

Textpassagen semantisch verknüpfen: Anwendungen in den Bereichen Mietrecht und Compliance

Jörg Landthaler, 21.09.2017, Munich

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
www.matthes.in.tum.de

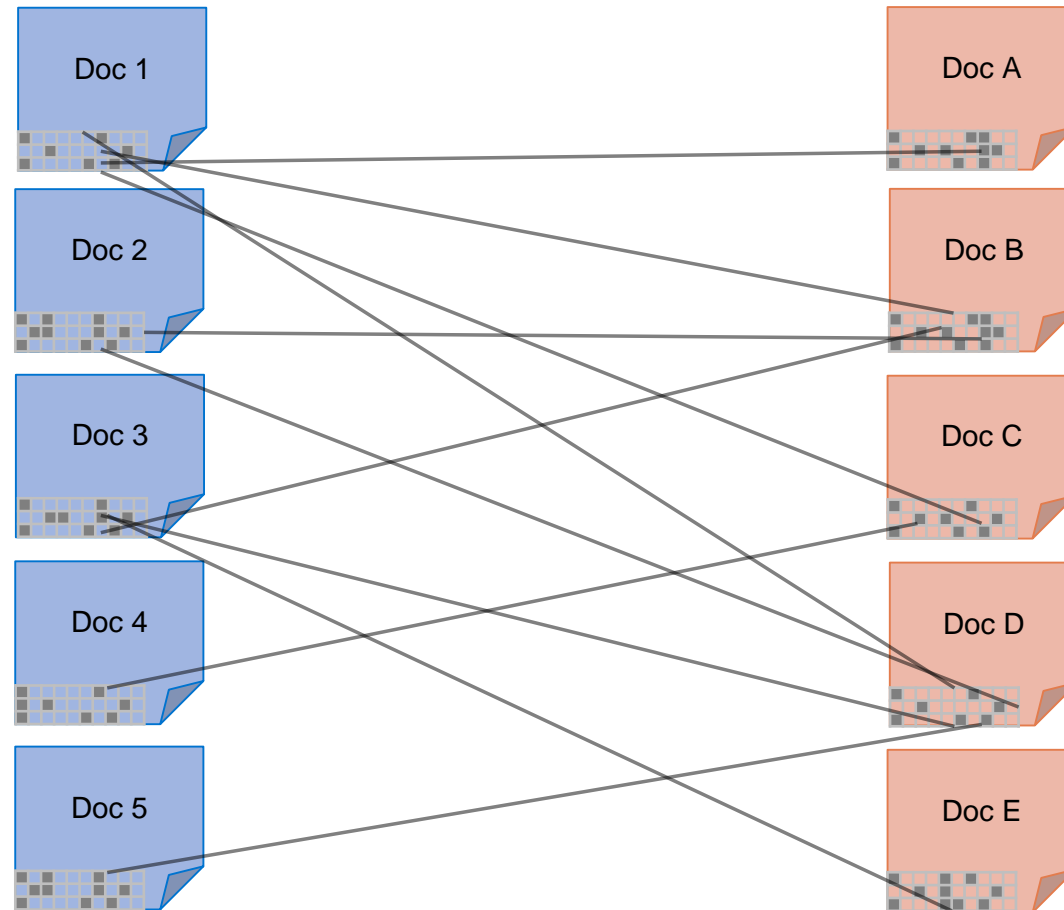
Semantic Text Matching

- Abstract Problem
- Solution Approach

Applications:

- Rental Law
- Compliance

The Semantic Text Matching Problem



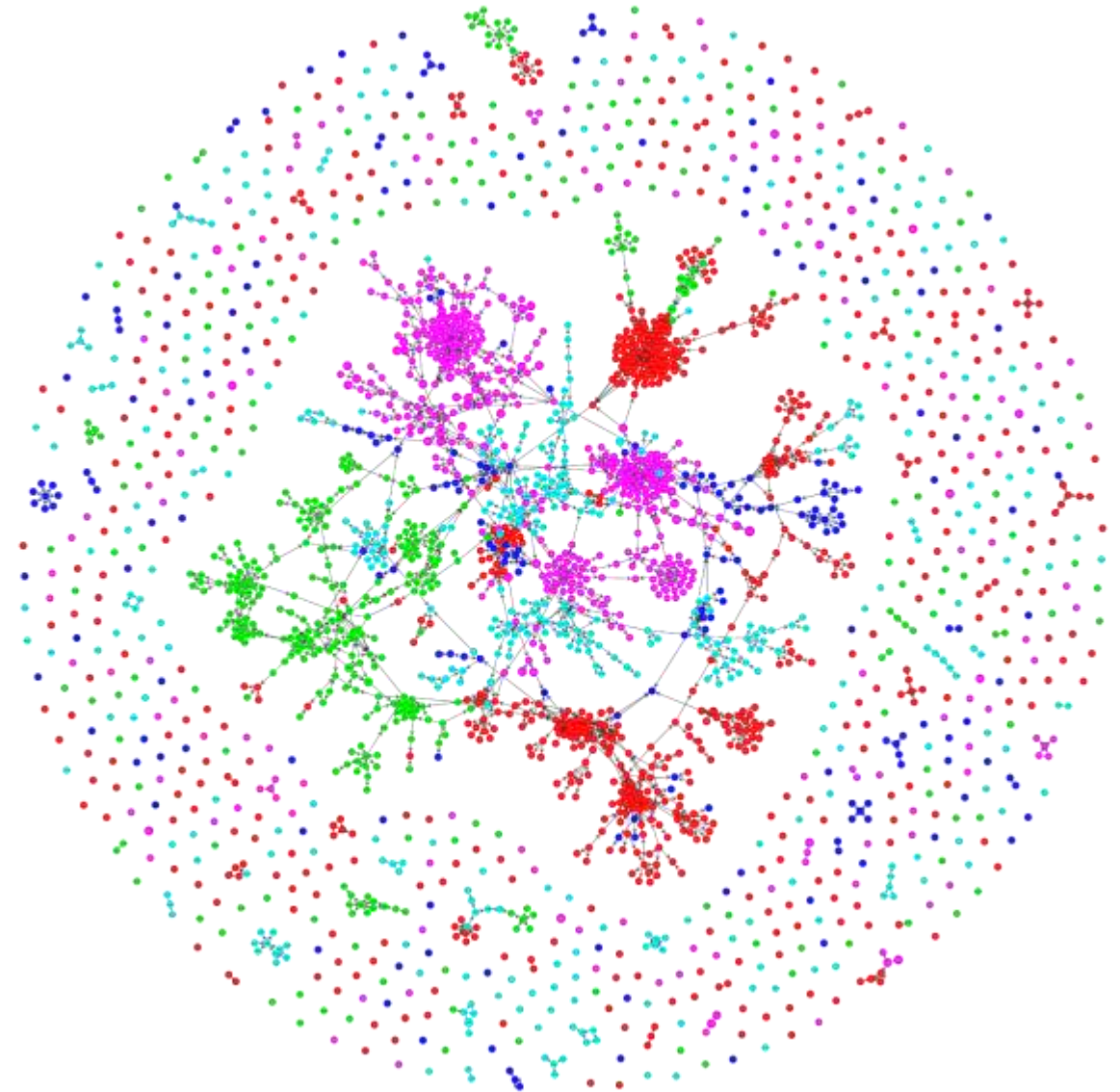
The identification of semantically and/or logically related text fragments in different documents (of different document types), e.g. contract clauses that need to satisfy certain judgements or norms in laws.

