

Auto-IA Workshop HMTM Hannover

TEXT MINING FOR SOCIAL SCIENTISTS

2. Example studies

Gregor Wiedemann | g.wiedemann@leibniz-hbi.de

Media Research Methods Lab

Leibniz-Institute for Media Research | Hans-Bredow-Institut

Andreas Niekler | aniekler@informatik.uni-leipzig.de

Abteilung Automatische Sprachverarbeitung, Institut für Informatik
Universität Leipzig



Covered in this lecture

1. Study „Verfassungsschutz“
2. Study „Post-democracy and neoliberalism“

Text Mining Application

An exemplary study on the German „Verfassungsschutz“

Background

Foto: Bundesamt für Verfassungsschutz



„Als Frühwarnsystem hat das BfV damit mehrfach kläglich versagt“ (NSU Untersuchungsausschuss)

Landeskriminalamt Thüringen

Fahndung



Das Landeskriminalamt Thüringen bittet um Mithilfe:

Vorbereitung eines Sprengstoffverbrechens

Beate Zschäpe, Uwe Böhnhardt, Uwe Mundlos werden wegen Vorbereitung eines Sprengstoffverbrechens im Bereich Jena von der Staatsanwaltschaft Gera mit Haftbefehl gesucht.



Personalien:

Beate Zschäpe, geb. Apel
23 Jahre alt
letzter Aufenthaltsort: Jena



Personalien:

Uwe Böhnhardt
20 Jahre alt
letzter Aufenthaltsort: Jena



Personalien:

Uwe Mundlos
24 Jahre alt
letzter Aufenthaltsort: Jena

Personenbeschreibung:

- 160 cm groß
- schlank, unauffälliges Äußeres
- dunkelblondes, schulterlanges, leicht gewelltes Haar

Personenbeschreibung:

- 186 cm groß
- schlank, hager
- dunkelblondes, kurz geschorenes Haar
- stark abstehende Ohren

Personenbeschreibung:

- 180 cm groß
- schlank, sportliche Figur
- dunkelbraunes, kurz geschorenes Haar

Faksimile: Fahndungsauftrag LKA Thüringen 1997, Repro NSU-Watch

Background



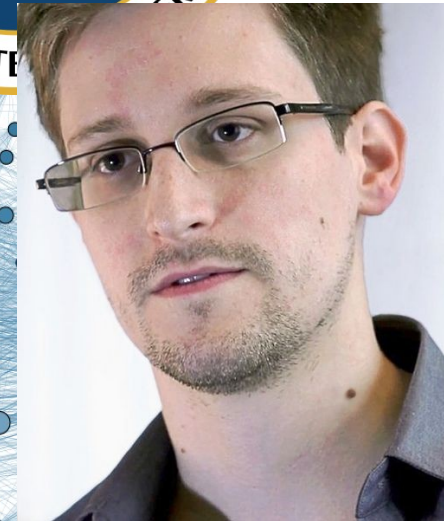
Landeskriminalamt Thüringen

Fahndung

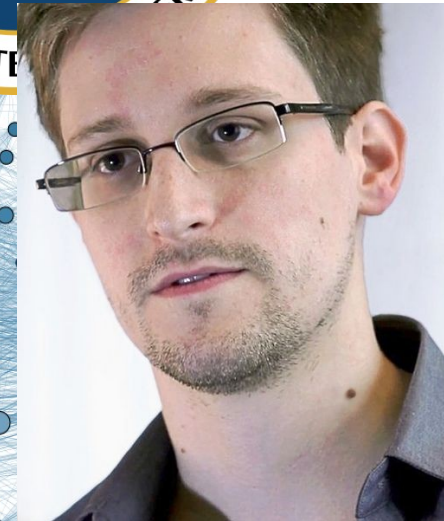
Das Landeskriminalamt Thüringen bittet um Mithilfe:
Vorbereitung eines Sprengstoffverbrechens

Beide Zecklpe, Uwe Böhnhardt, Uwe Mundlos werden wegen Vorbereitung eines Sprengstoffverbrechens im Bereich Jena von der Staatsanwaltschaft Gera mit Haftbefehl gesucht.

