Towards a Workflow Manager for Curation Technologies in the Legal Domain

Julian Moreno Schneider, Georg Rehm
Language Technology Lab, DFKI GmbH
Alt-Moabit 91c, 10559 Berlin, Germany
julian.moreno_schneider@dfki.de, georg.rehm@dfki.de

Abstract

We develop a system for the curation and further processing of documents from the legal domain. The platform is based on a legal knowledge graph. The overall project will result in three use-case-specific prototypes for different areas of the legal domain. For the purpose of designing the exact needs, demands, ideas, wishes and feature requests we currently collect the functional and non-functional requirements from the three use case partners. The objective of our work is the design and implementation of a generic, yet customisable, workflow management system for content and data curation services in the legal domain. In this article we describe and discuss how the inherent characteristics of a specific domain influence the design and development process of automatic workflows of text and data processing as well as curation components. Different techniques for the analysis and for collecting requirements are presented, followed by our survey and hybrid approach.

Keywords: curation technologies, requirements gathering, workflow definition

1. Introduction

European enterprises that operate internationally, especially small and medium-sized companies (SMEs), face multiple difficulties when attempting to offer and to market their products and services in other countries. Complying with regulatory and legal aspects is a hard challenge, which is usually delegated to law firms and consultancies. They have to identify, retrieve and process documents in multiple languages, from various sources and published by various institutions according to different criteria and formats. The further expansion and internationalisation of European SMEs is severely hindered by this situation. The potential of smart technologies to address the situation and to support these companies is enormous.

Current content and data analysis solutions are mature enough to be transferred to the market and to benefit from the new opportunities created by the Linked Data paradigm and the Open Data movement. Among these mature solutions are curation technologies, that enable and support the semantic analysis of documents with the help of automatic processes (Bourgonje et al., 2017) in order to extract information and to enrich single documents and whole document collections (Bourgonje et al., 2016). The goal is to make knowledge workers, who process and make use of these documents, more efficient and more effective in their day to day work, supporting them by delegating tasks that can be automised to the machine (summarisation, translation, report generation, named entity recognition, time expression analysis etc.) (Rehm et al., 2017a; Schneider et al., 2017; Rehm et al., 2017b; Rehm et al., 2018).

Documents that belong to the legal domain are highly interesting. Many types of legal documents exhibit rather fixed and clearly defined structures, which are typically considered an advantage when it comes to automatic analyses. Legal documents also contain multiple references to other documents, which make them difficult to read and fully comprehend. Often it is simply not feasible to read all documents referenced in a document. In addition to the high number of internal and external references, the ever-changing nature of law itself makes it important to have technologies that are capable of identifying these changes and reporting them whenever changes occur.

The objective of LYNX (Building the Legal Knowledge Graph for Smart Compliance Services in Multilingual Europe), a 36 months European Union project that started in December 2017, is the generation of a legal knowledge graph that contains different types of legal and regulatory data.[1] A set of advanced semantic services is currently under development to collect, to aggregate and to interlink data from heterogeneous sources and different jurisdictions, languages and orders. The project will eventually offer compliance-related functionalities that will be tested and validated in three use cases. The first pilot will be a legal compliance solution, where documents related to data protection are innovatively managed, analysed, and visualised across different jurisdictions. In the second pilot, LYNX will support the understanding of regulatory regimes, including norms and standards, related to energy operations. The third pilot will be a compliance solution in the domain of labour law, where legal provisions, case law, administrative resolutions, and expert literature will be interlinked, analysed, and compared to define legal strategies for legal practice.

In this article we describe the first steps towards the design and development of the underlying curation workflow manager by studying and analysing the requirements of the three pilots mentioned above. There are several important research questions that have to be answered to identify the needs of the three pilots:

1. Which are the specific needs of each use case?
2. Which datasets and common services are needed in each use case?
3. How can data and content be best organised and managed in the system so that the use case and corresponding pilot can be implemented?

[1] http://www.lynx-project.eu

4. Which services are needed? In which order and with which data and content will they be used?

5. What is the expected output of each use case?

In the EU project LYNX we apply curation technologies, applied to documents of several other domains in previous projects, to the legal domain. The main contribution of our work is a description of the first steps in the process of defining the workflows governing the curation processes, concretely, the requirements gathering process applied towards the definition of workflows in the legal domain.

The remainder of the article is structured as follows. Section 2. describes the concept of curation workflows and their application in the legal domain. Section 3. describes the solution for the requirements gathering and workflow definition processes. Section 4. concludes the article.

