

Semantic Storytelling: Towards Identifying Storylines in Large Amounts of Text Content

Georg Rehm, **Karolina Zaczynska**, Julián Moreno-Schneider

karolina.zaczynska@dfki.de

DFKI GmbH, Berlin

Text2Story, ECIR Workshop – Cologne, Germany, April 14th, 2019

Overview

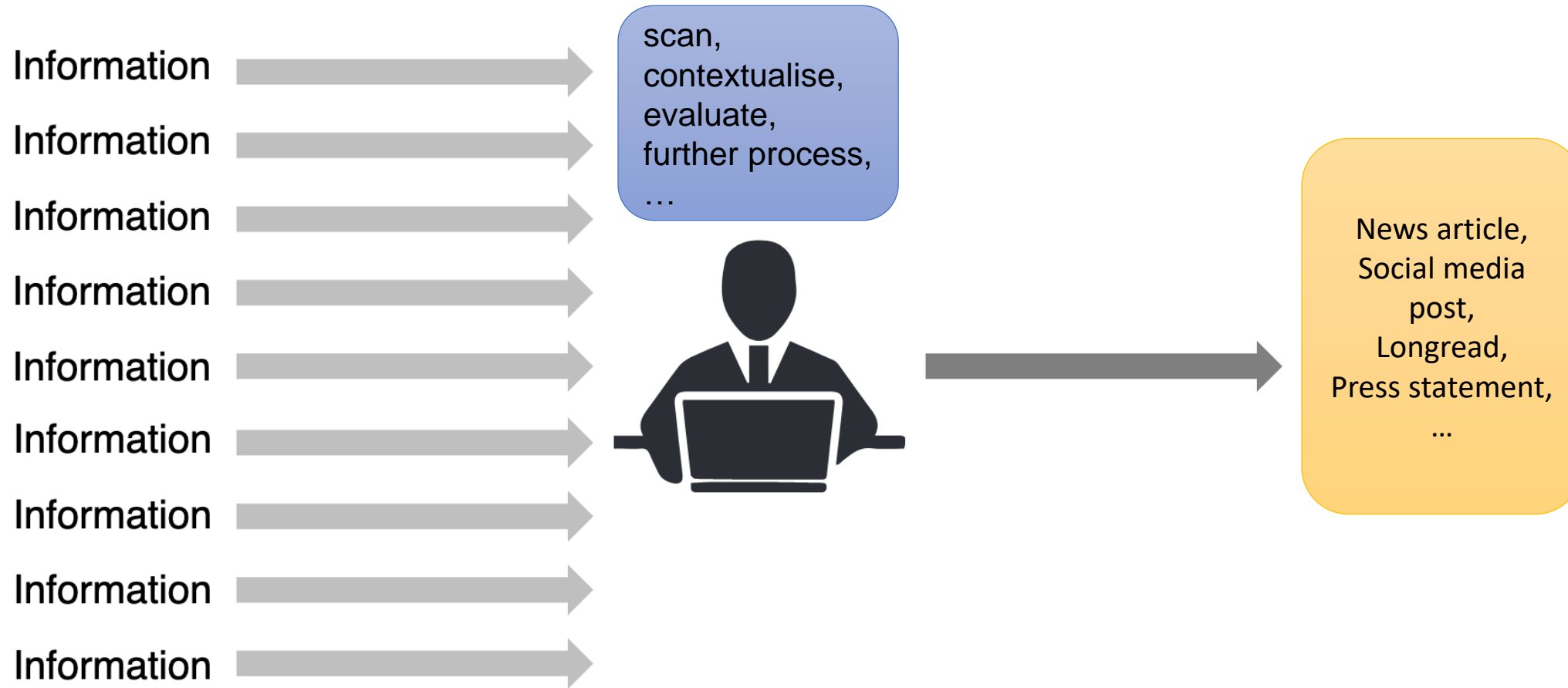
1. Introduction
2. Towards Semantic Storytelling
3. Architecture and Components
4. Use Case: Mendelsohn Letters
5. Conclusions and Future Work

Introduction

There are more and more knowledge workers, who's daily job is:

- to analyse or make use of incoming digital content
- to understand the, in a wider sense, storyline of an event or series of events
- to produce new piece of information

Knowledge worker example:



Introduction

Challenges:

- Increasing amount of online information
- Still need to do a great amount of work manually
- Available tools are often restricted to specific tasks

→ Big need for tools that support knowledge workers in **producing new storylines**

2. Towards Semantic Storytelling

What is Storytelling and how can we support it?

