

Networks of Names

Obtaining Lombardi's "Narrative Structures" by Combining Visual Analytics and Language Technology

Master's Thesis by

Artjom Kochtchi

supervised by

Prof. Dr. Chris Biemann

Dr. Tatiana von Landesberger

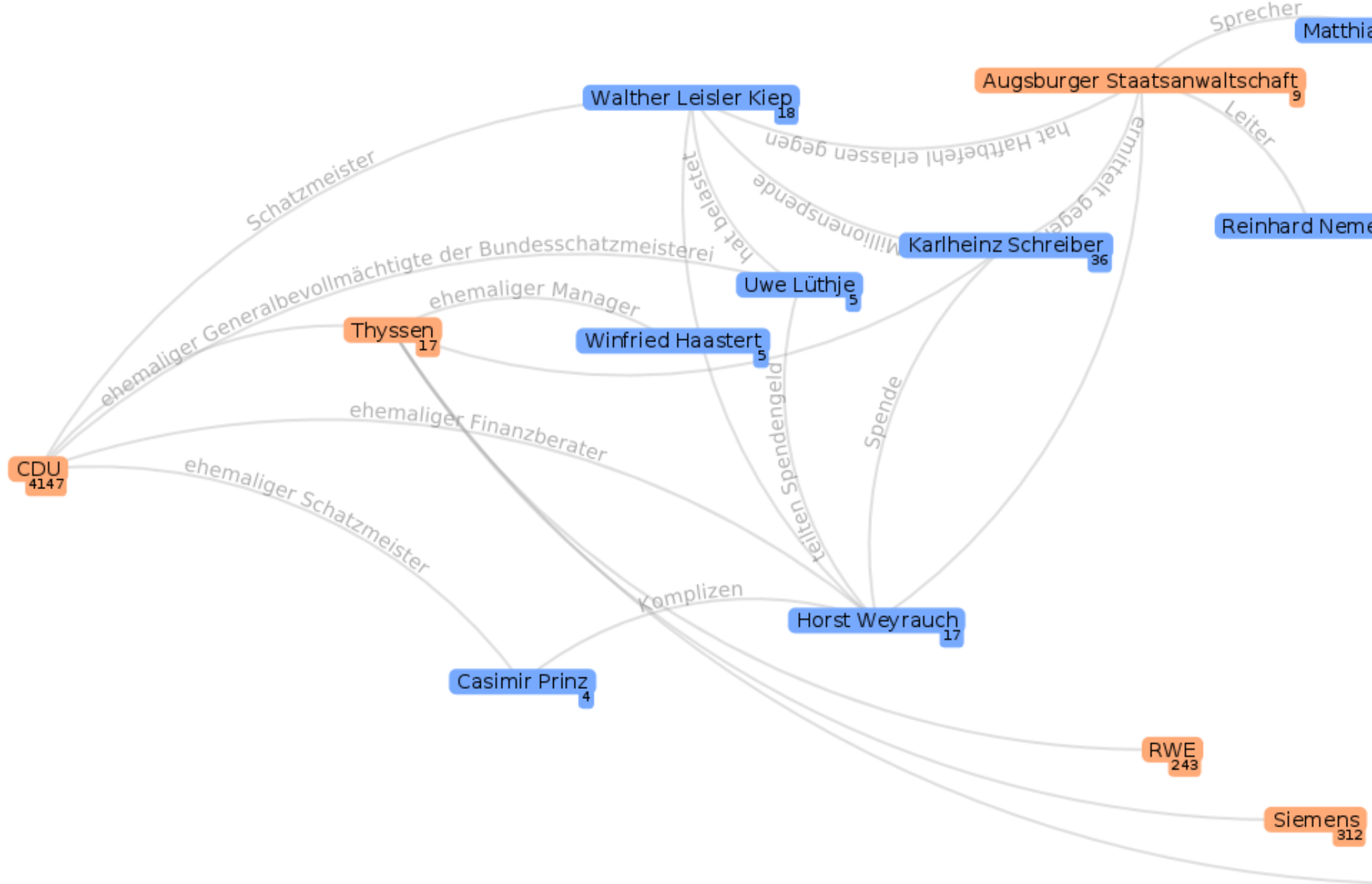


Narrative Structures

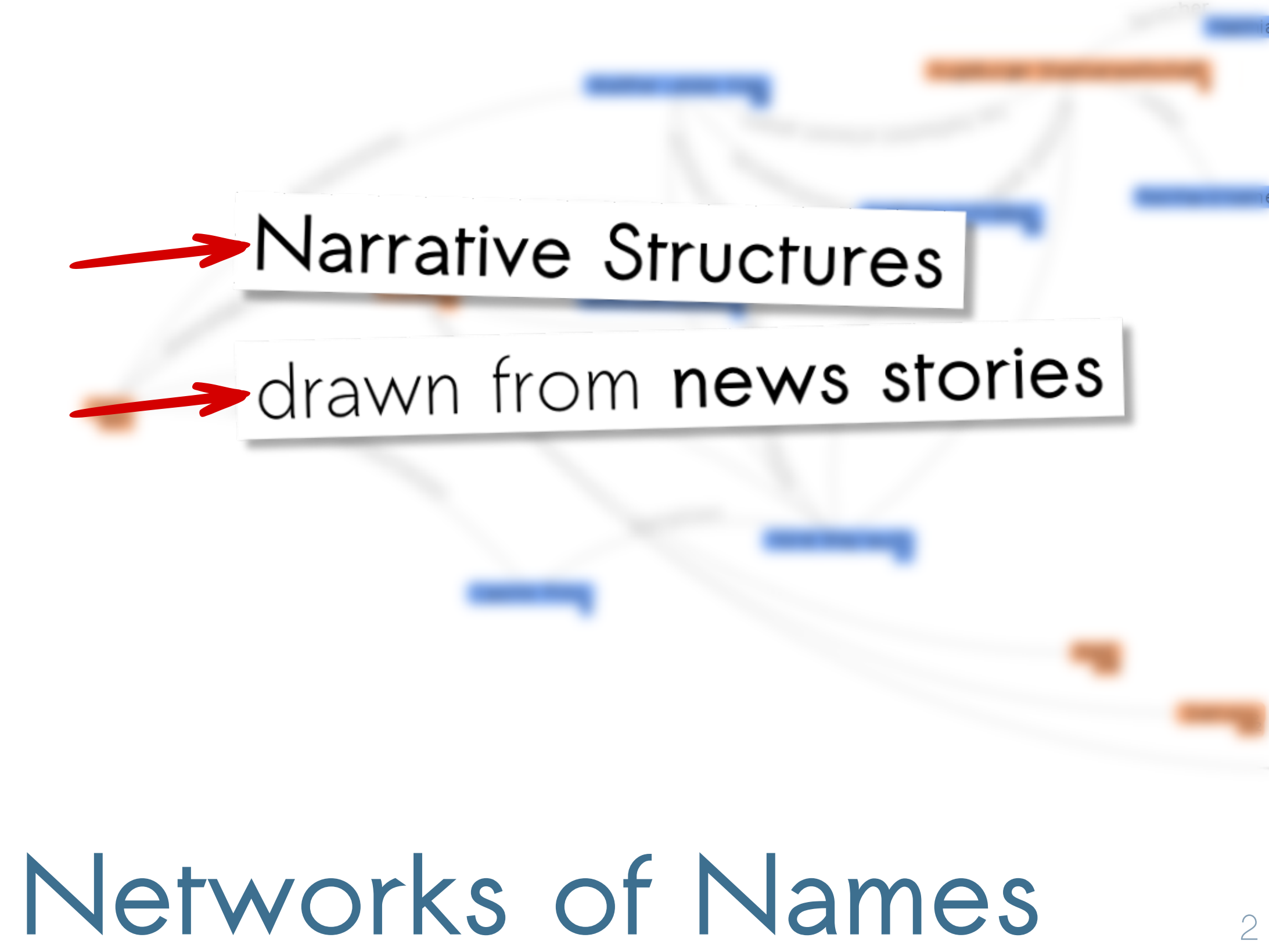
drawn from news stories