A Solution Approach: Reference Detection & Resolution

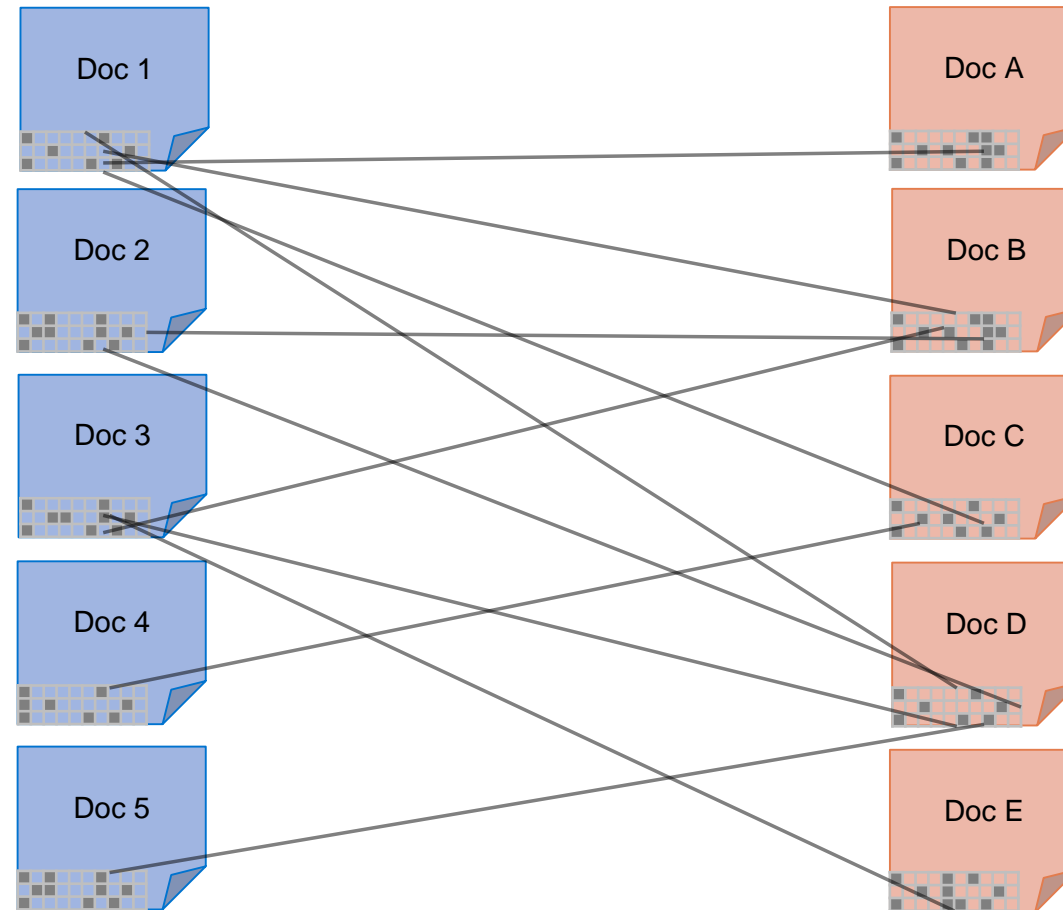
If **direct references** are present, the detection and resolution of references might suffice / is an „easier“ task.

Example on the right: internal references of the German Civil Code.

However, direct references are not always present...

