifies the process of linking entities across tables, while integrating human feedback to ensure high accuracy. By leveraging automation and offering real-time feedback, Koala-UI helps organizations efficiently manage large volumes of tabular data, ensuring that the enriched data is well-prepared for further analysis or downstream applications.

This paper presents the design, implementation, and real-world application of Koala-UI. It explores how this tool enables scalable and efficient tabular data linking and highlights its key features, including

RuleML+RR'24: Companion Proceedings of the 8th International Joint Conference on Rules and Reasoning, September 16–22, 2024, Bucharest, Romania

[*] Corresponding author.

[†] These authors contributed equally.

roberto.avogadro@sintef.no (R. Avogadro); iroshani.jayawardene@sintef.no (I. Jayawardene); xiang.ma@sintef.no (X. Ma); ahmet.soylu@kristiania.no (A. Soylu); dumitru.roman@sintef.no (D. Roman)
0009-0005-7404-4123 (R. Avogadro); 0000-0002-8297-9763 (I. Jayawardene); 0000-0001-6465-0254 (X. Ma)

human-in-the-loop verification, automated entity linking, and robust data handling capabilities for large datasets. The role of Koala-UI within a broader data enrichment ecosystem is also discussed, showcasing its value in real-world use cases such as public procurement, where it was utilized to reconcile and enrich tender data.

2. Related Work

Entity linking plays an important role in data enrichment, enabling the association of entities in a dataset with their corresponding counterparts in external knowledge bases. Several techniques and associated tools were developed to streamline this process, offering varying levels of functionality, user interface design, and integration with external services.

SemTUI[1] is a modular framework aimed at enhancing the semantic enrichment of tabular data [9]. It provides a flexible interface for instance-level and schema-level annotations by integrating with multiple external services, including reconciliation and extension systems. The tool's modular design allows for easy customization, making it adaptable to a wide range of data enrichment needs, particularly in more complex scenarios [10, 11].

OpenRefine[2] is another popular open-source tool, well-known for its data cleaning and transformation capabilities. It includes features for linking datasets to external databases such as Wikidata. While primarily focused on data cleaning [12], OpenRefine also offers reconciliation services for entity linking [13], though its specialization in this area is more limited compared to tools that are designed specifically for entity linking tasks.

DataGraft[3] offers an accessible platform for managing, transforming, and linking data. It supports linking entities to external knowledge graphs and provides a variety of data transformation tools, similar to OpenRefine [14]. DataGraft's user interface is designed to accommodate both technical and non-technical users, making it a versatile option for data linking tasks [15]. Its emphasis on user experience and ease of integration across diverse data sources positions it as a notable option in the field.

Despite the versatility and functionalities offered by such tools, they still present limitations when handling large-scale datasets or providing detailed interaction options for users. Koala-UI addresses these gaps by offering a user-friendly interface that is simple to use while capable of managing large tables efficiently, thanks to its paginated results feature. The tool further enhances the user experience by providing instruments to explore and refine entity linking results, such as candidate scores, sorting by confidence, and filtering by types. Additionally, it enables users to manually search for potential matches by typing strings, allowing for greater flexibility. Another advantage of Koala-UI is its ability to integrate with various backends, as long as they follow a consistent API response format, making it adaptable to a wide range of entity linking services.

3. Koala-UI Overview

Koala-UI[4] is an interactive web interface designed to simplify the entity linking process, enabling the integration of data from diverse sources into enriched datasets. A demo video of Koala-UI is available on YouTube[5], which showcases its core functionalities and user interface. Its primary goal is to make complex entity linking tasks accessible to users without extensive technical expertise, improving the quality and consistency of data for downstream applications such as analytics and decision-making.

Koala-UI was developed to support scalable and efficient data enrichment workflows. The design focuses on user-friendliness, automation, and flexibility, ensuring the platform can be applied in various

[1] https://i2tunimib.github.io/I2T-docs

[2] https://openrefine.org

[3] https://www.eubusinessgraph.eu/datagraft

[4] https://github.com/enRichMyData/koala_ui

[5] https://www.youtube.com/watch?v=4Nc8bMBQpyE

scenarios and domains.

The following subsections introduce Koala-UI's design principles and core features, including usability, scalability, and human-in-the-loop verification, highlighting how these elements enable efficient entity linking for large datasets.