hand-written database

Mark Lombardi



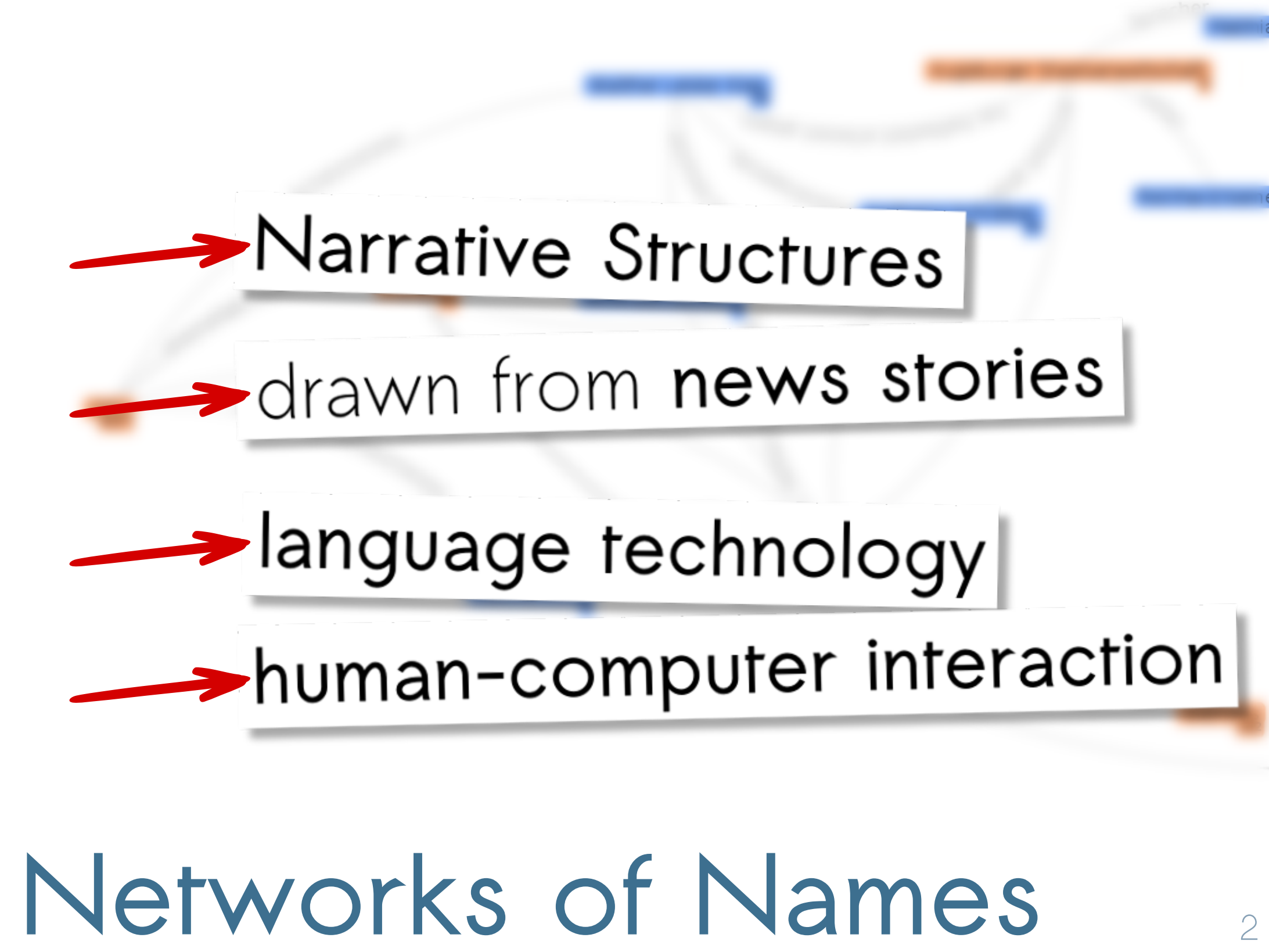
Networks of Names



Narrative Structures

drawn from news stories

Networks of Names



→ Narrative Structures

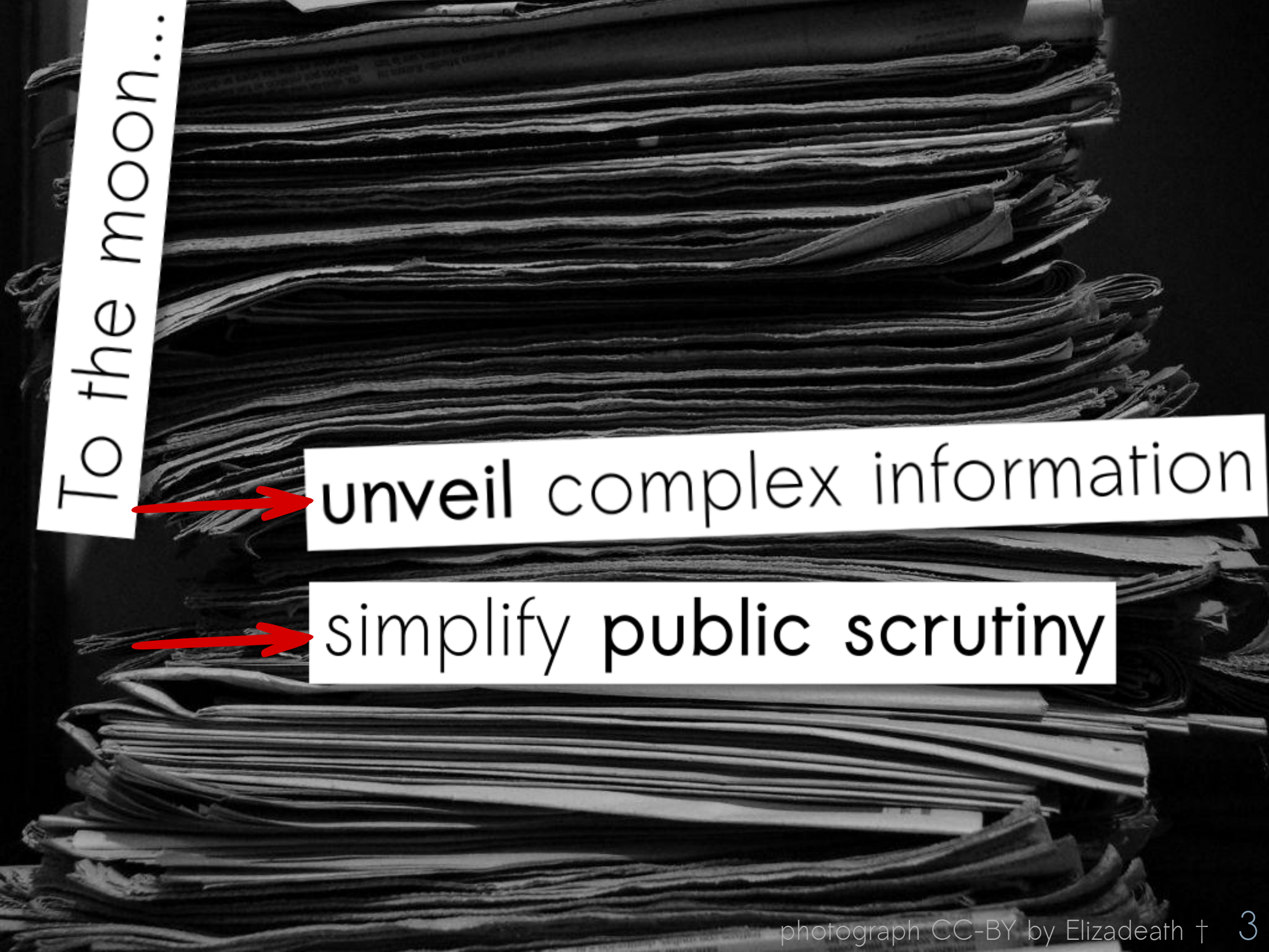
→ drawn from news stories

→ language technology

→ human-computer interaction

Networks of Names

To the moon...



To the moon...