2. Towards Semantic Storytelling

- Storytelling = human technique to order a series of events in the world and find meaningful patterns in them
- Relate events into a schematic structure (organized by topics, locality or causal relationships) and construct explanatory models of the world and events

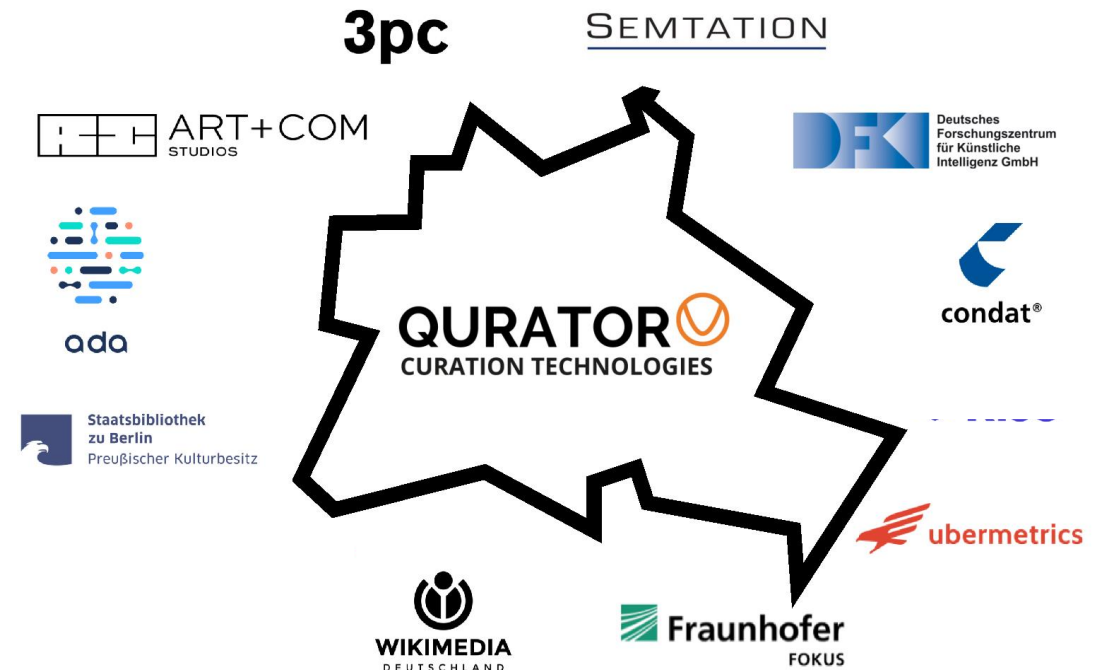
2. Towards Semantic Storytelling

Semantic Storytelling can be seen as the attempt to translate the theories of storytelling into a formal, and machine-processable scheme

→ Support knowledge workers so they can create new stories based on the content presented by our tools