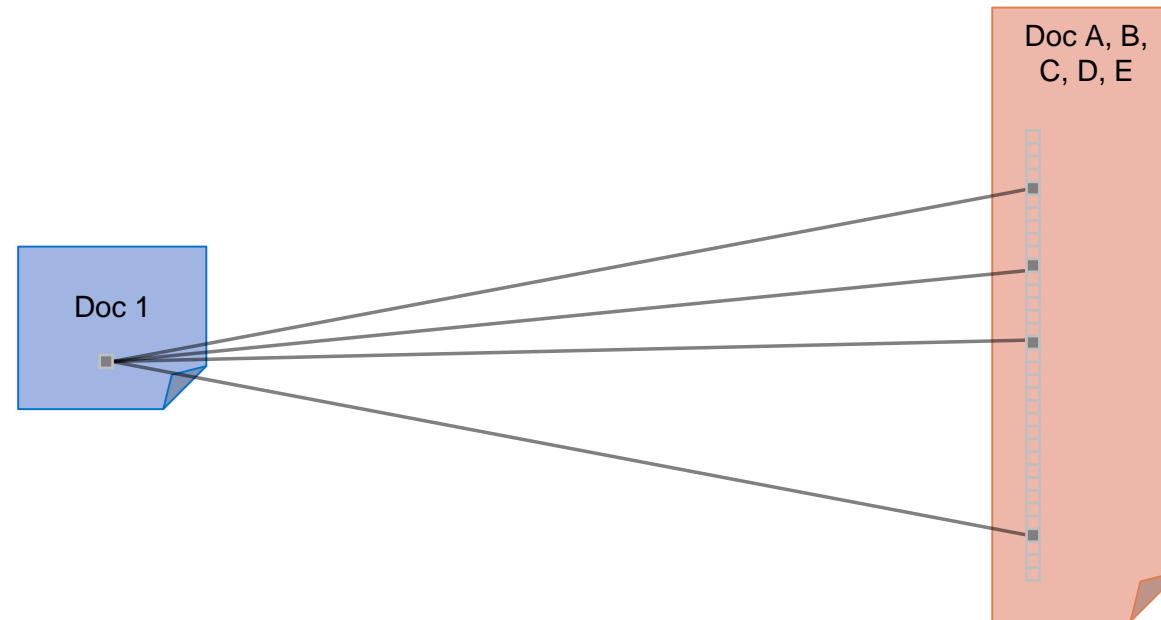


The Semantic Text Matching Problem



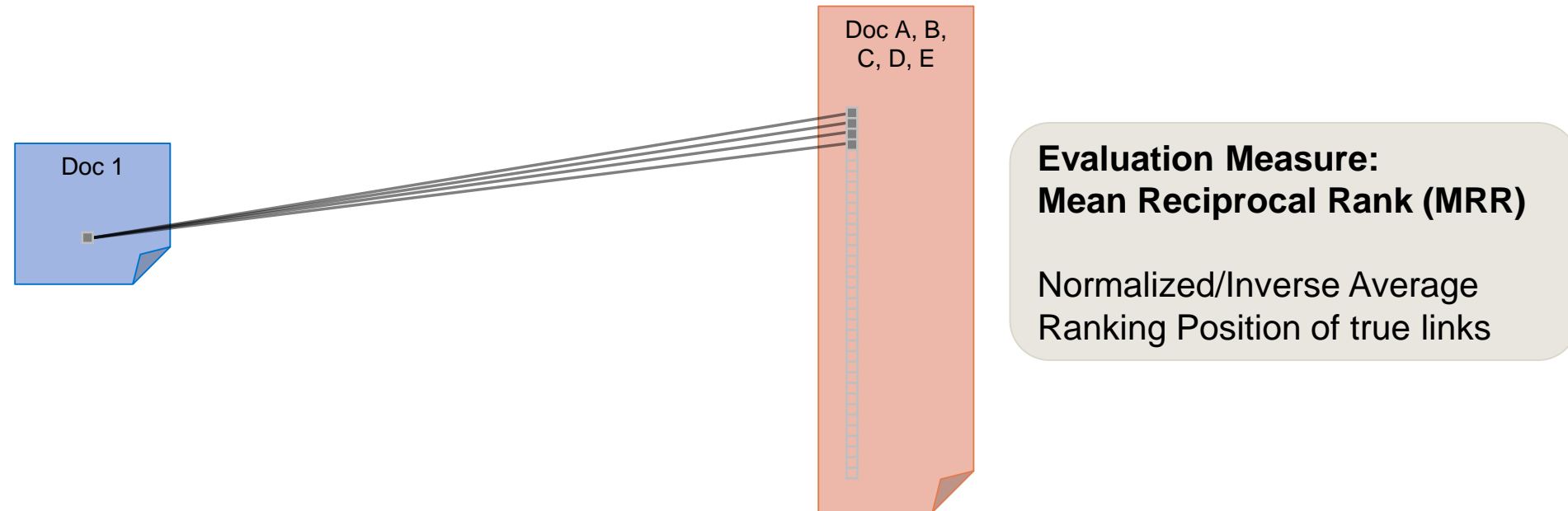
The identification of semantically and/or logically related text fragments of different documents (of different document types), e.g. contract clauses that need to satisfy certain judgements or norms in laws.

Solution Approach: Text Similarity Measures



- 1) Segment text fragments of documents of both types according to a heuristic (e.g. all newlines).
- 2) Pool all text fragments of the documents of the second document type

Solution Approach: Text Similarity Measures (cont'd)



3) Apply a text similarity measure to rank all text fragments of the pool to one text fragment of document 1

4) Iterate all text fragments of document 1, ..., N

TF (/TF-IDF)

John learns to read fast

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Sparse representation

Word Embeddings

John learns to read fast

2,34	4,87	1.01	8,34	3,22
-1,30	3,22	0.01	0.23	5.67
0.23	-1.0	2.44	2.34	0.01
1.33	3.9	3.84	1.04	-1.2
1.0	3.8	-2.08	4,55	2,66

Dense representation

Vocabulary

Manually chosen size

Beispiele für Vertragsklauseln über einen Kündungsverzicht:

Die Mietvertragsparteien **verzichten** wechselseitig bis zum Ende des _____ auf ihr **Recht zur ordentlichen Kündigung**.