3.1. Design Principles

The design of Koala-UI is guided by the following principles:

- **Usability**: The interface was built with a user-first approach, ensuring that users with varying technical expertise can interact with the tool efficiently. Koala-UI's interface simplifies complex entity linking tasks by providing intuitive UI components, visual cues, and real-time feedback, making the process as straightforward as possible.
- **Scalability**: Koala-UI was designed to handle large-scale datasets efficiently. The architecture ensures that performance remains high even when processing thousands of records, making it suitable for applications involving large volumes of data across diverse domains.
- **Flexibility**: Koala-UI was built with extensibility in mind. Future versions aim to support additional models and algorithms that can be easily plugged in, allowing users to adapt the tool to various entity linking scenarios and data enrichment needs.
- **Human-Centric Interaction**: The design emphasizes the role of human users in verifying and improving entity linking accuracy. The system is structured to allow users to provide feedback at critical points in the entity matching process, improving overall data quality.

3.2. Core Features

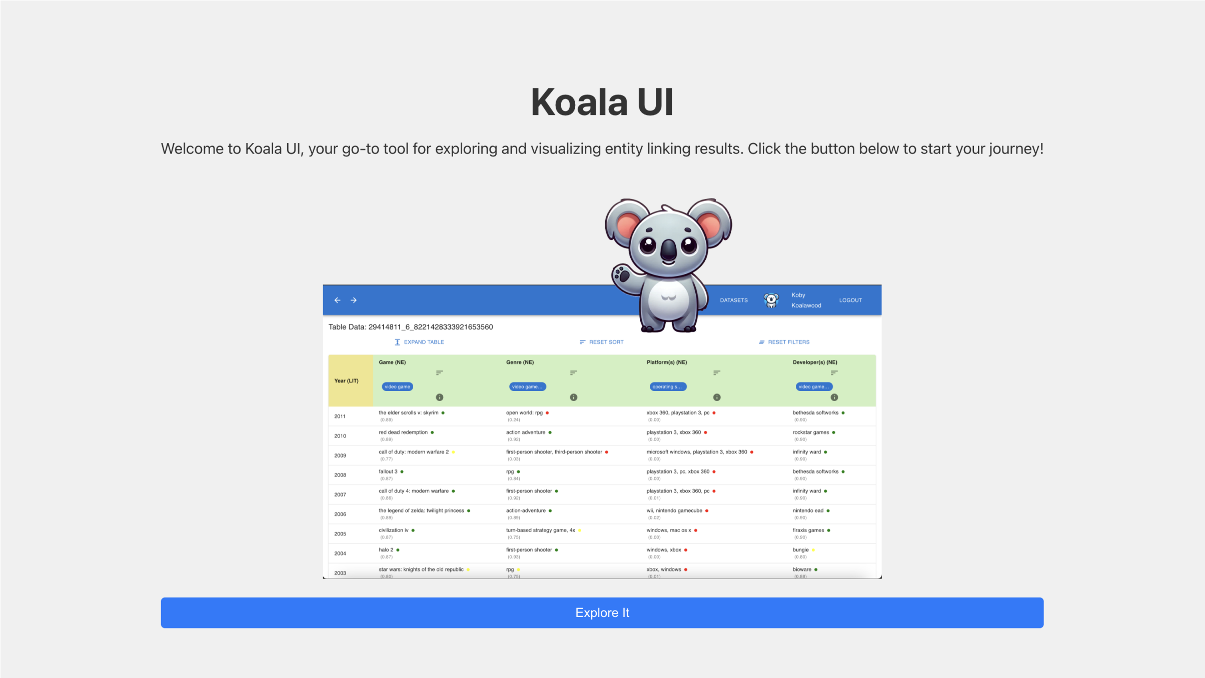
Koala-UI offers a range of features to streamline and improve the entity linking process:

- **Interactive Visualization**: Users are provided with real-time visual feedback on the entity linking process, including visual indicators for matches, confidence scores, and unresolved entities. This enables users to make informed decisions about entity verification and ensures data quality is maintained.
- **Automated Entity Linking**: The system uses advanced algorithms to automate much of the entity linking workflow, significantly reducing the need for manual intervention. The system provides confidence scores for each link, helping users focus on matches that need human validation.
- **Human-in-the-Loop Verification**: Koala-UI balances automation with human expertise by incorporating a verification mechanism. Users can review suggested entity links and correct or confirm matches, ensuring that the system learns and improves from user feedback.
- **Data Management for Large Datasets**: Koala-UI is optimized for large datasets, employing techniques such as lazy loading and efficient memory management. This ensures that users can manage and enrich large tables with minimal performance bottlenecks, making it suitable for industrial-scale applications.