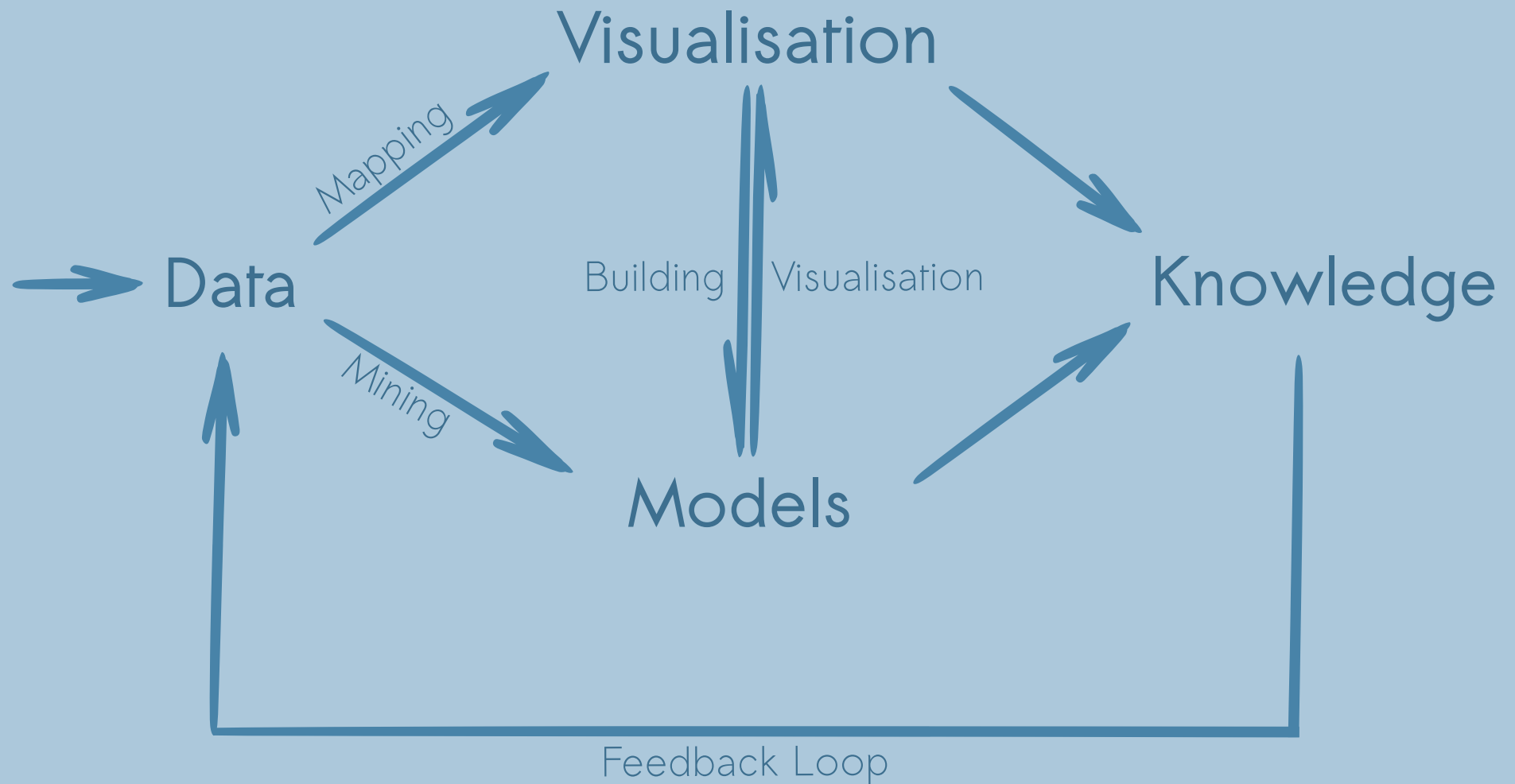
→ unveil complex information

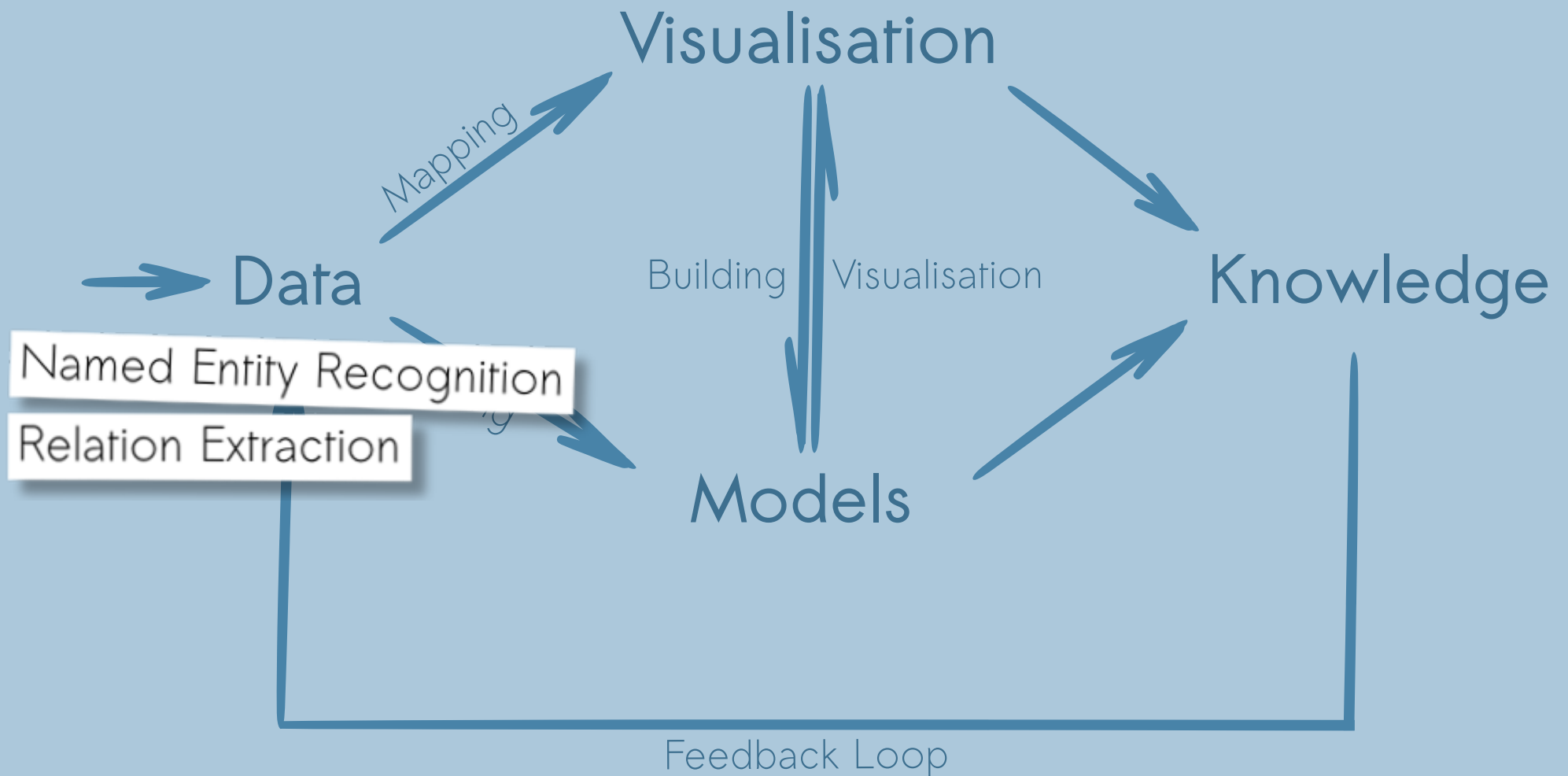
→ simplify public scrutiny

Contributions

Contributions

- integration and adaptation
- populatization of automation
- knowledge from HCI





Network Visualization

Search

Visualisation

Mapping

→ Data

Named Entity Recognition

Relation Extraction

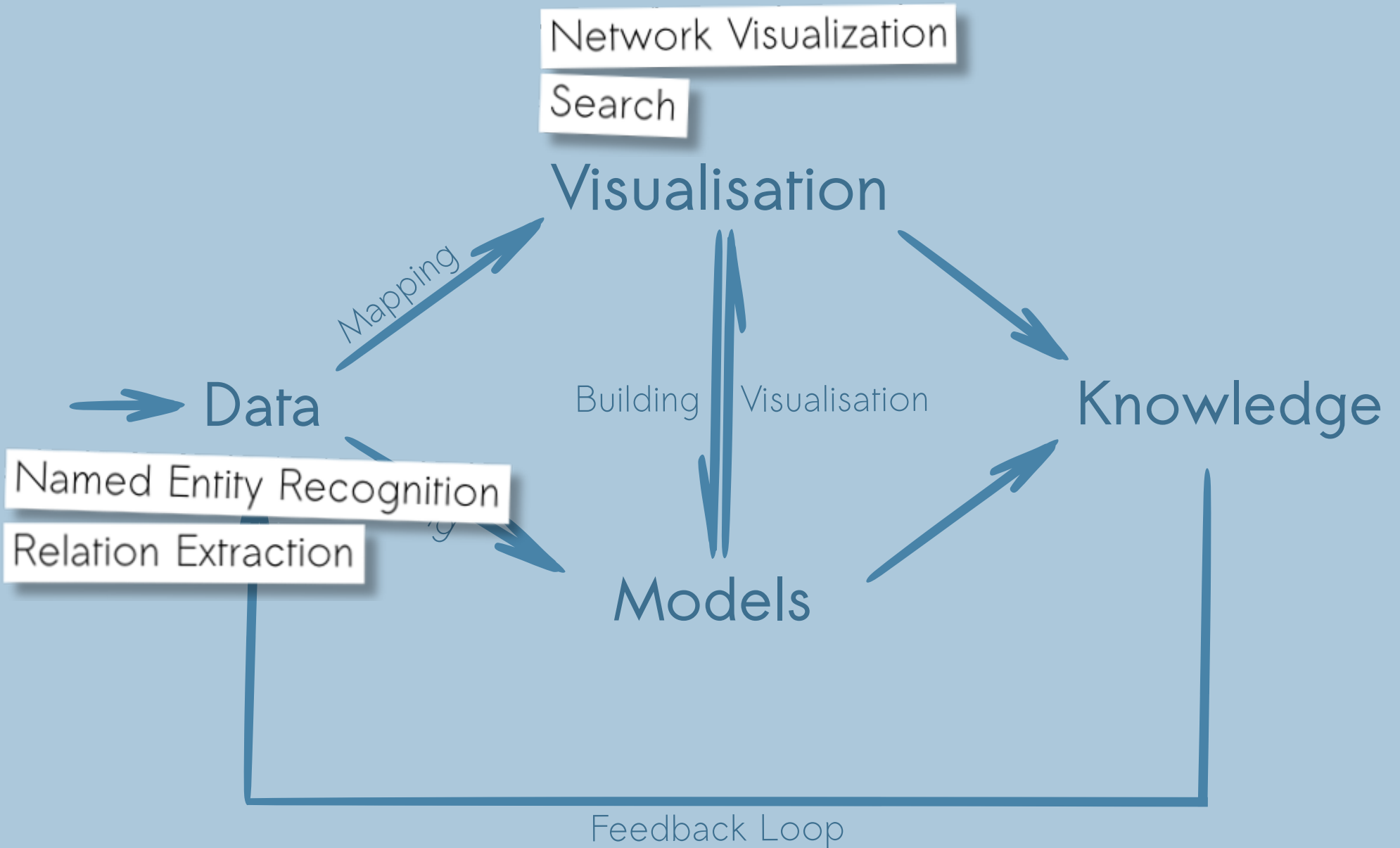