Während dieses Zeitraumes kann der Vertrag von beiden Seiten **nicht ordentlich gekündigt werden**.

TF (/TF-IDF)

John learns to read fast

Vocabulary

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Sparse representation

- Easy to use / Works out of the box (e.g Elasticsearch)
- Often works well
- Do not capture semantics
- No relevance criterion

Word Embeddings

John learns to read fast

2,34	4,87	1.01	8,34	3,22
-1,30	3,22	0.01	0.23	5.67
0.23	-1.0	2.44	2.34	0.01
1.33	3.9	3.84	1.04	-1.2
1.0	3.8	-2.08	4,55	2,66

Dense representation

Manually chosen size

- Do capture semantics
- Can determine a subset of relevant links using our intersection approach (relevance criterion)
- Need to be trained (dep. on corpus a few hours to days) & parameters need to be tuned

Semantic Text Matching

- Abstract Problem
- Solution Approach

Applications:

- Rental Law
- Compliance

Startup, munich-based

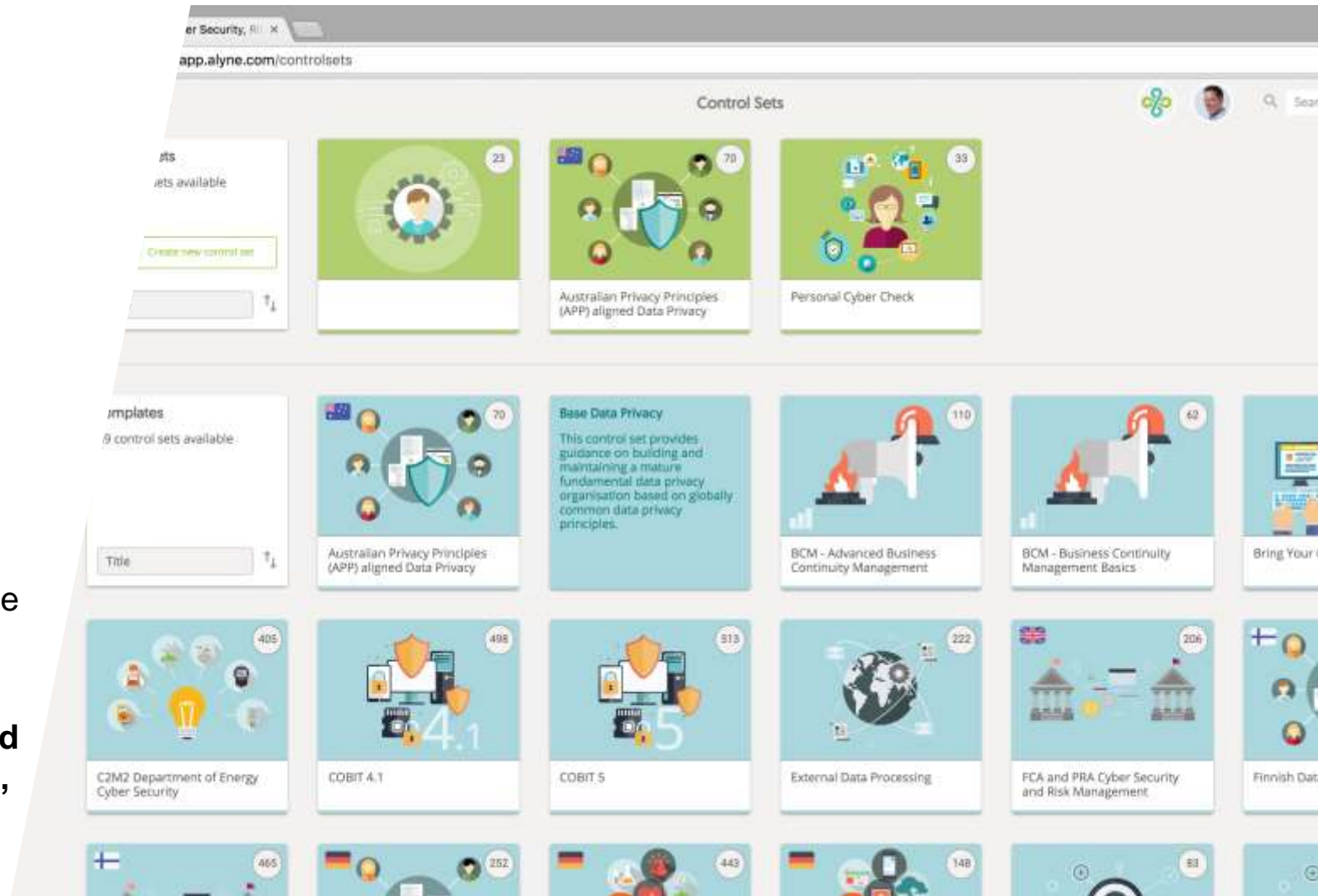
Offers a Software-as-a-Service (SaaS) Compliance Management & Risk Assessment tool.

Larger companies are obliged to use a Compliance Tool.