Figure 1 presents the splash screen of Koala-UI. Figure 2 illustrates the entity matches along with their respective confidence scores, and Figure 3 highlights the type details extracted from a specific column.

4. Implementation

Koala-UI was implemented using a combination of frontend and backend technologies designed to ensure scalability, efficiency, and ease of use. While the backend manages user-specific data, the core data enrichment tasks, such as entity linking, are handled through external APIs integrated directly into the frontend.



Koala UI

Welcome to Koala UI, your go-to tool for exploring and visualizing entity linking results. Click the button below to start your journey!

Table Data: 29414811_6_8221428333921653560

Year (LIT)	Game (NE) video game	Genre (NE) video game...	Platform(s) (NE) operating s...	Developer(s) (NE) video game...
2011	the elder scrolls v: skyrim ● (0.89)	open world: rpg ● (0.24)	xbox 360, playstation 3, pc ● (0.00)	bethesda softworks ● (0.90)
2010	red dead redemption ● (0.89)	action adventure ● (0.92)	playstation 3, xbox 360 ● (0.00)	rockstar games ● (0.90)
2009	call of duty: modern warfare 2 ● (0.77)	first-person shooter, third-person shooter ● (0.03)	microsoft windows, playstation 3, xbox 360 ● (0.00)	infinity ward ● (0.90)
2008	fallout 3 ● (0.87)	rpg ● (0.84)	playstation 3, pc, xbox 360 ● (0.00)	bethesda softworks ● (0.90)
2007	call of duty 4: modern warfare ● (0.86)	first-person shooter ● (0.92)	playstation 3, xbox 360, pc ● (0.01)	infinity ward ● (0.90)
2006	the legend of zelda: twilight princess ● (0.89)	action-adventure ● (0.89)	wii, nintendo gamecube ● (0.02)	nintendo ead ● (0.90)
2005	civilization iv ● (0.87)	turn-based strategy game, 4x ● (0.75)	windows, mac os x ● (0.00)	firaxis games ● (0.90)
2004	halo 2 ● (0.87)	first-person shooter ● (0.93)	windows, xbox ● (0.00)	bungie ● (0.80)
2003	star wars: knights of the old republic ● (0.80)	rpg ● (0.75)	xbox, windows ● (0.01)	bioware ● (0.88)

Figure 1: Koala-UI splash screen.