2. Curation Workflow Manager

A workflow is typically defined as "an orchestrated and repeatable pattern of business activity enabled by the systematic organization of resources into processes that transform materials, provide services, or process information" (BPM, 2009). This business activity is usually restricted to automatic processes. In our case we do not want to limit ourselves to automatic processes, but also include manual or intellectual activities carried out by experts. Therefore, the definition of workflows is not just a pipeline of automated components, but a complex structure or network of domain-specific steps.

For example, Figure 1 shows a workflow defined for the discovery and monitoring of legal documents. It is composed of seven tasks, of which four are automatic (blue boxes) and three are manual (orange boxes). The links between the tasks (their execution) depend on fulfilling the conditions established by the links themselves. Next to each task, a green box denotes the entities or roles involved in that task (human experts in case of manual tasks or systems in case of automatic ones).

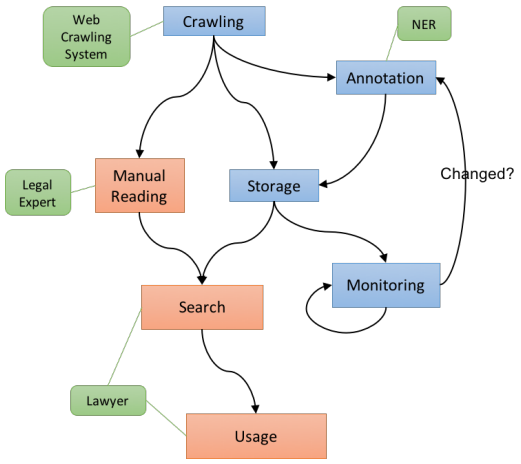

Figure 1: Workflow for the discovery and monitoring of legal documents inside an external database

This workflow is just an illustrative example, though, and we have only included the minimum tasks needed. A genuine workflow established in a company can be vastly more complex. We collaborate with three use case partner companies to get a better insight as to which are the specific workflows they currently use and follow and the requirements they have. Only with sufficient knowledge and insights into their processes and workflows can we begin the design of the curation workflow manager including its requirements and functionalities.

To provide a few other examples regarding the curation of digital documents and content in other domains, experts in a digital agency build mobile apps or websites for clients who provide the digital agency with documents, data, pictures, videos and other assets that are processed, sorted, augmented, arranged, packaged and then deployed. Journalists need to stay on top of the news stream including blogs, microblogs, newswires, websites etc. in order to produce a new article on a breaking topic, based on the information they collected, processed, sorted, evaluated, verified and synthesised (Rehm et al., 2017a).

The main tasks, apart from obtaining, analysing and organising documents, in data protection compliance are searching, browsing and commenting on documents. In regulatory compliance, companies have to support the current audit, verification and certification (including classification) of documents, apart from summarising larger documents and enabling commenting on them. Translating documents from other jurisdictions and comparing them among jurisdictions is also a common task. A labour law expert needs to access, aggregate and interlink relevant legal information, which starts by looking for relevant documents, manually creating links between legal provisions, case law, administrative resolutions and expert literature even across different jurisdictions, identifing relevant documents that may affect the case and tracing their changes, through the life cycle of the case.

The research and innovation project LYNX is currently in its first steps. Within LYNX, we design, define and develop a set of workflows for fulfilling the needed functionalities in the pilot use cases. The Curation Workflow Manager will be defined, including its main functionality to arrange, organise, orchestrate and combine all components in the platform in order to generate suitable workflows for every use case. The Curation Workflow Manager is located in between the pilot use cases and the common services.

Due to the fact that the platform will be based on a flexible service-oriented architecture, in which the basic services form a set of single services and data endpoints, the workflows act like a semantic middleware that integrates the core services in the use cases. This approach has the advantage of clearly separating the development of the distributed components of services from the actual use cases so that the different partners can develop their technologies at the same time without compromising the concurrent development work of other partners (as long as the agreed upon REST APIs remain stable and unchanged).