Contact: Manuel Reil (CTO)

The problem is even more challenging, if multiple regulatory regions are involved (e.g. EU, US)

At the core, **regulatory documents are mapped to manually maintained atomic requirements, called control statements.**



Startup, munich-based

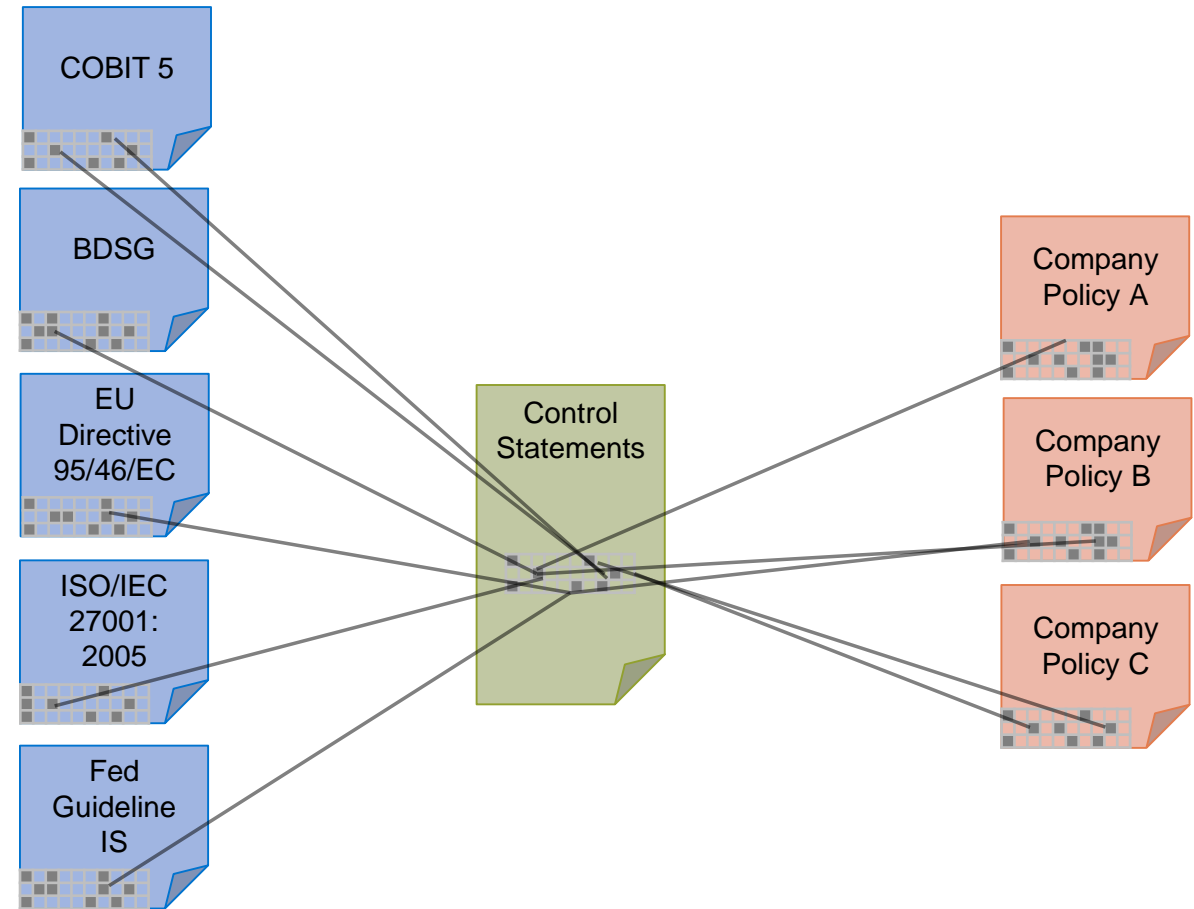
Offers a Software-as-a-Service (SaaS) Compliance Management & Risk Assessment tool.

Larger companies are obliged to use a Compliance Tool.

Contact: Manuel Reil (CTO)

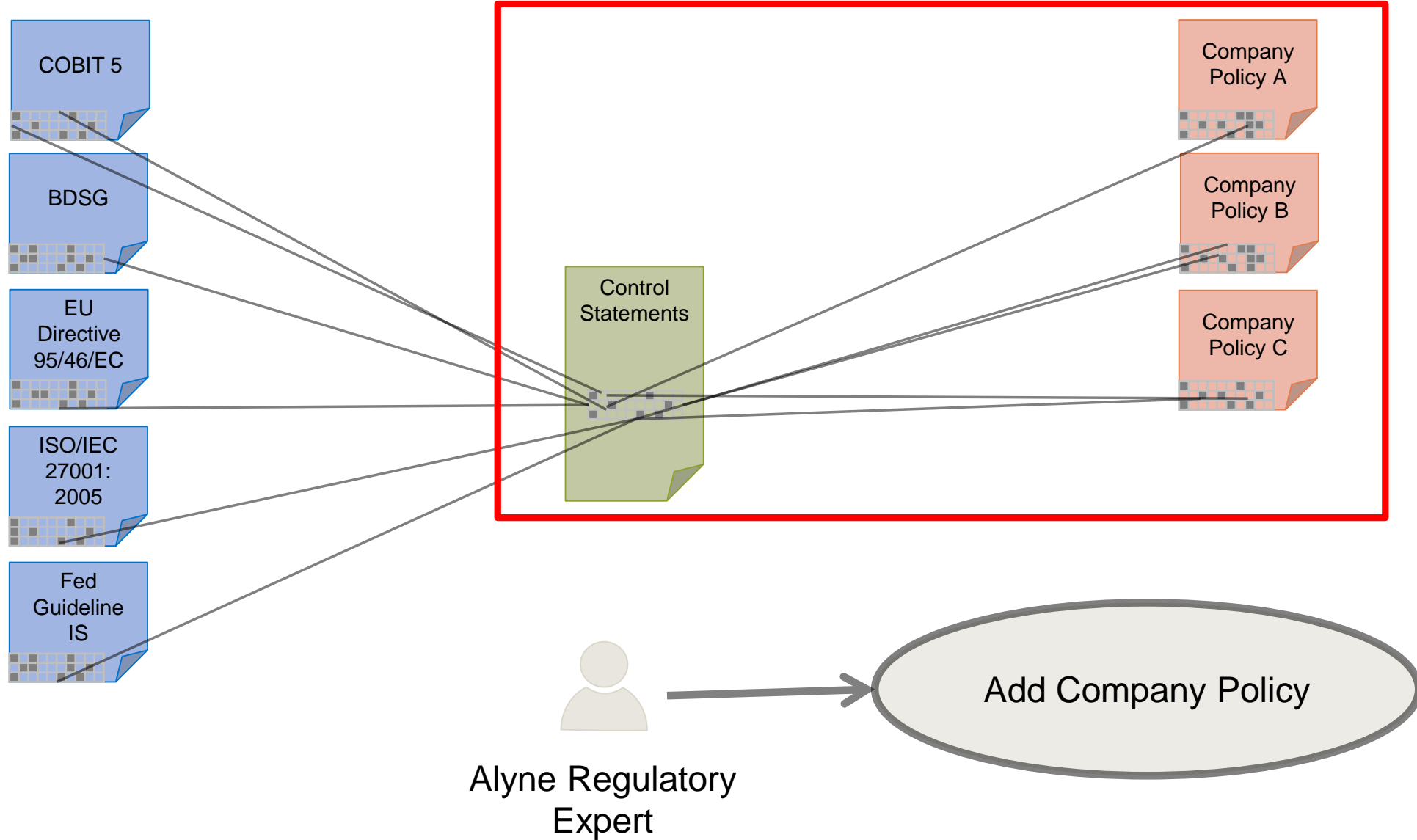
The problem is even more challenging, if multiple regulatory regions are involved (e.g. EU, US)