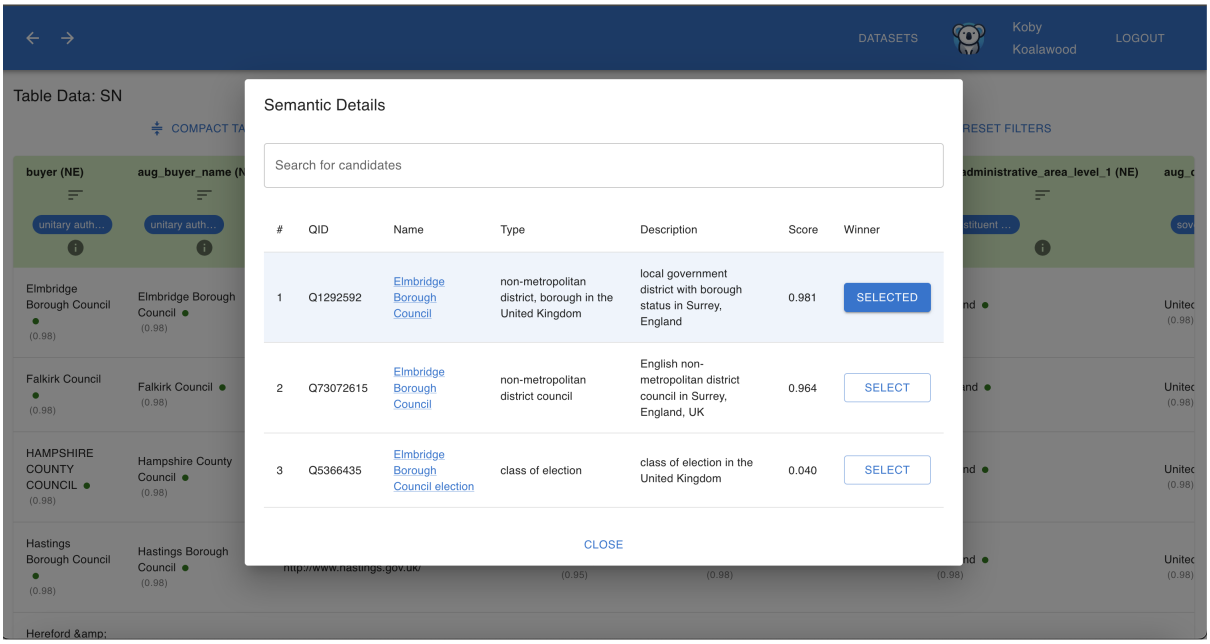


Table Data: SN

buyer (NE) unitary auth...	aug_buyer_name (N unitary auth...
Elmbridge Borough Council ● (0.98)	Elmbridge Borough Council ● (0.98)
Falkirk Council ● (0.98)	Falkirk Council ● (0.98)
HAMPSHIRE COUNTY COUNCIL ● (0.98)	Hampshire County Council ● (0.98)
Hastings Borough Council ● (0.98)	Hastings Borough Council ● (0.98)
Hereford &amp;	

Semantic Details

#	QID	Name	Type	Description	Score	Winner
1	Q1292592	Elmbridge Borough Council	non-metropolitan district, borough in the United Kingdom	local government district with borough status in Surrey, England	0.981	SELECTED
2	Q73072615	Elmbridge Borough Council	non-metropolitan district council	English non-metropolitan district council in Surrey, England, UK	0.964	SELECT
3	Q5366435	Elmbridge Borough Council election	class of election	class of election in the United Kingdom	0.040	SELECT

Figure 2: Koala-UI candidates details.

The frontend of Koala-UI was built using React[6] toensure a responsive and user-friendly interface. The design allows users to interact with the tool seamlessly across different platforms. Visual feedback, such as confidence scores and entity match suggestions, helps users efficiently assess and verify the quality of the entity linking process.

Koala-UI integrates two key external services within the frontend:

- *Alligator* [16], used for entity linking, accessed via its paginated API[7]. The API returns annotated