Building

Visualisation

Knowledge

Models

Feedback Loop



Network Visualization

Search

Visualisation

Mapping

Data

Named Entity Recognition

Relation Extraction

Building Visualisation

Knowledge

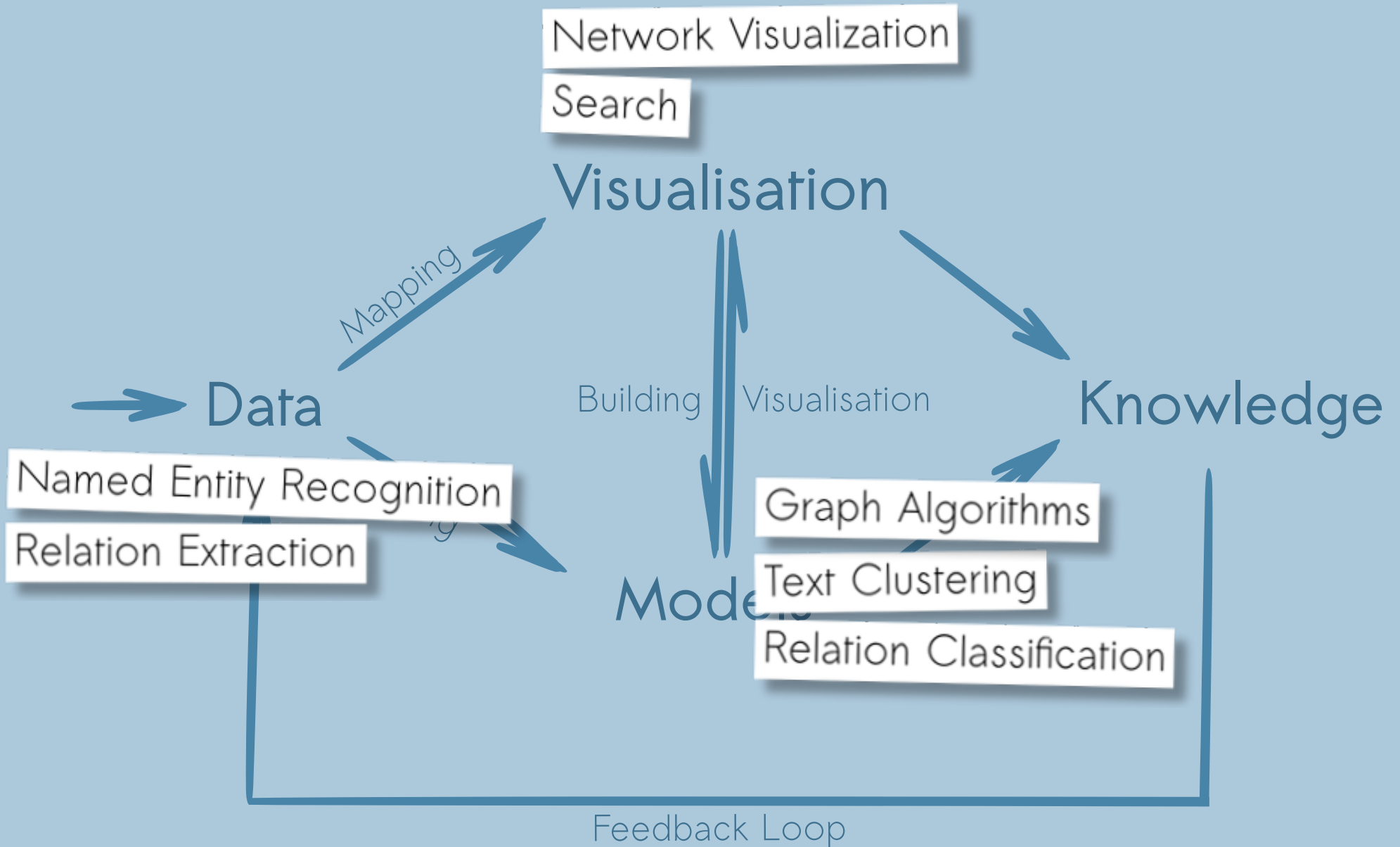
Graph Algorithms

Text Clustering

Relation Classification

Models

Feedback Loop



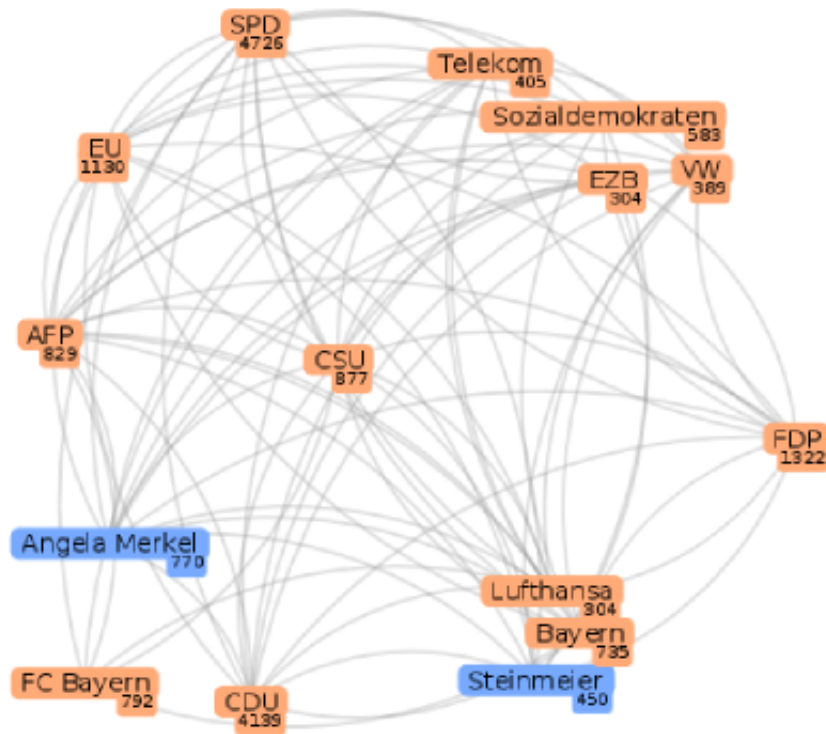
Overview,
Zoom and Filter,
Details on Demand