Workflows are sets of tasks (both manual and automatic) that are interconnected to complete a larger, more complex task. Among the automatic services that will become avail-

able during the course of the project and that will be used in the definition of the curation workflows are: named entity recognition, terminology identification, geolocation annotation, time expression analysis, translation, summarisation, semantic annotation, information extraction, linking and mapping, smart search, recommendation, alerts.
A comprehensive introduction to workflow management is offered by (Van Der Aalst and Van Hee, 2004), who present a basic overview of workflow terminology and organisation. The state of the art is presented by (Unertl, Kim M and Novak, Laurie L and Johnson, Kevin B and Lorenzi, Nancy M, 2010). Sequences of tasks and whole workflows are defined on a regular basis in many domains, which is why examples can be found in many different areas, among others, biomedical (BioNLP UIMA Component Repository (Baumgartner et al., 2008), JULIE Lab's UIMA Component Repository (JCoRe) (Hahn et al., 2008), Smntx (Chard et al., 2011), (Rak et al., 2012) or (Köster and Rahmann, 2012)), software development (Apache Oozie (Islam et al., 2012)) and in NLP, where many different frameworks for the definition of workflows exist: Taverna (Hull et al., 2006), Galaxy (Blankenberg et al., 2010), GATE (General Architecture for Text Engineering) (Cunningham et al., 2002), DKPro Core (de Castilho and Gurevych, 2014), U-Compare (Kano et al., 2009; Kano et al., 2011) and TextFlows (Perovšek et al., 2016).

3. Design, Definition, Development

The Curation Workflow Manager (CWM) is a component of the architecture that is responsible for orchestrating and managing the workflows and adapting them to the specific needs of the three use cases. In this project, the users of the system are the actual use case partners including, if applicable, their clients and other immediate stakeholders. Therefore, we need to collect and specify the workflows that are currently used and that are to be used in the future, given the new LYNX functionalities, and how we can realise them using the above mentioned semantic services.
The first step towards the design and definition of the CWM is a list of requirements obtained from the users. To that end, many different techniques can be applied. Some examples are: Document Analysis (evaluating the documentation of a present system), Feasibility Study (studying existing systems and the possibility of replacing them), Interview (with one more future users), Observation (studying users in their workplace), Prototyping (gathering preliminary requirements to build an initial prototype), Survey/Questionnaires (gathering information from a small or large amount of users), Brainstorming (identifying all possible solutions to problems) and Requirements Workshop (more organised and structured than a brainstorming session). A complete description of requirement gathering techniques can be found in (Fricker et al., 2015).
Based on the existing and commonly used techniques, and also taking the typical constraints of a research project into account, we opt for a hybrid approach that consists of the following steps:

- First we define a survey the main objective of which is to collect a first set of requirements, needs, ideas and visions the use case partners have.
- Second, we use the results obtained from the survey to design a first, still coarse-grained specifications of the workflows for each of the three use cases.
- Based on the three sets of coarse-grained specifications we plan several brainstorming workshops, in which we will collect the requirements on a much more detailed level from the partners. There are several options how to organise these workshops. One fundamental distinction relates to the question if a new, even conceptually, GUI needs to be implemented or if the semantic services are to be integrated in existing systems and GUIs. In addition, if the survey results from two or maybe all three use case partners are similar, there may be no need for bilateral workshops; this result would also be indicative of the emergence of a shared mutual understanding of an application type, which could be called, for example, "legal data and content curation system".
- Finally, the results obtained in the one, two or three workshops will be translated into requirements for the design and implementation of the Curation Workflow Manager.

Such a user-centred design approach allows the inclusion of the users (in our case, three use case partners) in the requirements gathering process, because they contribute to the initial definition of requirements through the survey, and also in its concretization through the workshops.

3.1. Survey

The first step of the requirements gathering process is a survey that will help us to define the general needs of the use case partners (and their clients). The survey is divided into several parts described in the next sections.

3.1.1. Non-functional Requirements

The non-functional requirements part of the survey has the goal to sketch the most general and abstract needs the pilot use case partners have in relation to the project and the overall platform (see Question 1. in Section 1.).

1. Please describe, as specifically as possible, your use case (or use cases): what kind of functionality or processing capabilities do you want to realise or achieve with the help of the Lynx platform?
2. What kind of devices do you work with predominantly? (Desktops/laptops, touch-interface devices, speech interfaces etc.)
3. Do you plan to integrate the Lynx platform into existing in-house systems and graphical user interfaces (GUIs)?

If the answer to question 3 is "Yes", please also reply to questions 4 and 5:

4. Please specify the system into which you want to integrate Lynx. Please provide screenshots or screencasts of the system.

5. Do you currently use a stand-alone application with a GUI or web-based GUI?

If the answer to question 3 is "No", please also reply to questions 6 and 7:

6. How are you planning to use the services developed in Lynx? (REST API calls, Web services, Web browser, Mobile phone/tablet applications, Other)

7. Would your preference be to develop a new (web-based) GUI to connect to the Lynx services or would you prefer some other way?