[6] https://react.dev

[7] https://github.com/unimib-datAI/alligator

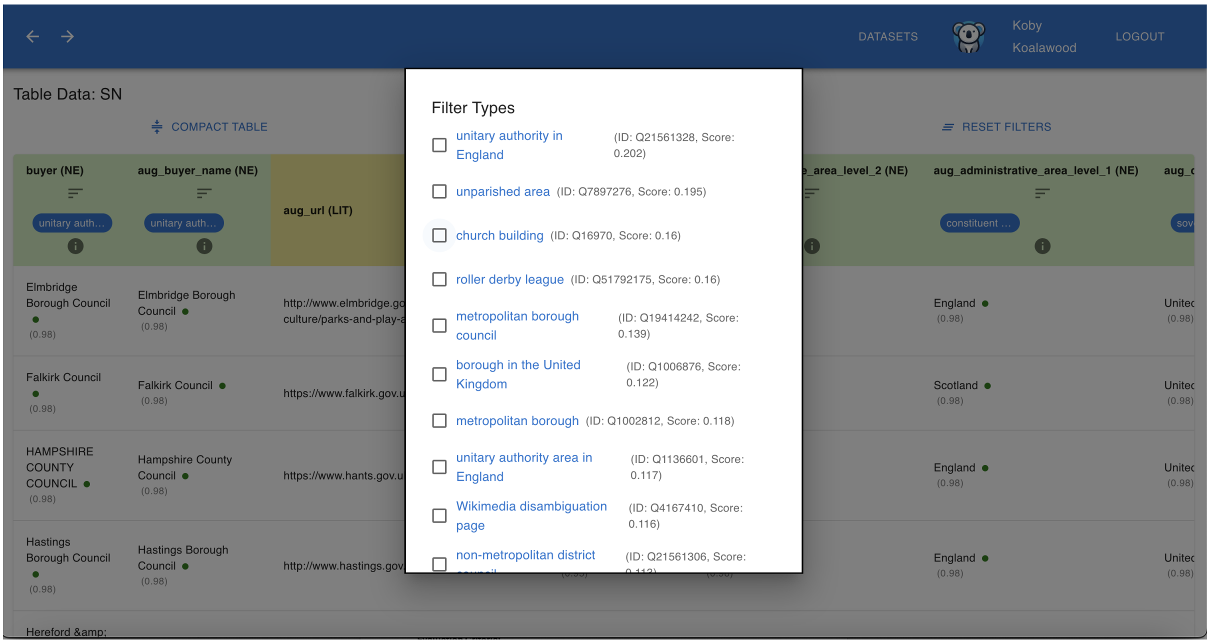


Table Data: SN

Filter Types

- □ unitary authority in England (ID: Q21561328, Score: 0.202)
- □ unparished area (ID: Q7897276, Score: 0.195)
- □ church building (ID: Q16970, Score: 0.16)
- □ roller derby league (ID: Q51792175, Score: 0.16)
- □ metropolitan borough council (ID: Q19414242, Score: 0.139)
- □ borough in the United Kingdom (ID: Q1006876, Score: 0.122)
- □ metropolitan borough (ID: Q1002812, Score: 0.118)
- □ unitary authority area in England (ID: Q1136601, Score: 0.117)
- □ Wikimedia disambiguation page (ID: Q4167410, Score: 0.116)
- □ non-metropolitan district council (ID: Q21561306, Score: 0.113)

buyer (NE) unitary auth…	aug_buyer_name (NE) unitary auth…	aug_url (LIT)	_area_level_2 (NE)	aug_administrative_area_level_1 (NE) constituent …	aug_c sov
Elmbridge Borough Council ● (0.98)	Elmbridge Borough Council ● (0.98)	http://www.elmbridge.go culture/parks-and-play-a		England ● (0.98)	Unitec (0.98)
Falkirk Council ● (0.98)	Falkirk Council ● (0.98)	https://www.falkirk.gov.u		Scotland ● (0.98)	Unitec (0.98)
HAMPSHIRE COUNTY COUNCIL ● (0.98)	Hampshire County Council ● (0.98)	https://www.hants.gov.u		England ● (0.98)	Unitec (0.98)
Hastings Borough Council ● (0.98)	Hastings Borough Council ● (0.98)	http://www.hastings.gov.		England ● (0.98)	Unitec (0.98)
Hereford &amp;					

Figure 3: Koala-UI types details.

tables as output, and data is processed incrementally using MongoDB. The paginated system allows Koala-UI to handle large datasets over time by breaking down responses into manageable parts, which can be processed one page at a time.

- *LAMAPI* [17] provides functionality for manually looking up entities by typing a string. This feature enables users to search for potential matches, offering greater flexibility when automated suggestions are insufficient.

Both services are called directly from the Koala-UI frontend, allowing it to handle entity linking and lookups without needing to process the data in its own backend.

The backend of Koala-UI is limited to managing user data. It supports the login system and stores data regarding different users, ensuring that user-specific settings and history are preserved across sessions.

The virtual machine (VM) running Alligator, which handles the bulk of the entity linking workload, was configured as follows:

- **Memory**: 30 GiB (16 GiB + 14 GiB DIMMs)
- **CPU**: AMD EPYC-Milan Processor, 8 cores, 2 GHz, 64-bit architecture
- **Firmware**: SeaBIOS, version 1.16.1-1.el9, 96KiB
- **Vendor**: KubeVirt (RHEL-9.2.0 PC)
- **Disk**: 944 GiB storage (EXT4 filesystem)

To handle large datasets efficiently, Alligator employs a paginated API system. This allows Koala-UI to request data incrementally, retrieving only a subset of the data per request. The following is an example of how paginated data for a table is retrieved via Alligator's API:

```
curl -X 'GET' 'https://alligator.hel.sintef.cloud/dataset/Spend%20Network/table/SN?page=1&per_page=10&token
     =alligator_demo_2023' \
-H 'accept: application/json'
```

In this request, data from the table `SN` of the `Spend Network` dataset is retrieved, with pagination parameters set to display 10 rows per page. The API responds with the table's data, including pagination details:

```
{
  "data": {
    "datasetName": "Spend Network",
    "tableName": "SN",
    "header": [
      "buyer",
      "aug_buyer_name",
      "aug_url",
      "aug_postal_town",
      "aug_administrative_area_level_2",
      "aug_administrative_area_level_1",
      "aug_country"
    ],
    "rows": [
      {
        "idRow": 26,
        "data": [...]
      },
      ...
    ]
  },
  "pagination": {
    "currentPage": 1,
    "perPage": 10,
    "totalPages": 9,
    "totalItems": 86
  }
}
```

The response provides the table's header, row data, and the pagination information. The pagination object specifies the current page (`currentPage`), the number of rows per page (`perPage`), the total number of pages (`totalPages`), and the total number of items in the dataset (`totalItems`). This structure allows Koala-UI to load and process the dataset in chunks, ensuring efficient handling of large datasets.