hen Sporthochschule nd Eichl Dr Berliner Wasserbetriebe ortbund and Berlin NGA on KND Meiser O Carla Jörg Schönbohm W Joachim F
ederer Birzele nski Roade Dr. Eberhard Engelhardt fenschel E Max Mirnyi iter Rensing S
sen Euro Deutsch In Deutschen Oper Zach Brian Eno S APEC Manch Tanja Gön E Schumann 3-Geheimdienst CIA Max Mirnyi iter Rensing S
Aznar Mike Kostunica David Lynch Marku Max Eberl Behrend Grit Schank Grien Schwimm-Verband Tamara
Strauss-K Jim Morrison id Odonkor MRT Till Backhaus ue Marseille nnes Bong Nick Schust Carl nid Lida Strombach
Rafael Nadal Be Bisky Inivereität K Schulzes Three Lions Tietmeyer Radovan Karadzic Charly ard nieber-Jastran
my Connors Babs chen Meisterschaften mermann zky einz Fä Strieder Hin Bachmann Brieger Paul Wolfowitz
MSF C Schano Hillen chon 1000 ngine Rogowski ola Bauer Schumachers lt con Valley Tausend Eur Roger Wittmann Ka Silke Laute
F Pavel Kuka Gre IG-Metall Nick Nolte Eurostat nari James Jones a Habin Pavel Levine LEU sportl
el Gross Bund Bridget Jones Charlotte na nsky Christoph Waltz RAF Max Ernst ats Schwab zader BA
E Andrzej Juskowiak Kinkel ich Ni Thysser Marlies Schnid ax chneling Max Ernst ats Schwab zader BA
hultze der Max Goldt Har Bar Bellini inhof Schul MF Er David Kelly Hilde Melanie olfgang Streibl user Benjamin wermelskirche
Birkholz Opel Baader Jörg Pilawa Schulze el la Te Werner Gegenbauer necht stoph Schlingensief Breuer ka Srpska grar Nick F
resemann Grit F Hanneleore Korn analgaler Europo Schumann Werner Gegenbauer necht stoph Schlingensief Breuer ka Srpska grar Nick F
Volkgang Gehrke SSV Jahn Regensburg y James Joyce IFC Deutschen Tennis Ri Hilmar Kopper St. Pa
Nick B Olyr Strutz Lyon Deutschen Olympischen Sportbundes Dats Ganger aus Matras Karl Krai Deutschen Skiverband I Studmans
Ch Christo Vilmar bert d Pout Bachirou Saou verner Hackmann li SPS Marku Markus Weiss
Me Krab Sri Lanka im Hörster Sci Tiffany Mil L Bachirou Saou verner Hackmann li SPS Marku Markus Weiss
her Korte dienstes Lydia Hartl ald Rumsfeld us Altegoer AOL Europe Le Birkenstock Bisc Ba Be Rohde Er Efo-Instit
Hannele Wolfgang Ji MSI pelis-Sperk Ger SSC Teutonia ly Altegoer AOL Europe Le Birkenstock Bisc Ba Be Rohde Er Efo-Instit
itsche Oper Wolfgang Roh Deutschen Oper Berlin Kostas Ko Brian Wilson Les Boch W Tanja Szewczenko ichthon Markus
larco Straller Ichibo Roland and K Rafael Benitez Eurosport SSV Hapag Alberto Fujimori e Agentur TI Lutz Sikora Schwanhol
r Heilmann a m Alberto James Prieotte K Knabach Donald Tusk Franz Gret Rafa Lehne Birgit Schrowan
Deutsche Roten Kreuz Nick Heidfeld David Letterman Kari marx Hannes Heer Gorki Deutsch Fraport AG unde Mov Miguel Indurain
ncent Birthler Igor Iwanow Schwan rs Eurowings Deuts Struck id McAllister fuller IEA art Lyndon B. Johnson Molfgang
ang Hedera De Heinz Weisener Idemur upa ik A Paulo Sousa er Donald eutschen Tourenwage Led Zennel Strunk IG Chemie KPRF
hwitz Petra F Deutschen Roten Kreuz Hanni geney yball V Cornelia Piener Albert Schmidt M Ganne Pavel Nedved UAI nda Strunzer Leip
Meisterschaft IL Violetta Gerhard Schmid Lydia wel Gratschow and Isse Alphonse Tchami tall gurd Max Karlheinz Stoc Si Biskys euther Fürth
Bretton Woods Strauss che S Sigmund Max Rauffer elika Mann s ndk IBA Schur Repnik Nick Cave ilk
M Virnil Hill duogan Al Le David Rosset Korschakow k pri Par Donizetti Brian Jones Caritas Deutschen Snorthundes
Benjamin Auer Frattini Nick Faldo Floyd Landis KPMG Focus Clinton SSC Neapel FN Chaplin idth TVG
Sachs Davi Frattini Nick Faldo Floyd Landis KPMG Focus Clinton SSC Neapel FN Chaplin idth TVG
Fredrick Bat Hans Werner Henze Mieterbundes Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks Bismarcks
P. Volk Ern IDS Scheer Rari Rohwein an Roger Mo Tiedemann Cora Frost Cora Frost Cora Frost Cora Frost Cora Frost Cora Frost Cora Frost Cora Frost Cora Frost
Radio Vatikan Th Hilpert Roger Willen Marco van Basten Müller-Wohlfahrt Marconi Karl-Josef Laumann Schumanns
Sa Wolfgang Joop Squadra Azzurra Mike Bullard iockemö Mark Tawin Erdogan waltscraft Augsburg ard Wi SR
Frie LDK Birk hultz itani Kim Clijsters SB Lear imere r AG Han Donald Duck di Sprri Wern Struve g o ike Bahr
Markus Cha Hilsberg on nil Wolfgang Gerke SS Christoph Spycher Sabrina mockennauf id Alker A valmal n Schwarzeneggers
my Carter Karlh SSV Reutlingen Birkler Karl Jaspers Carla Del Ponte Springstein Schle Be FMG a Wolfgang Gerhardt s
thhaus acher a Vinc ICC Tanja De Apel Karl Jaspers Carla Del Ponte Springstein Schle Be FMG a Wolfgang Gerhardt s Florian Hal

Search,
Show Context,
Expand on Demand

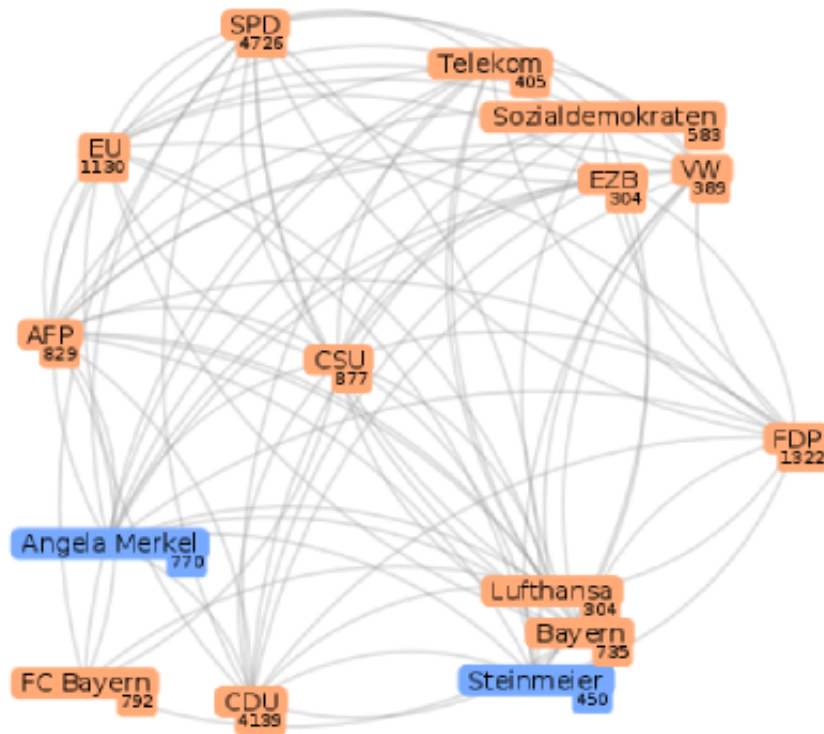
Search,
Show Context,
Expand on Demand

The “Angela Merkel Problem”