At the core, **regulatory documents are mapped to manually maintained atomic requirements, called control statements.**



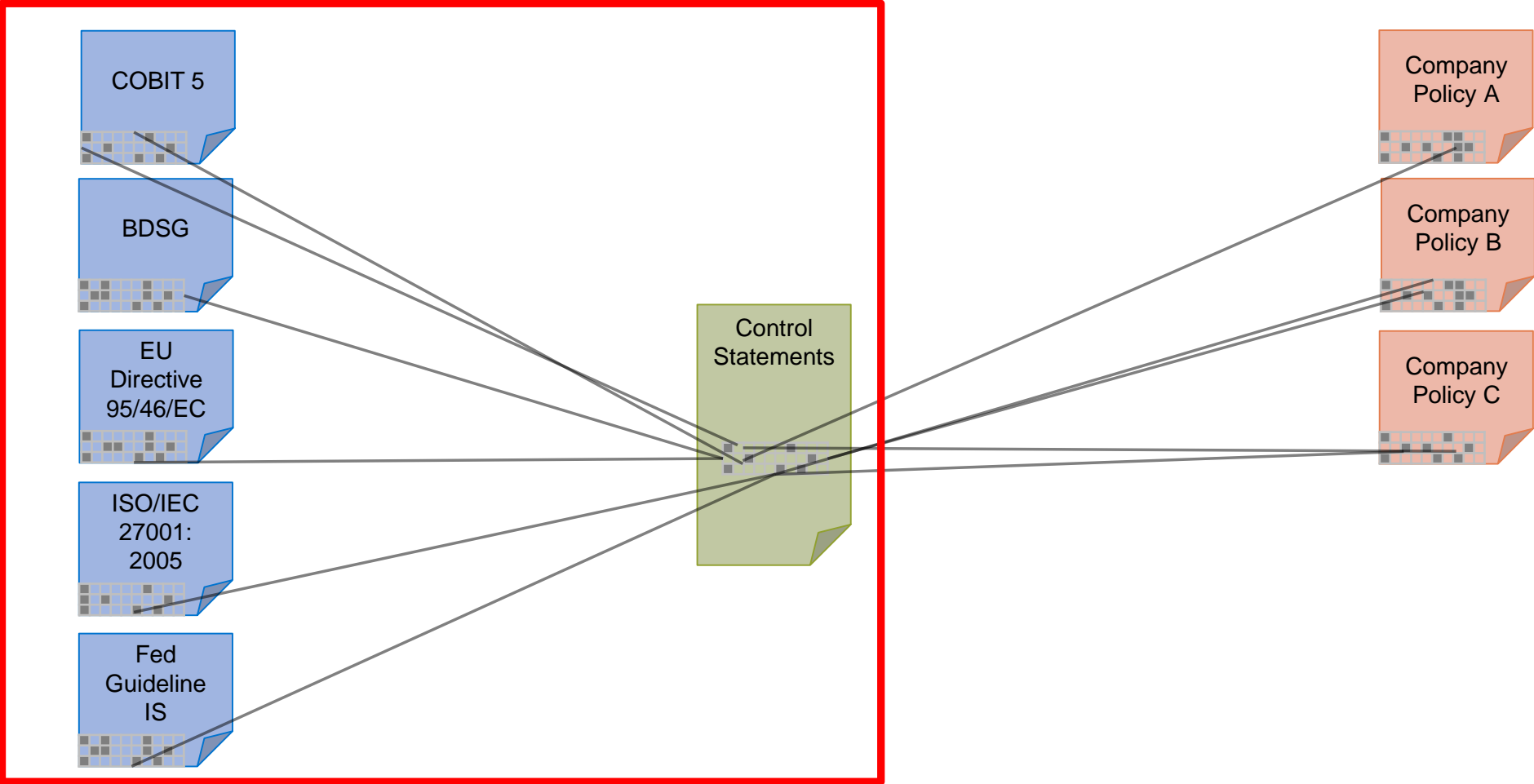
Application: Compliance

Matching Regulatory Documents & Company Policies



Application: Compliance

Matching Regulatory Documents & Company Policies



Document	TFIDF	Word Embeddings
BDSG	0.181	0.168
COBIT 5	0.096	0.124
COSO	0.094	0.074
Finnish Personal Data Act	0.161	0.175
ISO 22301	0.047	0.066
ISO 27001_2013	0.105	0.092
ISO 31000_2009	0.018	0.050
NIST	0.042	0.039

Mean Reciprocal Rank (MRR)
A value close to 1 is best.