In addition to pagination, Koala-UI uses frontend techniques such as lazy loading and caching. These techniques ensure that the interface remains responsive, even when processing substantial volumes of information. The system's ability to handle large-scale data processing is largely driven by MongoDB's incremental processing model, which is leveraged through Alligator.

By combining a responsive frontend, a user management backend, and integration with external APIs such as Alligator and LAMAPI, Koala-UI is able to manage both user-specific data and large datasets effectively.

5. Use Case: Koala-UI applied on tenders data

Koala-UI was applied in a use case involving the organization Spend Network[8]. The objective of this use case was to create a European register of public entities, consolidating public sector entities from across Europe. This register is intended to serve as a canonical index for both public and private sector stakeholders, supporting applications such as compliance monitoring, Know Your Customer (KYC) processes, and facilitating cross-departmental collaborations.

Spend Network collects tender and contract data from over 700 sources globally, standardizing this data to the Open Contracting Data Standard (OCDS). The goal is to create a detailed European register of public entities. A major challenge was reconciling and validating inconsistent, incomplete, or ambiguous entity data. For instance, different names or abbreviations might refer to the same organization (e.g., ”CHU Rennes” and ”Centre Hospitalier Universitaire de Rennes”). Additionally,

[8] https://www.spendnetwork.com

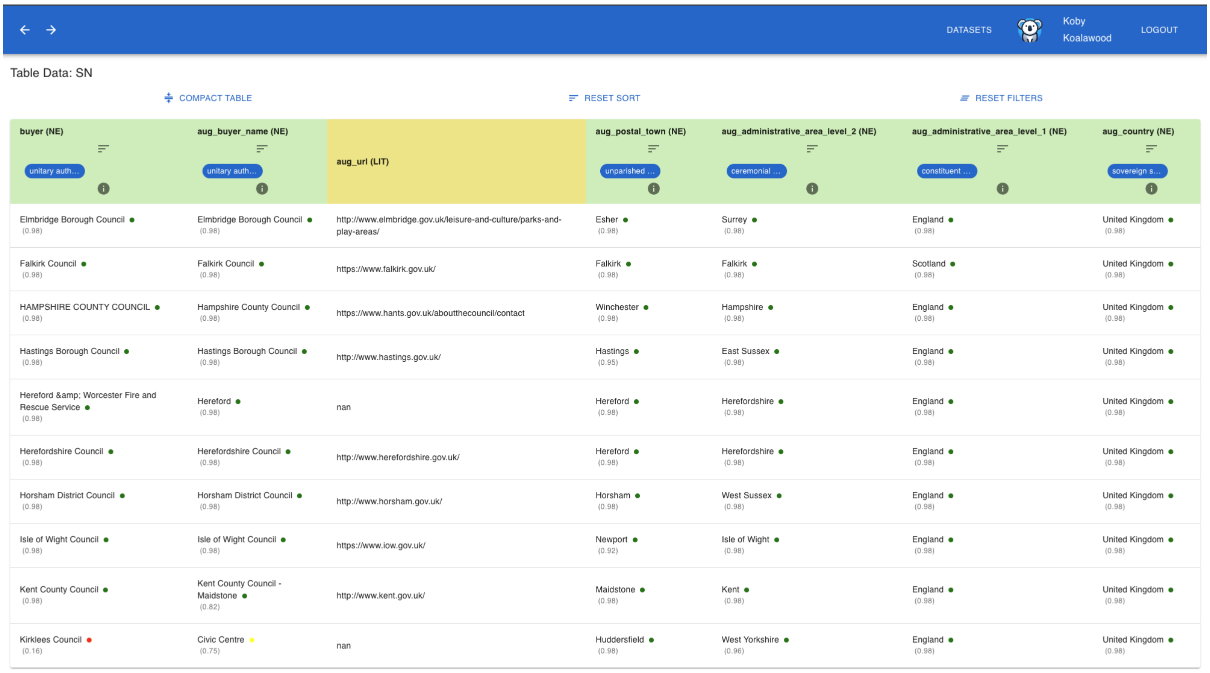


Table Data: SN

buyer (NE) unitary auth…	aug_buyer_name (NE) unitary auth…	aug_url (LIT)	aug_postal_town (NE) unparished …	aug_administrative_area_level_2 (NE) ceremonial …	aug_administrative_area_level_1 (NE) constituent …	aug_country (NE) sovereign s…
Elmbridge Borough Council ● (0.98)	Elmbridge Borough Council ● (0.98)	http://www.elmbridge.gov.uk/leisure-and-culture/parks-and-play-areas/	Esher ● (0.98)	Surrey ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Falkirk Council ● (0.98)	Falkirk Council ● (0.98)	https://www.falkirk.gov.uk/	Falkirk ● (0.98)	Falkirk ● (0.98)	Scotland ● (0.98)	United Kingdom ● (0.98)
HAMPSHIRE COUNTY COUNCIL ● (0.98)	Hampshire County Council ● (0.98)	https://www.hants.gov.uk/aboutthecouncil/contact	Winchester ● (0.98)	Hampshire ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Hastings Borough Council ● (0.98)	Hastings Borough Council ● (0.98)	http://www.hastings.gov.uk/	Hastings ● (0.95)	East Sussex ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Hereford &amp; Worcester Fire and Rescue Service ● (0.98)	Hereford ● (0.98)	nan	Hereford ● (0.98)	Herefordshire ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Herefordshire Council ● (0.98)	Herefordshire Council ● (0.98)	http://www.herefordshire.gov.uk/	Hereford ● (0.98)	Herefordshire ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Horsham District Council ● (0.98)	Horsham District Council ● (0.98)	http://www.horsham.gov.uk/	Horsham ● (0.98)	West Sussex ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Isle of Wight Council ● (0.98)	Isle of Wight Council ● (0.98)	https://www.iow.gov.uk/	Newport ● (0.92)	Isle of Wight ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Kent County Council ● (0.98)	Kent County Council - Maidstone ● (0.82)	http://www.kent.gov.uk/	Maidstone ● (0.98)	Kent ● (0.98)	England ● (0.98)	United Kingdom ● (0.98)