Project **QURATOR (2018-2021):** Create an ecosystem for curating technologies

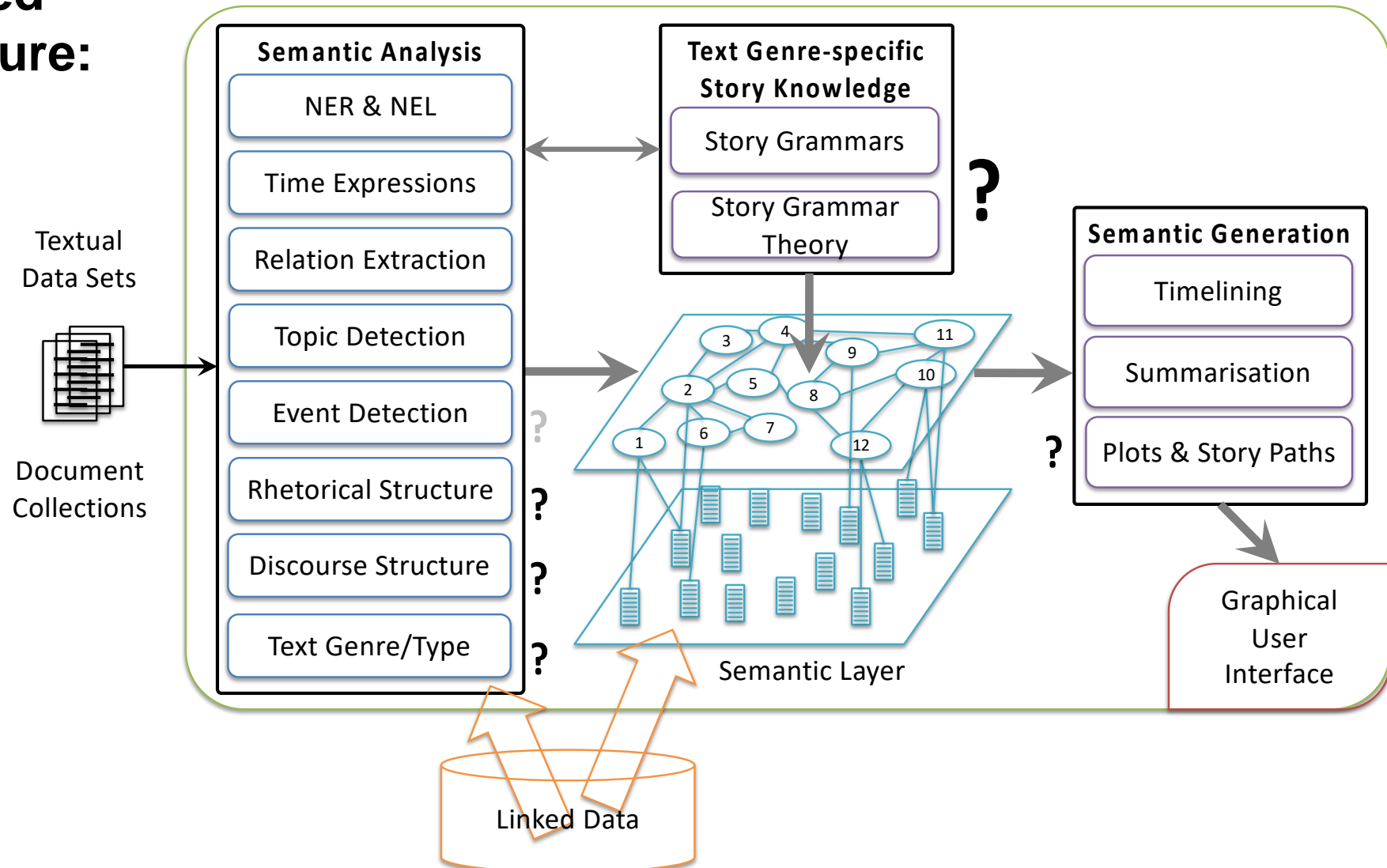
- 10 industry and research partners
- Create smart AI supported tools & architecture to support knowledge workers



3. Architecture and Components

One possible architecture for an Semantic Storytelling Platform:

Suggested architecture:



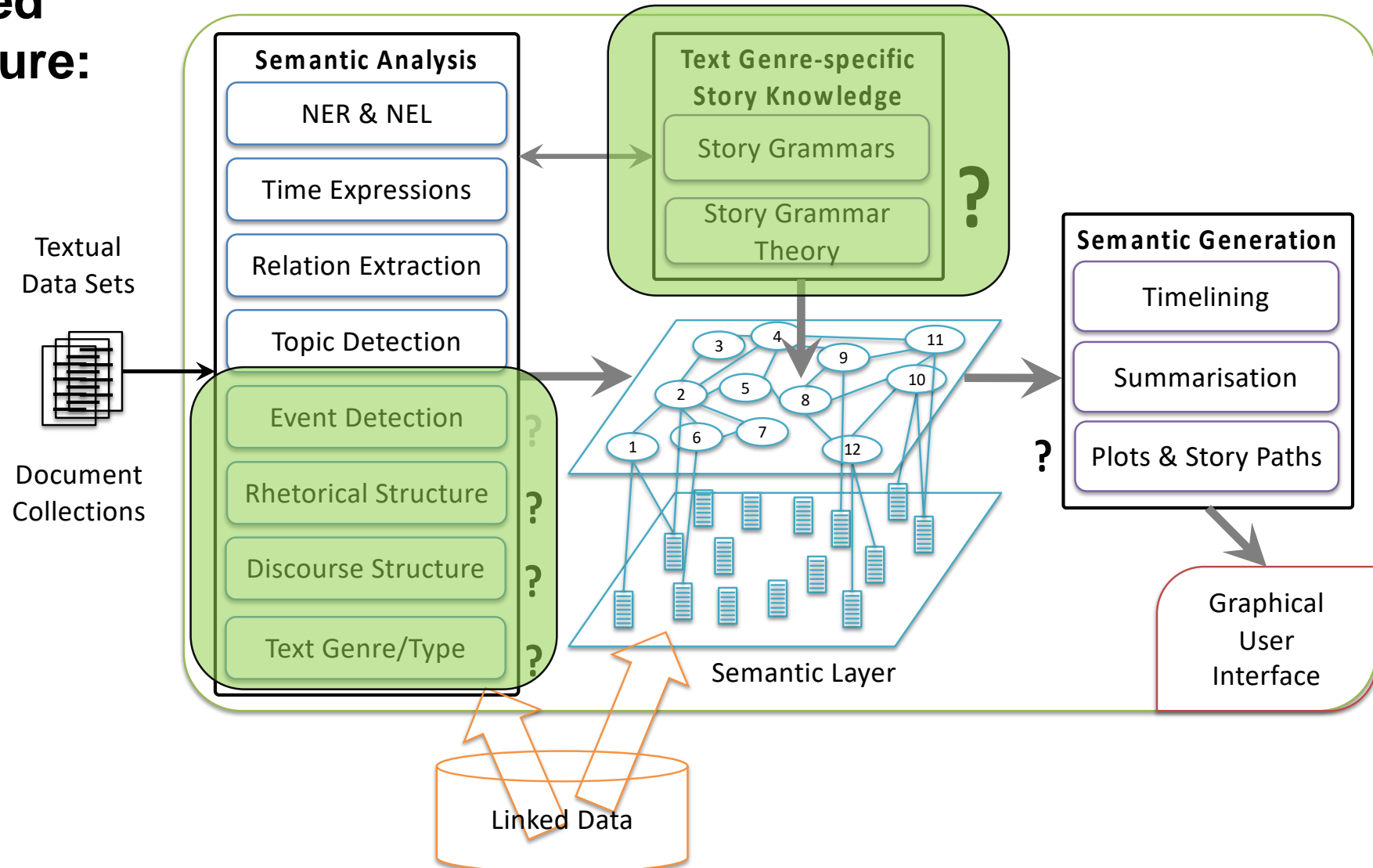
3. Architecture and Components

What works well?

- Named Entity Recognition and Linking
- Time Expression Analysis
- Topic Detection
- Text Classification

→ Have been used in many different applications for years

Suggested architecture:



3. Architecture and Components:

Generic Text Genre processing

- Plan: Use knowledge of text genres
- For advanced text and discourse informed processing optimized by the characteristics for each respective text genre

3. Architecture and Components:

Generic Text Genre processing

- **Text Genre:** represent main communicative aims with standardized lexicogrammatical relations
 - News – inform about an event
 - Journal articles – discuss a scientific claim
 - ...
- Extract more purposeful most important informative parts in texts
- Support more focused storyline-production