3.1.2. Actual Usage of Automatic Processing

This part of the survey is intended to analyse the current usage of automatic processing techniques and tools inside the use case partners environments and their customers (legal firms).

8. How do you analyse or process legal documents in your company? (For example, with the help of human experts, fully automatically, semi-automatically etc.? Please be as specific and descriptive as possible.)

9. Do you use automatic solutions and tools for analysing and processing legal documents in your company? If yes, which ones?

10. What kind of documents from the legal domain (or your use case domain) do you work with (official law texts, letters, case law, EU regulations and directives, client specifications etc.)?

11. If you already use software for processing legal documents, please provide screenshots or screencasts of your software/GUIs.

12. In terms of use cases and workflows, please specify all (or a representative set of) typical workflows that you use in-house (e. g., types of documents, types of analysis, types of processing, types or approaches of producing new content, etc.).

The questions in this part of the survey are rather abstract and general, because we need to get an overview of the workflows that the use case partners currently use, without paying too much attention to the implementation or concrete details (with which we will deal in later steps in the project).

3.1.3. Users and Profiles

Although this part of the survey does not have many questions, they are important for the development of the Curation Workflow Manager, because depending on the amount and type of users that can use the platform (workflows) the whole management implementation has to be adapted.

13. What types of users are going to use Lynx services (e. g., JavaScript developers, lawyers, knowledge workers, customers, etc.)?

14. Do you need a multi-user solution?

15. Do you need authentication (login/password)?

16. Do you need access control lists with different roles and different permissions?

3.1.4. Data Sets

This part of the survey is more concrete and tries to get a better understanding about the concrete datasets needs in every use case (addressing Questions 2. and 4. in Section 1.). The idea is to determine which datasets are needed and in which format for every use case.

17. What kind of reference materials or reference data sets do you use on a regular basis?

18. Which online data sets or reference materials would help you in your daily work?

19. File Formats: Which are the formats of files that you want to process with Lynx? Do you want the same file format in the request you send to Lynx as well as in the responses you get back from Lynx?

3.1.5. Common Services

This part of the survey is more concrete and tries to get a better understanding about the concrete services needs in every use case. The idea is to determine those services that are needed overall and those that are specific to only one concrete case. This section addresses Questions 2., 3. and 4. from Section 1..

20. Do you need a tool that can identify and highlight named entities (persons, locations, organizations, etc.) in legal documents? For example, this could result in a colour-based highlighting of person, location, organisation names in documents or the filtering of document collections based on the names contained in them.

21. Do you need a tool that can identify and highlight time expressions and normalize them? Such a function could enable a timeline view of a large document collection, for example, of a series of letters or correspondence.

22. Do you need a tool that can identify and highlight geographical information related to locations in legal documents? For example, the output of such a function could be an interactive map containing all documents or content of the documents.

23. Do you need a tool that can identify and highlight events (or other types of important keywords) in legal documents? For example, the output of such a function could be a list of events (words, phrases, expressions, etc.) that require some kind of action or reaction from the reader.

24. Do you need a tool that can identify relations between entities (some judge is related to a criminal because they are involved in a court case) in legal documents? For example, the output of such a function could result in capabilities for searching documents containing relations through certain entities.

25. Do you need a tool that can identify specific domain terminology (legal terms, oil & gas related terms, etc.) in legal documents?

26. Do you need a tool that can recognize citations, references and relations between legal documents? For example, the output of such a function could be an interactive graph display showing the relations between all the documents of a court case or piece of legislation.

27. Do you need a tool that can disambiguate the sense of a term determining if it is referring to labour law (as an example) or any other domain in legal documents? For example, the output of such a function could be used for better determining concrete topics the document is talking about.

28. Do you need a tool that can translate legal documents to other languages (if yes, which languages and language pairs?)?

29. Do you need a tool that can summarise documents or sets of documents in the legal domain?

30. Do you need a tool that can search through collections of legal documents?

31. Do you need a tool that can recommend other legal documents related to a certain task?

32. Do you need a tool that can alert you about changes in existing legal documents or the appearance of new legal documents?

33. Do you need a tool that can determine the main topic of a legal document or part of a document (paragraphs, etc.)? For example, the output of such a function could help in searching documents for certain legislations.

34. Do you need a tool that can determine the main type of a legal document (e. g., letter, law, contract, technical report, case report etc.)? For example, the output of such a function could help further process and visualise a large and heterogeneous set of documents.