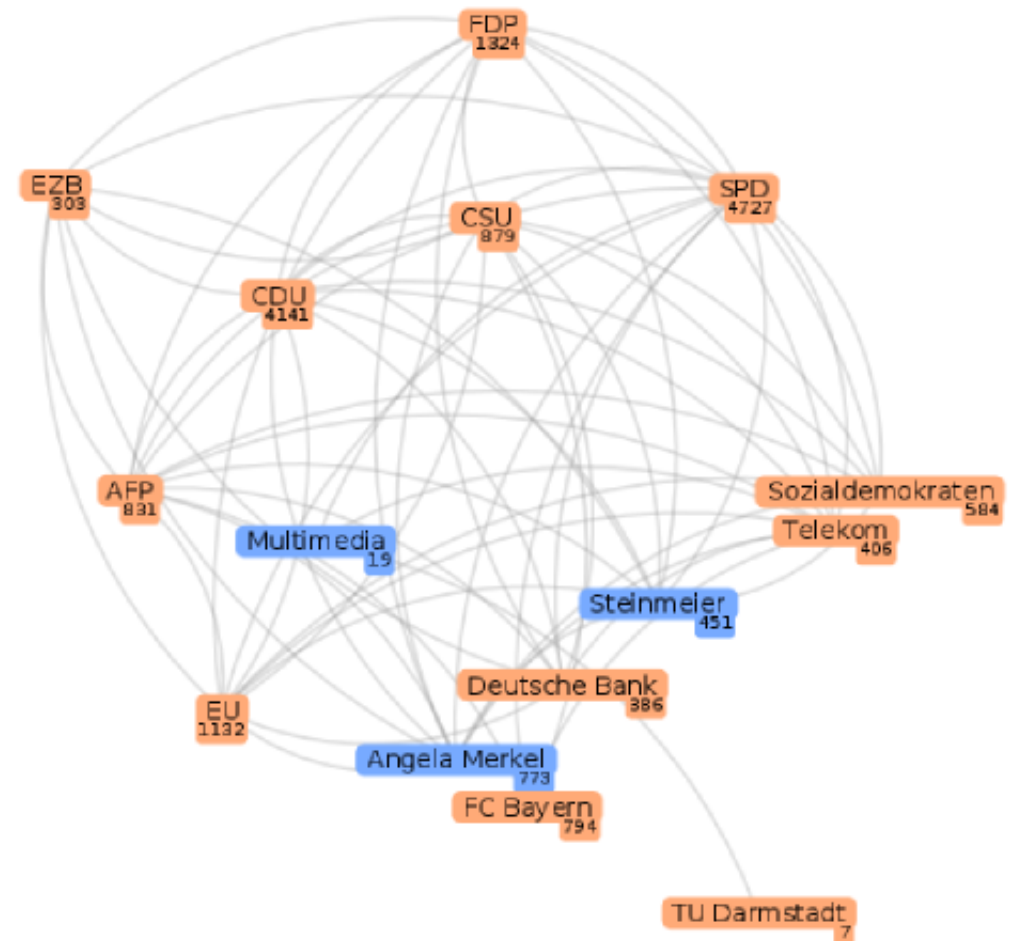
Angela Merkel

The “Angela Merkel Problem”



Angela Merkel

TU Darmstadt





Degree of Interest

+ a-priori interest
+ user focus
- distance

 Degree of Interest 

+ a-priori interest

+ user focus

- distance

→ Degree of Interest } for
edges
+ a-priori interest
- user focus
distance

→ Pointwise Mutual
Information

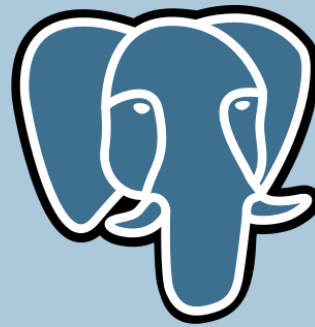
→ Degree of Interest } for
edges

+ a-priori interest
+ user focus
- distance

→ Pointwise Mutual
Information

→ don't expand hubs

 **Scala**
play 



Demo



D3.js  
 *CoffeeScript*

„Gegenkandidat des 53-jährigen Mathematikers war der bisherige
Vizepräsident der TU Darmstadt, Johannes Buchmann.“

Johannes
Buchmann



Vizepräsident der

TU Darmstadt

„Gegenkandidat des 53-jährigen Mathematikers war der bisherige Vizepräsident der TU Darmstadt, Johannes Buchmann.“

Johannes
Buchmann  Vizepräsident der TU Darmstadt

„Gegenkandidat des 53-jährigen Mathematikers war der bisherige
Vizepräsident der TU Darmstadt, Johannes Buchmann.“

Johannes
Buchmann  Vizepräsident der TU Darmstadt

„Gegenkandidat des 53-jährigen Mathematikers war der bisherige
Vizepräsident der TU Darmstadt, Johannes Buchmann.“