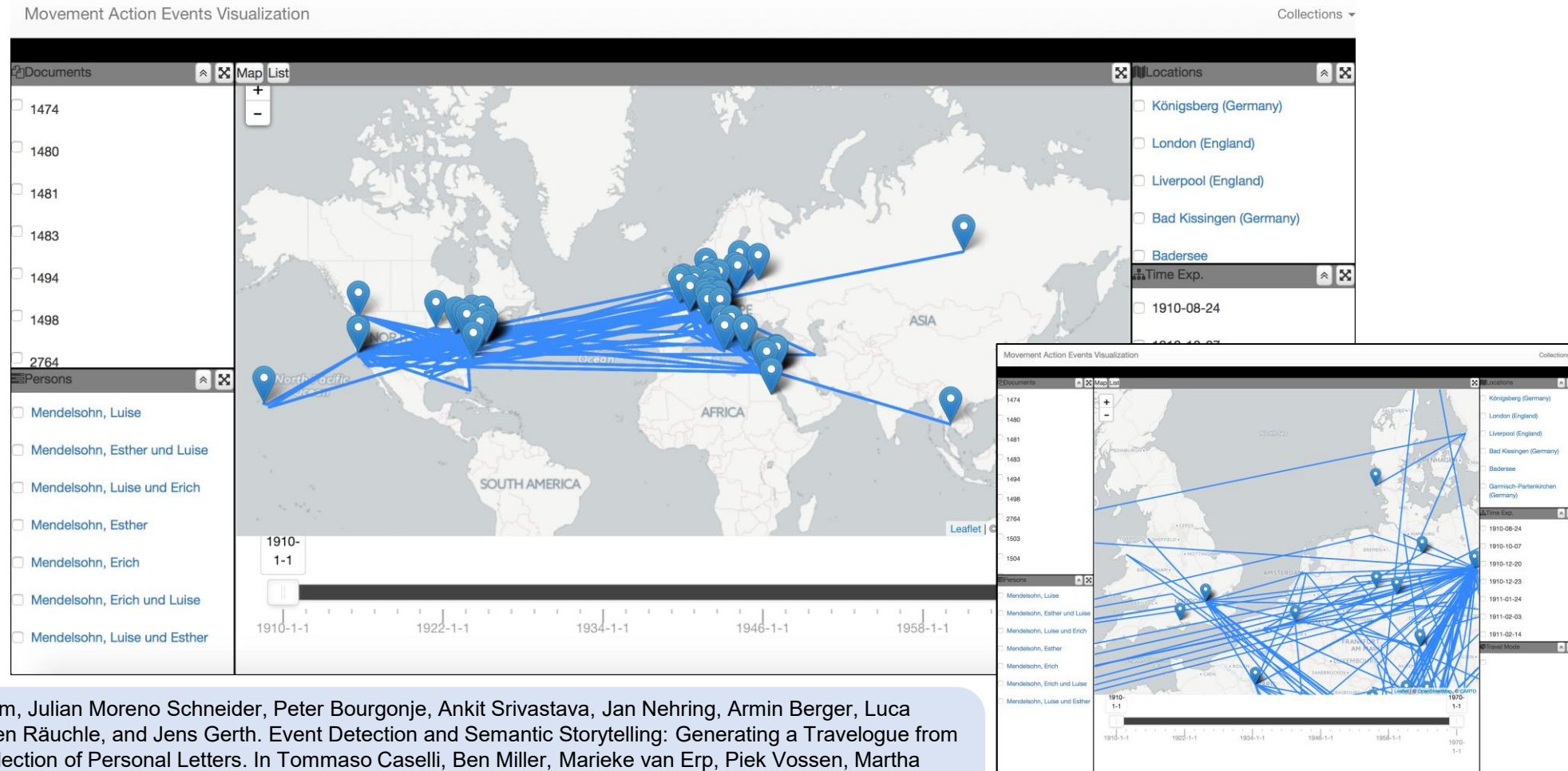
TIMELINE-BASED EVENT ORDERING

CLAIM EXTRACTION

4. Experiments: Mendelsohn Letters

- First experiments from previous project DKT
Digitale Kuratierungstechnologien,
Project 2015-2017 funded by German Federal Ministry of Education and Research
- Input: 2796 letter correspondences between Luise and Erich Mendelsohn (2796, from 1910 to 1953)
- Extracted **Movement Action Events (MAE)** from a large collection of personal letters

Experimental Storytelling Dashboard based on MAE:



Georg Rehm, Julian Moreno Schneider, Peter Bourgonje, Ankit Srivastava, Jan Nehring, Armin Berger, Luca König, Sören Räuchle, and Jens Gerth. Event Detection and Semantic Storytelling: Generating a Travelogue from a large Collection of Personal Letters. In Tommaso Caselli, Ben Miller, Marieke van Erp, Piek Vossen, Martha Palmer, Eduard Hovy, and Teruko Mitamura, editors, Proceedings of the Events and Stories in the News Workshop, Vancouver, Canada, August 2017. Association for Computational Linguistics. Co-located with ACL 2017.

4. Use Case: Mendelsohn Letters

- 3pc created a mock-up for a possible storytelling tool
= authoring tool for knowledge workers (editors, authors, ...)