35. Do you want to combine several automatic processing steps? For example: When you get a document, the first thing you do is to translate (if it is in a language other than English), then you read it to learn which people are mentioned (locations and time expressions are also important but first are people). After that you focus on the references of other laws and finally you try to identify arguments and events.

3.1.6. Additional Requirements

The last part of the survey is an open question for including any information that is missing in the previous questions and that the use case partners what to include.

36. Please write down any additional requirements you may have that are not covered by the questions above.

3.2. Workshops

Once the survey has been circulated to the pilot use case partners and they have filled it in, we will analyze them to define the first sets of requirements. With these we will be in a good position to define and plan the workshops in which we will concretize (clean and filter) the requirements with the partners.

The development of the workshops depend directly on the results obtained from the surveys. If the implementation of a completely new and redesigned GUI is considered important (through Question 5), the workshop will be established as a graphical design workshop where the main focus will be put on the generation of mockups and wireframes of the new interface. Here, the output of the workshop will not only be a list of requirements, but also a set of mockups of the new interface. Depending on the results of the survey, it has to be decided if there will be only one workshop with the three use case partners, developing a common GUI as well as individual solutions, or if it is better to have independent workshops, one with each use case partner and then extrapolating a common interface. If the surveys reveal, on the other hand, that an integration of the new components, services and workflows in existing systems is needed, three individual workshops will be organised. These will be focused on the study and analysis of the currently used technologies and how the users interact with them, as well as how the future workflows can be integrated into the current processes.

4. Summary and Conclusion

We define the concept of the Curation Workflow Manager (CWM), which refers to the management of specifying different curation workflows in an adaptive platform. The definition of workflows is a complex task that requires a close collaboration among all involved stakeholders by means of requirements gathering processes. We apply a hybrid approach that consists of a requirements gathering survey, together with face-to-face workshops with the pilot use case partners. The survey has been design with the main goal of gathering concrete information about the three pilot use cases in LYNX. The first part collects general information about the intended use of LYNX technologies. The second part is designed to learn more details about the use case partners' current workflows, in addition to determining if automatic processes are used. The survey also includes questions on the intended users of the system and regarding the functional requirements. The last part collects information on the necessary infrastructure for each of the use cases. We are still in the process of design the Curation Workflow Manager. This paper includes the final version of the survey. First results of the requirements gathering phase will be presented at the workshop and in follow-up publications.

5. Acknowledgements

The project LYNX has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 780602. More information is available online at http://www.lynx-project.eu.

6. Bibliographical References

Baumgartner, W. A., Cohen, K. B., and Hunter, L. (2008). An open-source framework for large-scale, flexible evaluation of biomedical text mining systems. *Journal of biomedical discovery and collaboration*, 3(1):1.

Blankenberg, D., Kuster, G. V., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. *Current protocols in molecular biology*, pages 19–10.

Bourgonje, P., Schneider, J. M., Rehm, G., and Sasaki, F. (2016). Processing Document Collections to Automatically Extract Linked Data: Semantic Storytelling Technologies for Smart Curation Workflows. In Aldo Gangemi et al., editors, *Proceedings of the 2nd International Workshop on Natural Language Generation and the Semantic Web (WebNLG 2016)*, pages 13–16, Edinburgh, UK, September. The Association for Computational Linguistics.

Bourgonje, P., Schneider, J. M., and Rehm, G. (2017). Domain-specific Entity Spotting: Curation Technologies for Digital Humanities and Text Analytics. In Nils Reiter et al., editors, *CUTE Workshop 2017 – CRETA Unshared Task zu Entitätenreferenzen. Workshop bei DHd2017*, Berne, Switzerland, February.

BPM, B. P. M. (2009). Glossary. Technical report, Center of Excellence (CoE).

Chard, K., Russell, M., Lussier, Y. A., Mendonça, E. A., and Silverstein, J. C. (2011). A cloud-based approach to medical NLP. In *AMIA Annual Symposium proceedings*, volume 2011, page 207. American Medical Informatics Association.

Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V. (2002). A framework and graphical development environment for robust NLP tools and applications. In *ACL*, pages 168–175.

de Castilho, R. E. and Gurevych, I. (2014). A broad-coverage collection of portable NLP components for building shareable analysis pipelines. In *Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT*, pages 1–11.

Fricker, S. A., Grau, R., and Zwingli, A., (2015). *Requirements Engineering: Best Practice*, pages 25–46. Springer International Publishing, Cham.