Johannes
Buchmann

Vizepräsident der

TU Darmstadt

<WORD> der <ORGANIZATION>, <PERSON>

<WORD> der

User Study

User Study

→ test software

→ explore impact of vis

→ generate data

User Study

- ① instructions
- ② conduct search
- ③ explore graph
- ④ label relationships

User Study

① instructions

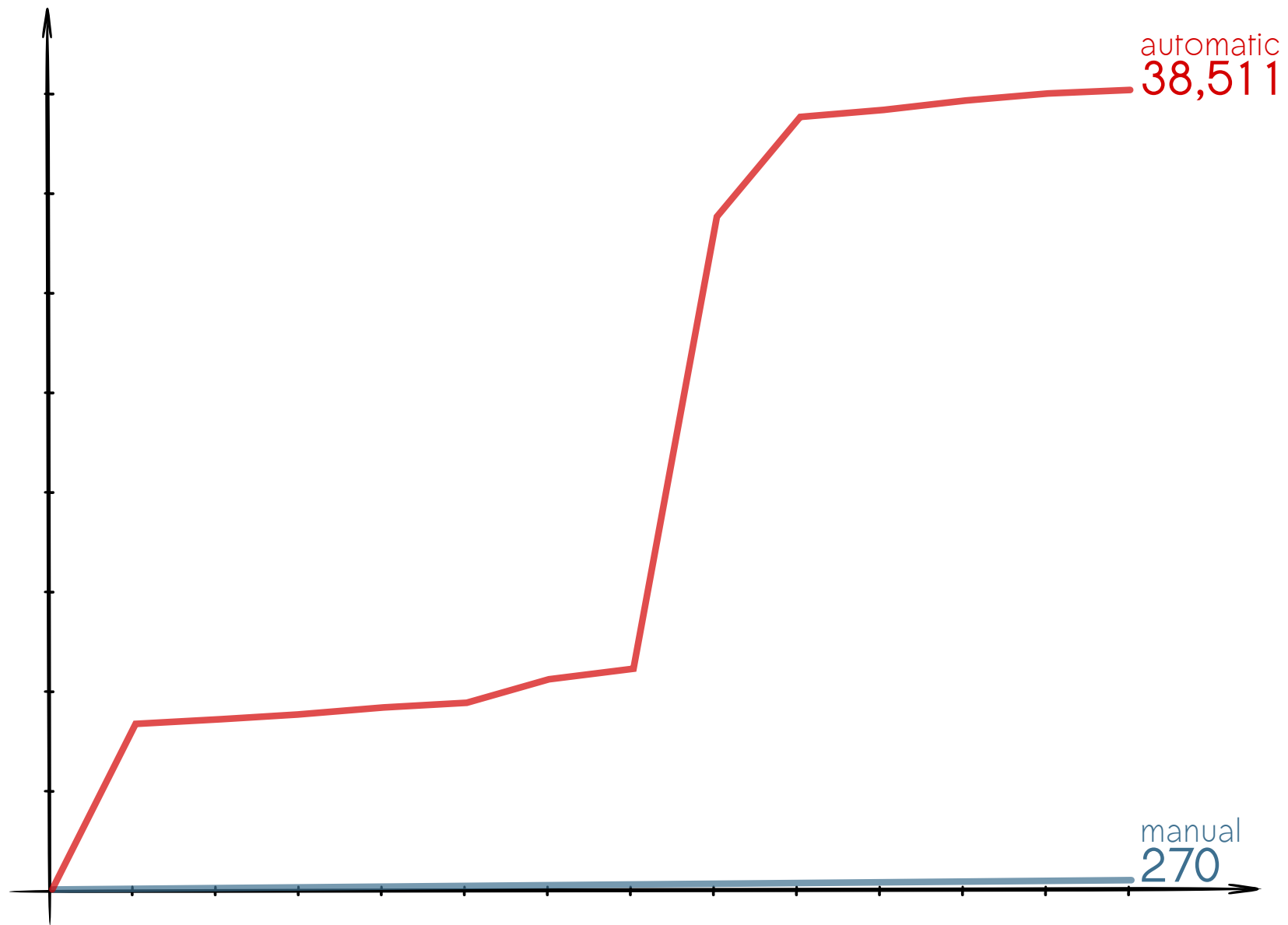
② conduct search

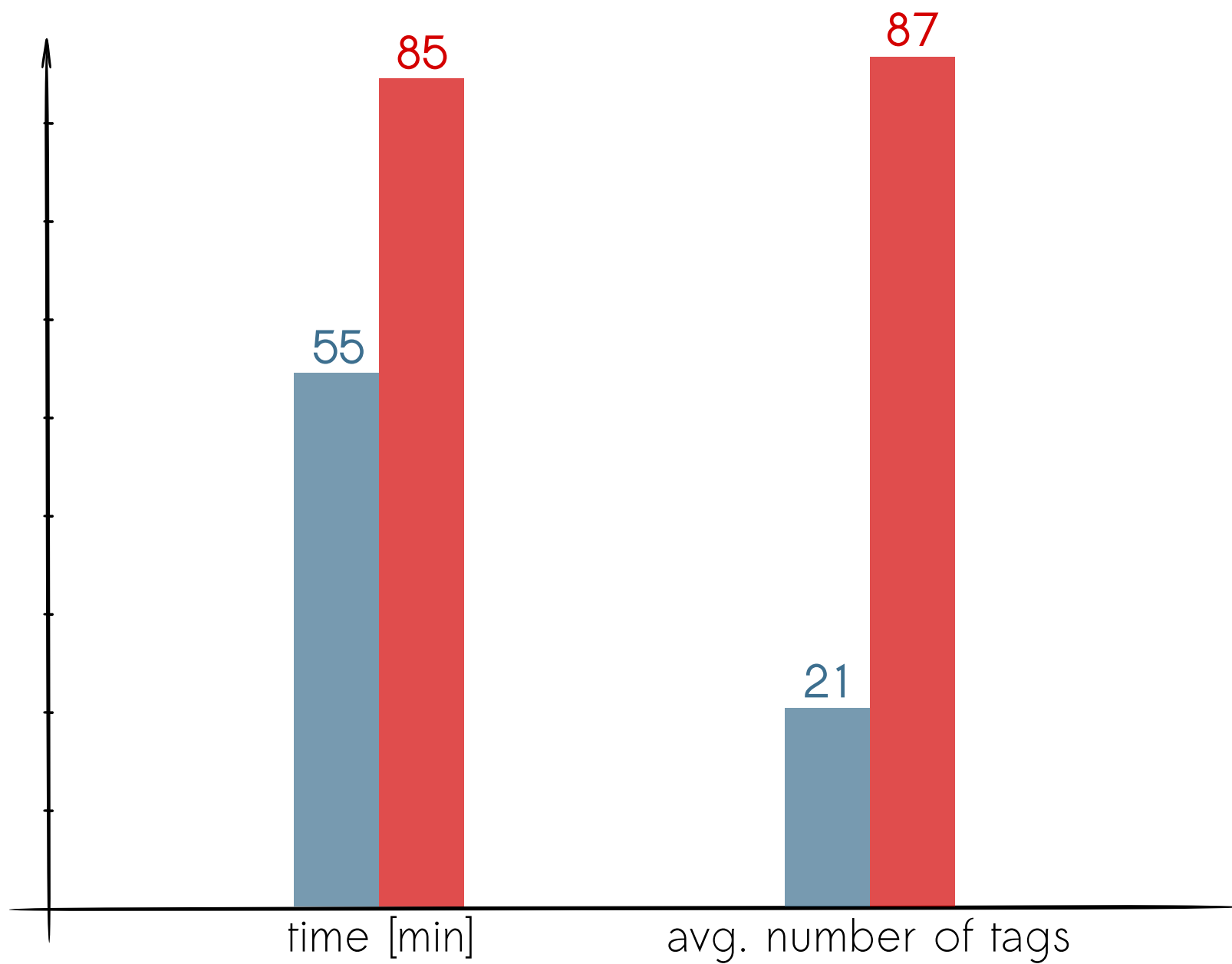
③ explore graph

④ label relationships

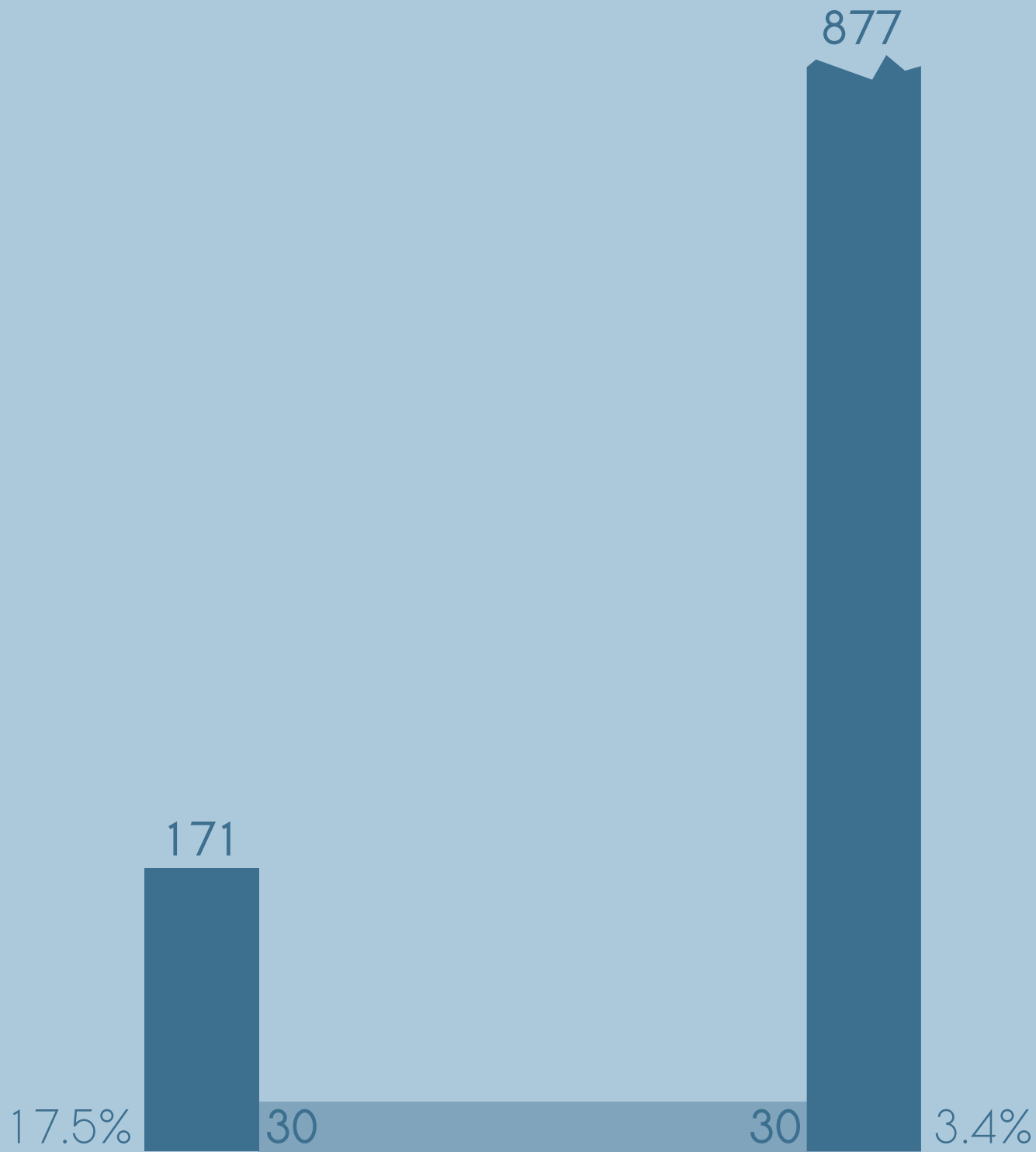
① instructions

② label relationships





	Visualization	Text
# manual tags	270	1,128
# automatic tags	38,511	161,857
# labels	171	877
avg. # words	1.53	2.67