4. Use case: Mendelsohn Letters

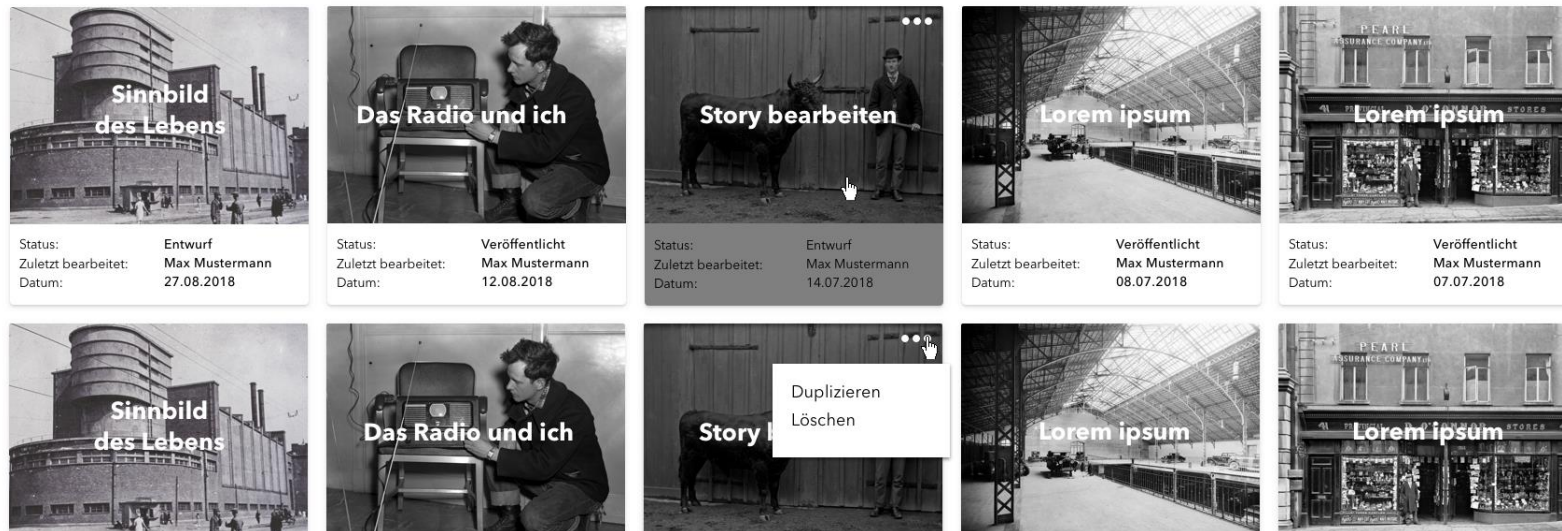
Redaktionstool

Max Mustermann
Redakteur

Neue Story



Alle / Entwurf / Veröffentlicht



4. Use case: Mendelsohn Letters

Redaktionstool

Max Mustermann
Redakteur

Neue Story

Leer

Alle / Entwurf / Veröffentlicht

Sinnbild
des Lebens

Status: Entwurf
Zuletzt bearbeitet: Max Mustermann
Datum: 27.08.2018

Das Leben des Erich Mendelsohn

Beschreibung

Eine Biografie geschrieben von Jeremy Larson und Kathleen Montgomery

Tags

Erich Mendelsohn

Story anlegen

Veröffentlicht
Max Mustermann
07.07.2018

Sinnbild
des Lebens

Das Radio und ich

Lorem ipsum

Lorem ipsum

Lorem ipsum

4. Use case: Mendelsohn Letters

←

Das Leben des Erich Mendelsohn

Eine Biografie geschrieben von Jeremy Larson und Kathleen Montgomery

Letzte Änderung vor 2 Minuten

★

Vorschläge

🗪

Stories

🔖

Merkliste

🔍 Archiv durchsuchen

A

🖼️

🎬

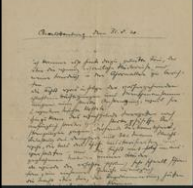
+

Drag & Drop!