- Run parameter studies
- Investigate different NLP pre-processing technologies like stemming, POS-tagging, etc.
- Investigate different corpora for word embeddings training (Wikipedia, Laws, EUROVOC, etc.)
- Investigate, if weighted context can support the approach
- Investigate intersection approach for the identification of relevant subsets
- Evaluate approach with Alyne

Legal Publisher & Digital Media Group

Ca. 1900 Employees

One large Customer Group: Lawyers

Strong focus on Digital Products & Services

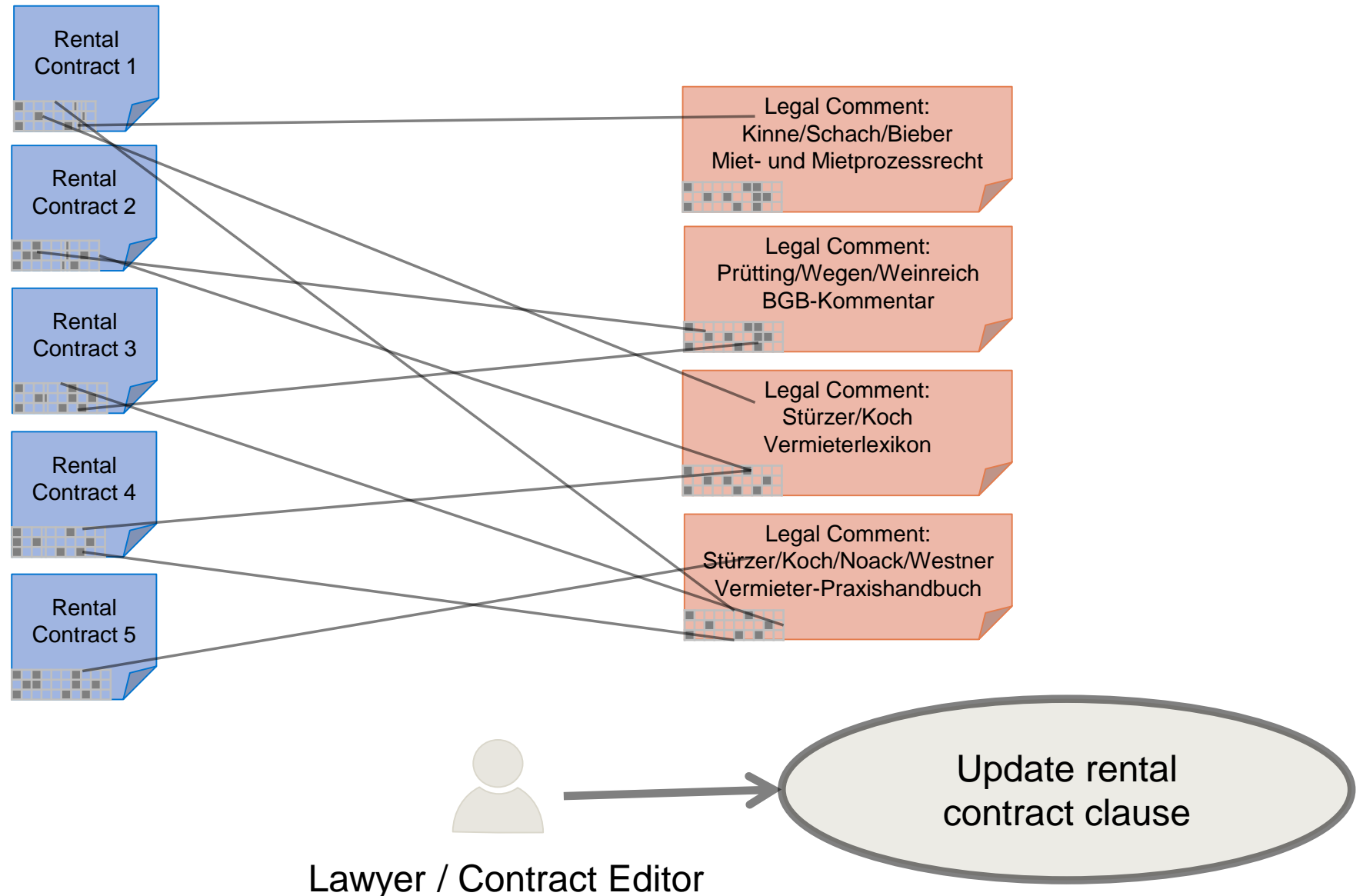
Contact: Hans Lecker (Product Owner)

Mehr über die Haufe Group unter:
<https://www.haufegroup.com>



Application: Tenancy Law

Matching Rental Contract Clauses & Legal Comment Chapters



Application: Tenancy Law

Example Rental Contract Clauses

Andere/zusätzliche* Vereinbarungen zur Kautio: _____

§10 Schönheitsreparaturen

Schönheitsreparaturen als Mieterpflicht Comm.

Die regelmäßigen Schönheitsreparaturen während der Mietzeit übernimmt auf eigene Kosten der Mieter.

Schönheitsreparaturen: Umfang Comm.

Die Schönheitsreparaturen umfassen insbesondere das Tapezieren und Anstreichen der Wände und Decken, das Lackieren der Heizkörper und Heizungsrohre und das Anstreichen der Innentüren sowie der Fenster und Außentüren von innen.