Kirklees Council ● (0.16)	Civic Centre ● (0.75)	nan	Huddersfield ● (0.98)	West Yorkshire ● (0.96)	England ● (0.98)	United Kingdom ● (0.98)

Figure 4: Koala-UI interface for entity linking in the spend network dataset.

organizations evolve over time, with entities merging, splitting or changing names, which complicates the entity reconciliation process.

Previously, Spend Network relied on manual, rule-based processes to manage this data, which limited the scalability and efficiency of the system. Expanding this register across Europe required a more automated approach, which Koala-UI provided through its entity linking capabilities. Koala-UI was used to reconcile entity names from tender data with external knowledge graphs such as Wikidata, helping to unify the various representations of the same entities. Its human-in-the-loop functionality allowed users to verify, review and adjust matches where necessary, ensuring a highe level of accuracy and consistency.

Using Koala-UI, Spend Network was able to:

- Reconcile entities like buyer names and URLs with knowledge graphs such as Wikidata.
- Enrich tender data with additional metadata sourced from knowledge graphs, such as geographic locations, administrative divisions, and other relevant data available in the external sources.
- Process large datasets efficiently, handling thousands of tenders from different countries.
- Leverage human-in-the-loop features to allow users to manually verify and adjust automated suggestions.

The application of Koala-UI improved the accuracy and scalability of the entity linking process, making the register more comprehensive and reliable for use in analytics and decision-making. Figure 4 shows an annotated table that comes from the Spend Network use case.

6. Summary and Outlook

Koala-UI was developed with the vision of evolving into a modern tool for tabular data linking, with a focus on simplicity, ease of use, and the ability to efficiently manage large datasets. The primary design goals were to create a user-friendly interface that simplifies complex tasks, making it accessible to a wide range of users while ensuring it can handle the demands of large-scale tabular data.

Currently, Koala-UI relies on the Alligator API for handling scalability and large datasets. However, looking ahead, the next phase of Koala-UI's development will involve transitioning from Alligator to a dedicated Python library. This library will be designed to handle big data directly, allowing for more flexibility and control over the data linking process. Once this Python library is developed, it will be integrated into a backend, providing its functionalities via a unified API.

The modular approach envisioned for the future will support the integration of multiple Python libraries, each representing different algorithms. This will allow users to select from traditional entity linking methods or more advanced algorithms, such as those based on large language models (LLMs). By wrapping these libraries into a single backend, Koala-UI will offer a versatile solution capable of adapting to various data linking needs and evolving technologies.

Additionally, improvements to the pagination system are planned. Currently, pagination relies on MongoDB's skip operator, which slows down when processing large tables. Future updates will address this by restricting navigation to one page at a time, enhancing performance with large datasets. New search functionalities will also be introduced, allowing users to match specific keywords across entire tables, thus enabling more efficient data exploration.