Hahn, U., Buyko, E., Landefeld, R., et al. (2008). Language Resources and Evaluation Workshop, Towards Enhanced Interoperability Large HLT System: UIMA NLP. *Marrakech, Morocco*, pages 1–8.

Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M. R., Li, P., and Oinn, T. (2006). Taverna: a tool for building and running workflows of services. *Nucleic acids research*, 34(suppl_2):W729–W732.

Islam, M., Huang, A. K., Battisha, M., Chiang, M., Srinivasan, S., Peters, C., Neumann, A., and Abdelnur, A. (2012). Oozie: Towards a Scalable Workflow Management System for Hadoop. In *Proceedings of the 1st ACM SIGMOD Workshop on Scalable Workflow Execution Engines and Technologies*, SWEET '12, pages 4:1–4:10, New York, NY, USA. ACM.

Kano, Y., Baumgartner Jr, W. A., McCrohon, L., Ananiadou, S., Cohen, K. B., Hunter, L., and Tsujii, J. (2009). U-Compare: share and compare text mining tools with UIMA. *Bioinformatics*, 25(15):1997–1998.

Kano, Y., Miwa, M., Cohen, K. B., Hunter, L. E., Ananiadou, S., and Tsujii, J. (2011). U-Compare: A modular NLP workflow construction and evaluation system. *IBM Journal of Research and Development*, 55(3):11–1.

Köster, J. and Rahmann, S. (2012). Snakemake – a scalable bioinformatics workflow engine. *Bioinformatics*, 28(19):2520–2522.

Perovšek, M., Kranjc, J., Erjavec, T., Cestnik, B., and Lavrač, N. (2016). TextFlows: A visual programming platform for text mining and natural language processing. *Science of Computer Programming*, 121:128–152.

Rak, R., Rowley, A., Black, W., and Ananiadou, S. (2012). Argo: an integrative, interactive, text mining-based workbench supporting curation. *Database*, 2012:bas010.

Rehm, G., He, J., Schneider, J. M., Nehring, J., and Quantz, J. (2017a). Designing User Interfaces for Curation Technologies. In Sakae Yamamoto, editor, *Human Interface and the Management of Information: Information, Knowledge and Interaction Design, 19th International Conference, HCI International 2017*, number 10273 in Lecture Notes in Computer Science (LNCS), pages 388–406, Vancouver, Canada, July. Springer. Part I.

Rehm, G., Schneider, J. M., Bourgonje, P., Srivastava, A., Nehring, J., Berger, A., König, L., Räuchle, S., and Gerth, J. (2017b). Event Detection and Semantic Storytelling: Generating a Travelogue from a large Collection of Personal Letters. In Tommaso Caselli, et al., editors, *Proceedings of the Events and Stories in the News Workshop*, pages 42–51, Vancouver, Canada, August. Association for Computational Linguistics. Co-located with ACL 2017.

Rehm, G., Schneider, J. M., Bourgonje, P., Srivastava, A., Fricke, R., Thomsen, J., He, J., Quantz, J., Berger, A., König, L., Räuchle, S., Gerth, J., and Wabnitz, D. (2018). Different Types of Automated and Semi-Automated Semantic Storytelling: Curation Technologies for Different Sectors. In Georg Rehm et al., editors, *Language Technologies for the Challenges of the Digital Age: 27th International Conference, GSCL 2017, Berlin, Germany, September 13-14, 2017, Proceedings*, number 10713 in Lecture Notes in Artificial Intelligence (LNAI), pages 232–247, Cham, Switzerland, January. Gesellschaft für Sprachtechnologie und Computerlinguistik e.V., Springer. 13/14 September 2017.

Schneider, J. M., Bourgonje, P., and Rehm, G. (2017). Towards User Interfaces for Semantic Storytelling. In Sakae Yamamoto, editor, *Human Interface and the Management of Information: Information, Knowledge and Interaction Design, 19th International Conference, HCI International 2017*, number 10274 in Lecture Notes in Computer Science (LNCS), pages 403–421, Vancouver, Canada, July. Springer. Part II.

Unertl, Kim M and Novak, Laurie L and Johnson, Kevin B and Lorenzi, Nancy M. (2010). Traversing the many paths of workflow research: developing a conceptual framework of workflow terminology through a systematic literature review. *Journal of the American Medical Informatics Association*, 17(3):265–273.

Van Der Aalst, W. and Van Hee, K. M. (2004). *Workflow management: models, methods, and systems*. MIT press.