Schönheitsreparaturen: Fristen Schönheitsreparaturen: Fristbeginn Comm.

Die Schönheitsreparaturen sind je nach Art der Mieträume bei tatsächlichem Renovierungsbedarf vorzunehmen. Dies ist im Allgemeinen - gerechnet vom Beginn des Mietverhältnisses oder dem Zeitpunkt der letzten vom Mieter durchgeführten Schönheitsreparatur - nach folgenden Zeitabständen der Fall. - in Küchen, Bädern und Duschen in der Regel alle 3 Jahre, - in Wohn- Ess- und Schlafräumen, Fluren, Dielen und Toiletten in der Regel alle 5 Jahre, - in allen anderen Räumen sowie Lackanstriche an Fenstern, Türen und Heizkörpern in der Regel alle 7 Jahre.

§11 Kleinreparaturen

Kleinreparaturen: Kostenübernahme Comm.

Der Mieter trägt die Kosten für kleinere Instandhaltungsarbeiten während der Mietzeit an Teilen der

Prototypical implementation: Mark arbitrary text selection with cursor...

Andere/zusätzliche* Vereinbarungen zur Kautio: _____

§10 Schönheitsreparaturen

Schönheitsreparaturen als Mieterpflicht: Comm.

Die regelmäßigen Schönheitsreparaturen während der Mietzeit übernimmt auf eigene Kosten der Mieter.

Schönheitsreparaturen: Umfang: Comm.

Die Schönheitsreparaturen umfassen insbesondere das Tapezieren und Anstreichen der Wände und Decken, das Lackieren der Heizkörper und Heizungsrohre und das Anstreichen der Innentüren sowie der Fenster und Außentüren von innen.

Schönheitsreparaturen: Fristen: Schönheitsreparaturen: Fristbeginn: Comm.

Die Schönheitsreparaturen sind je nach Art der Mieträume bei tatsächlichem Renovierungsbedarf vorzunehmen. Dies ist im Allgemeinen - gerechnet vom Beginn des Mietverhältnisses oder dem Zeitpunkt der letzten vom Mieter durchgeführten Schönheitsreparatur - nach folgenden Zeitabständen der Fall. - in Küchen, Bädern und Duschen in der Regel alle 3 Jahre, - in Wohn- Ess- und Schlafräumen, Fluren, Dielen und Toiletten in der Regel alle 5 Jahre, - in allen anderen Räumen sowie Lackanstriche an Fenstern, Türen und Heizkörpern in der Regel alle 7 Jahre.

§11 Kleinreparaturen

Kleinreparaturen: Kostenübernahme: Comm.

Der Mieter trägt die Kosten für kleinere Instandhaltungsarbeiten während der Mietzeit an Teilen der

Die regelmäßigen Schönheitsreparaturen während der Mietzeit übernimmt auf eigene Kosten der Mieter.



Kinne/Schach/Bieber: Einführung (HI1688369)

Einführung

[83] Der Begriff der Schönheitsreparaturen wird im BGB nicht erwähnt. In § 535 ist lediglich die Verpflichtung des Vermieters festgehalten, die Mietsache dem Mieter in einem zum vertragsgemäßen Gebrauch geeigneten Zustand zu überlassen und sie während der Mietzeit in diesem Zustand zu erhalten. Schönheitsreparaturen gehören dazu. Die Ausführung von Schönheitsreparaturen gehört daher grundsätzlich zur Instandhaltungspflicht des Vermieters nach § 535 und ist damit eine schuldrechtliche Hauptpflicht. Die Ausführung der Schönheitsreparaturen gehört zu diesen Instandsetzungspflichten. Die Ausführung von Schönheitsreparaturen gehört daher grundsätzlich zur Instandhaltungspflicht des Vermieters nach § 535 und ist damit eine schuldrechtliche Hauptpflicht. Dieses festzuhalten ist besonders wichtig, weil bei unwirksamen Überwälzungsklauseln der Schönheitsreparaturen auf den Mieter die Pflicht zur Ausführung der Schönheitsreparaturen eben beim Vermieter verbleibt...

Kinne/Schach/Bieber: Überwälzungsklauseln (HI1688372)

Überwälzungsklauseln

[90] Die Übertragung der Schönheitsreparaturen durch eine Allgemeine Geschäftsbedingung auf den Mieter muss eindeutig sein.

Eine Klausel Der Mieter ist verpflichtet, die während des Mietverhältnisses anfallenden Schönheitsreparaturen auf eigene Kosten durchzuführen, ist eindeutig und unbedenklich. Dasselbe gilt für eine Klausel

Der Mieter verpflichtet sich, die Schönheitsreparaturen auf eigene Kosten durchzuführen.

... and (hopefully) see relevant legal comment chapters on the same screen.

- Finish mapping of legal comments to clause tags (1800+ chapters, 200+ tags)
- Investigate different NLP pre-processing technologies like stemming, POS-tagging, etc.
- Run parameter studies
- Investigate different corpora for word embeddings training (Wikipedia, Laws, EUROVOC, etc.)
- Integrate library into Lexia workbench
- Evaluate approach with Haufe

- **Identification of other STM use cases**
- **Classification of rental contract clauses**
- **Ontology creation on rental law** (e.g. knowledge-graph expansion using semi-supervised graph mining with word embeddings, Google Expander)
- **Human controlled/refined semi-supervised machine learning using visualizations**



M.Sc. (TUM)

Jörg Landthaler

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for Business
Information Systems

Boltzmannstraße 3
85748 Garching bei München

Tel +49.89.289. 17139

Fax +49.89.289.17136

joerg.landthaler@tum.de
www.matthes.in.tum.de