Vorstand

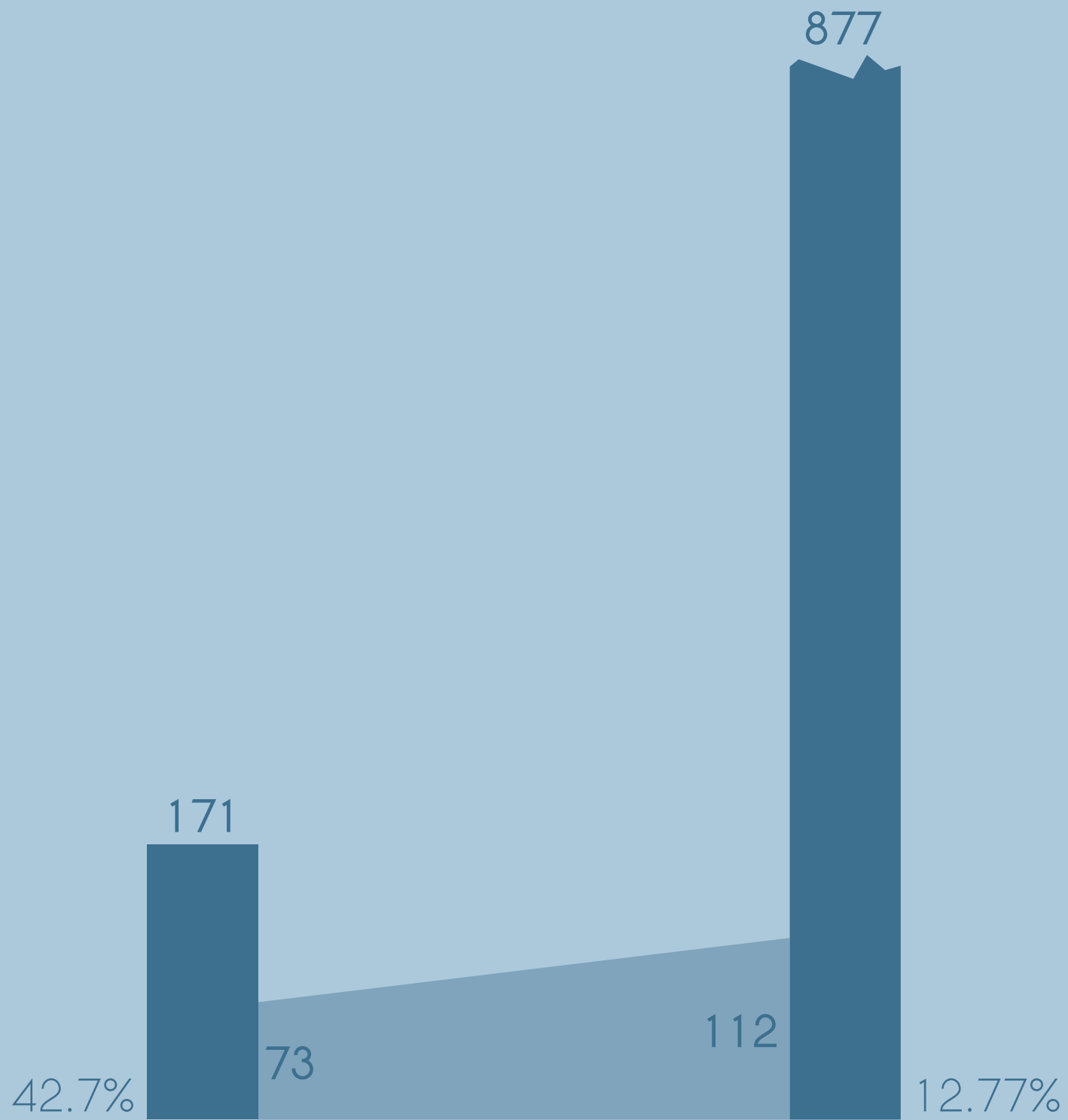
ist
war
wird
hat
hatte

ggf.
vielleicht
offenbar
möglicher
langjähriger
ehemaliger
alter
neuer
Ex-

Vorstand

smitglied
svorsitzender
skollege

des
der
von
für
an
in
auf
mit
bei
...

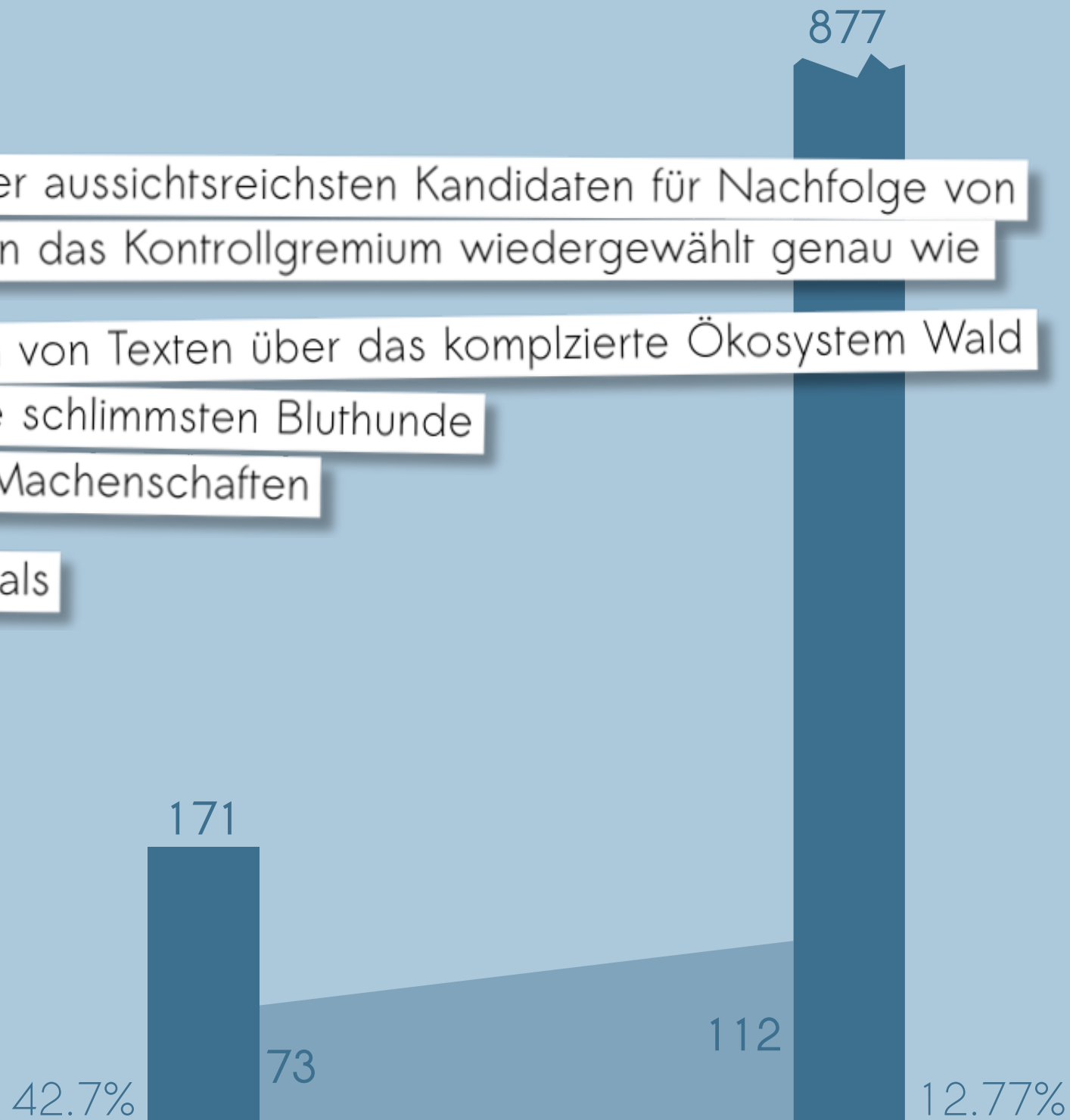


einer der aussichtsreichsten Kandidaten für Nachfolge von
wurde in das Kontrollgremium wiedergewählt genau wie

Autoren von Texten über das komplizierte Ökosystem Wald
sind die schlimmsten Bluthunde

haben Machenschaften

anders als



Classifier Precision

Classifier Precision

- separate from study
- based on text study data
- allow deviating phrasing

Classifier Precision

- separate from study
- based on text study data
- allow deviating phrasing
- **precision** of 0.53 (0.63)

<WORD> von <ORGANIZATION>, <PERSON>
<PERSON>, <WORD> von <ORGANIZATION>
<ORGANIZATION>-<WORD> <PERSON>

CEO

Sprecher

Chef

Abgeordneter

Vorstandsvorsitzender

<WORD> wie <PERSON> und <PERSON>
<WORD>, darunter <PERSON> und <PERSON>

Wissenschaftler

Stars

Diktatoren

<WORD> von <ORGANIZATION>, <PERSON>
<PERSON>, <WORD> von <ORGANIZATION>
<ORGANIZATION>-<WORD> <PERSON>

CEO Sprecher
Chef Abgeordneter
Vorstandsvorsitzender

<WORD> wie <PERSON> und <PERSON>
<WORD>, darunter <PERSON> und <PERSON>

Wissenschaftler
Stars Diktatoren

<ORGANIZATION>, Oberst <PERSON>
<ORGANIZATION>, US-General <PERSON>

Kommandeur

<PERSON> und Gitarrist <PERSON>

Bandmitglieder

<PERSON> <WORD> <PERSON>

trifft hilft
sollte gilt

<PERSON> und <WORD> <PERSON>

Innenminister
zuletzt Jürgen

<PERSON> <WORD> <PERSON>

trifft hilft
sollte gilt

<PERSON> und <WORD> <PERSON>

Innenminister
zuletzt Jürgen

<PERSON> hat die Wahl von General <PERSON>

begrüßt Ernennung

<ORGANIZATION> <WORD> mit 35 Milliarden
Euro für die Rettung von <ORGANIZATION>

bewilligt Rettung

<ORGANIZATION>, Goldman Sachs
und <ORGANIZATION>

gemeinsames
Geschäftsfeld

reasonable precision

patterns as in literature

not predefined

language-independent

no special training

Future Work

Future Work

- normalization of names
- current sources
- exclusion of non-relations?

Future Work

- **directions** and multi-edges
- integrate existing **ontologies**
- **refine** classifier

Future Work

- infos on entities
- addition of sources
- on-the-fly timeframe change



Questions?