+


−

Bilder Videos Audio Dokumente




Brief No. 1 an Luise Mendelsohn

Dokument




Portrait Erich Mendelsohn

Bild



Portrait Erich und Luise Mendelsohn

Bild

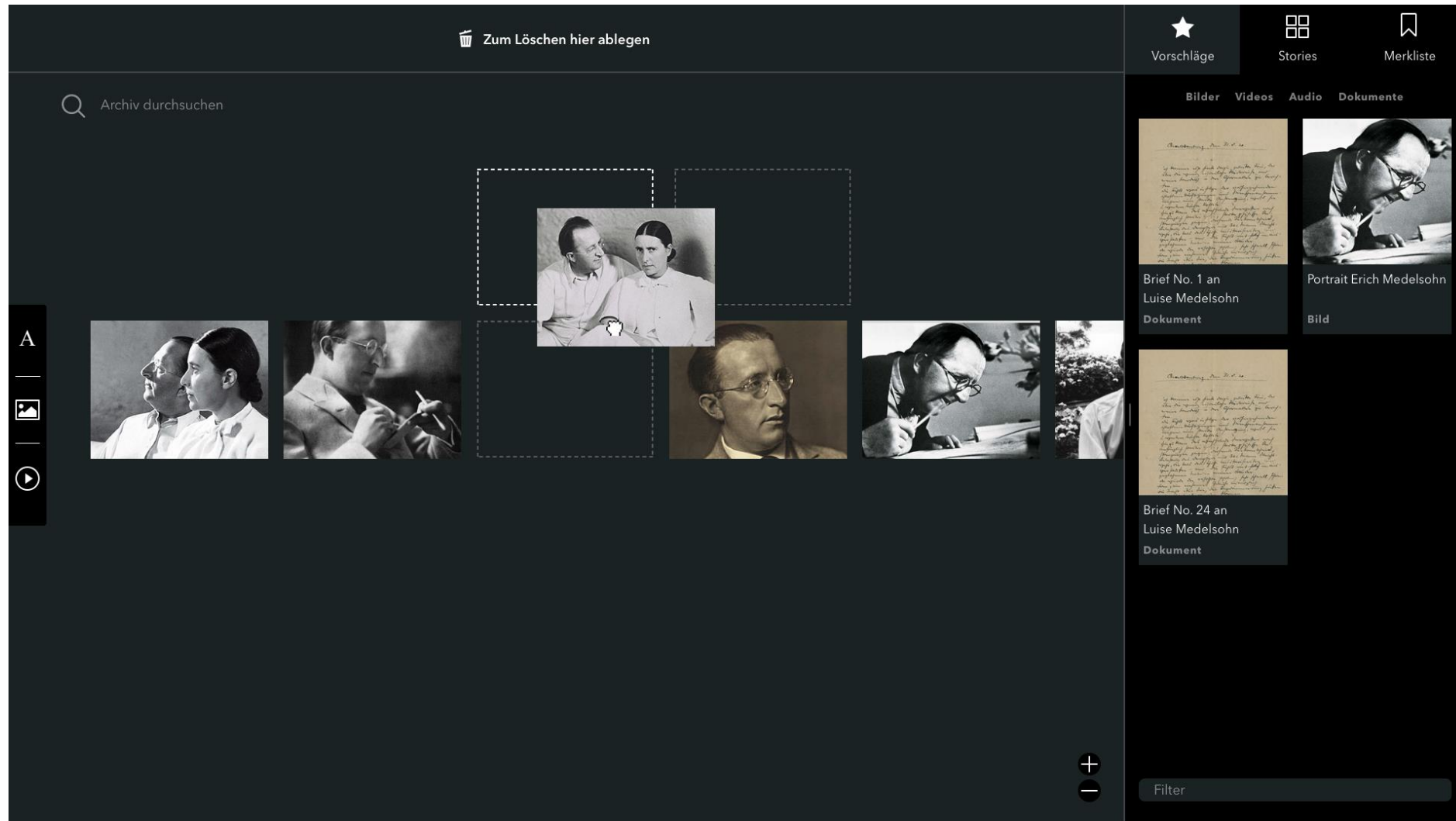


Brief No. 24 an Luise Mendelsohn

Dokument

Filter

4. Use case: Mendelsohn Letters



4. Use case: Mendelsohn Letters

←

Das Leben des Erich Mendelsohn

✎

Eine Biografie geschrieben von Jeremy Larson und Kathleen Montgomery

Letzte Änderung vor 2 Minuten

Q Archiv durchsuchen

A

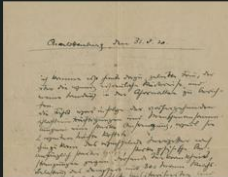

🖼️

▶

Text eingeben

+

-



Text eingeben

Text

×

Schriftgröße:

H1

H2

H3

H4

p

Schriftfarbe:

🟢

⬜

⬛

🔴

🟡

Ausrichtung:

≡

≡

≡

Textfeld:

Text eingeben

4. Use case: Mendelsohn Letters

Das Leben des Erich Mendelsohn

Eine Biografie geschrieben von Jeremy Larson und Kathleen Montgomery

Letzte Änderung vor 2 Minuten

Erch Mendelsohn

Reise

Luise Mendelsohn

Expressionismus

Moderne

Architektur

Emigration

Reise

Amerika

Palästina

Israel

London

Deutschland

München

Berlin / Potsdam

Gerhard Schröder

Luise Mendelsohn

Erich Mendelsohn

Luise Mendelsohn

Bilder Videos Audio Dokumente

Portraits von Erich Mendelsohn

Bilder

Brief No. 1 an Luise Medelsohn

Dokument

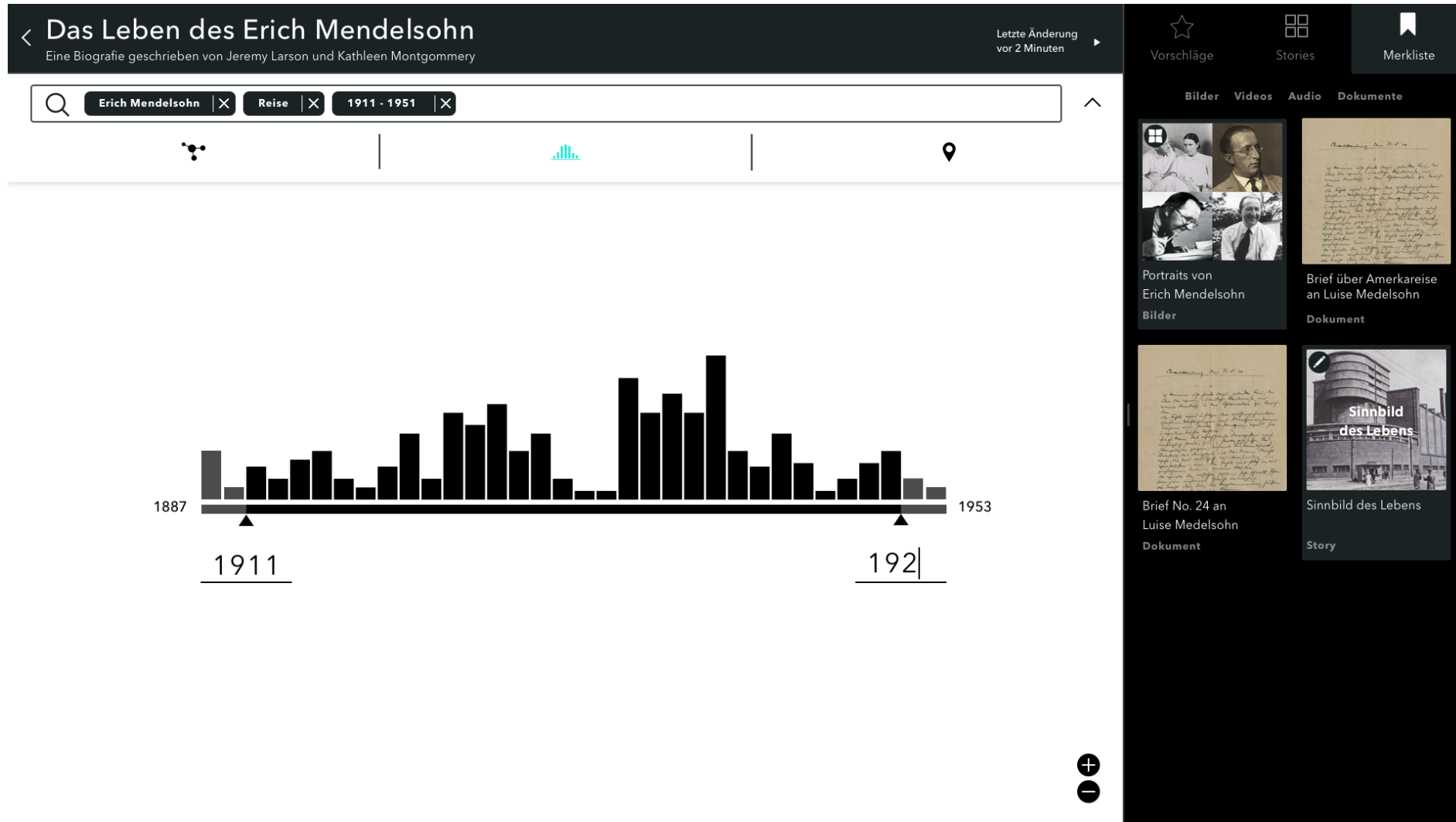
Brief No. 24 an Luise Medelsohn

Dokument

Sinnbild des Lebens

Story

4. Use case: Mendelsohn Letters



5. Conclusions and Future Work

- Goal is to develop tools that will support knowledge workers by:
 - handling, evaluating, sorting and processing document collections
 - enable users to identify faster interesting stories based on the (extracted) information available
 - in order to generate new pieces of content

5. Conclusions and Future Work

Future work:

- Concentrate on the missing steps inside our architecture, like generic text genre processing, discourse rhetorical structure theory etc.
- Look through document collections, create a corpus, experiment with annotations

5. Conclusions and Future Work

- Find the right balance between basic research...
- ...and the need to build soon robust prototypes to develop corresponding technologies and knowledge representation approaches

Thank you for your attention!



Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH



Digitale
Kuratierungs-
technologien

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

The project QURATOR is supported by the German Federal Ministry of Education and Research (BMBF), “Unternehmen Region”, instrument “Wachstums Kern” (grant no. 03WKDA1A).