A critical objective for the future is to implement a complete data enrichment lifecycle within Koala-UI that incorporates human feedback at various stages. This iterative approach will enable continuous improvements in entity linking and data processing, ensuring that Koala-UI evolves in response to new challenges and opportunities in the field of data enrichment.

Acknowledgments

The work in this paper received partial funding from the projects enRichMyData (HE 101070284), Graph-Massivizer (HE 101093202) and UPCAST (HE 101093216).

References

[1] D. Stutz, J. T. de Assis, A. A. Laghari, A. A. Khan, N. Andreopoulos, A. Terziev, A. Deshpande, D. Kulkarni, E. G. Grata, Enhancing security in cloud computing using artificial intelligence (ai), Applying Artificial Intelligence in Cybersecurity Analytics and Cyber Threat Detection (2024) 179–220.

[2] A. Bécue, I. Praça, J. Gama, Artificial intelligence, cyber-threats and industry 4.0: Challenges and opportunities, Artificial Intelligence Review 54 (2021) 3849–3886.

[3] J. Zhao, X. Qu, Y. Wu, M. Fowler, A. F. Burke, Artificial intelligence-driven real-world battery diagnostics, Energy and AI 18 (2024) 100419.

[4] B. O. Antwi, B. O. Adelakun, A. O. Eziefule, Transforming financial reporting with ai: Enhancing accuracy and timeliness, International Journal of Advanced Economics 6 (2024) 205–223.

[5] A. Layegh, A. H. Payberah, A. Soylu, D. Roman, M. Matskin, Wiki-based prompts for enhancing relation extraction using language models, in: Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing, 2024, pp. 731–740.

[6] M. Mountantonakis, Y. Tzitzikas, Large-scale semantic integration of linked data: A survey, ACM Computing Surveys (CSUR) 52 (2019) 1–40.

[7] J. Hendler, Data integration for heterogenous datasets, Big data 2 (2014) 205–215.

[8] L. Wang, Heterogeneous data and big data analytics, Automatic Control and Information Sciences 3 (2017) 8–15.

[9] M. Ripamonti, F. De Paoli, M. Palmonari, Semtui: a framework for the interactive semantic enrichment of tabular data, arXiv preprint arXiv:2203.09521 (2022).

[10] I. Krasteva, D. Petrova-Antonova, F. De Paoli, E. Hristov, M. Borukova, M. Ciavotta, R. Avogadro, Geospatial enrichment of urban data for advanced city planning: a pilot study, in: 2023 IEEE International Conference on Big Data (BigData), IEEE, 2023, pp. 3139–3143.

[11] F. De Paoli, M. Ciavotta, R. Avogadro, E. Hristov, M. Borukova, D. Petrova-Antonova, I. Krasteva, An interactive approach to semantic enrichment with geospatial data, Data & Knowledge Engineering 153 (2024) 102341.

[12] K. Ham, Openrefine (version 2.5). http://openrefine. org. free, open-source tool for cleaning and transforming data, Journal of the Medical Library Association: JMLA 101 (2013) 233.

[13] T. F. Kusumasari, et al., Data profiling for data quality improvement with openrefine, in: 2016 international conference on information technology systems and innovation (ICITSI), IEEE, 2016, pp. 1–6.

[14] D. Roman, M. Dimitrov, N. Nikolov, A. Putlier, D. Sukhobok, B. Elvesæter, A. Berre, X. Ye, A. Simov, Y. Petkov, Datagraft: Simplifying open data publishing, in: The Semantic Web: ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29–June 2, 2016, Revised Selected Papers 13, Springer, 2016, pp. 101–106.

[15] D. Roman, N. Nikolov, A. Putlier, D. Sukhobok, B. Elvesæter, A. Berre, X. Ye, M. Dimitrov, A. Simov, M. Zarev, et al., Datagraft: One-stop-shop for open data management, Semantic Web 9 (2018) 393–411.

[16] R. Avogadro, M. Ciavotta, F. De Paoli, M. Palmonari, D. Roman, Estimating link confidence for human-in-the-loop table annotation, in: 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE, 2023, pp. 142–149.

[17] R. Avogadro, M. Cremaschi, F. D'Adda, F. De Paoli, M. Palmonari, et al., Lamapi: a comprehensive tool for string-based entity retrieval with type-base filters., in: OM@ ISWC, 2022, pp. 25–36.