Personenfoto	Personenfoto	Personenfoto
Zecklpe, gpb. April 23 Jahre alt. keine Aufenthaltsort: Jena	Uwe Böhnhardt 27 Jahre alt. keine Aufenthaltsort: Jena	Uwe Mundlos 24 Jahre alt. keine Aufenthaltsort: Jena
Personenbeschreibung: <ul style="list-style-type: none">- 180 cm groß- schlank, mittelgroße Augen- dunkelbraune, schulterlangen, leicht gewellte Haar	Personenbeschreibung: <ul style="list-style-type: none">- 180 cm groß- schlank, langes- dunkelbraune, kurz geschorene Haar- stark absteigende Ohren	Personenbeschreibung: <ul style="list-style-type: none">- 180 cm groß- schlank, sportliche Figur- dunkelbraune, kurz geschorene Haar



Background



Public history of German intelligence

September 1950: Die westlichen Alliierten lockern das Besatzungsstatut und gestatten der Bundesrepublik neben der Gründung des BKA auch die Einrichtung eines Inlandsgeheimdienstes. Noch im selben Monat verabschiedet der Bundestag das längst vorbereitete Bundesverfassungsschutzgesetz (BVerfSchG).

→ Seit 1950 öffentliche Berichterstattung über den VS



Public history of German intelligence

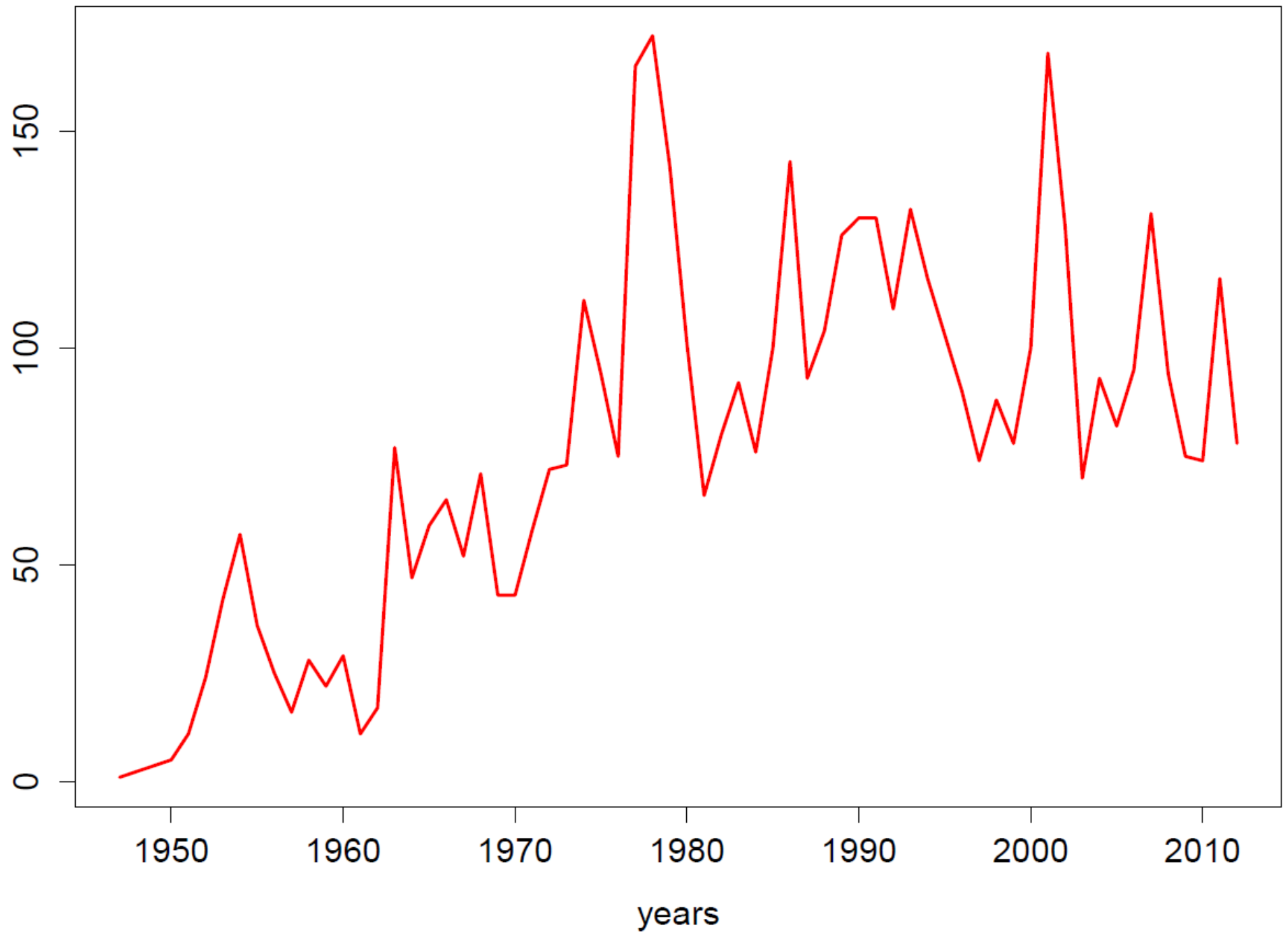
September 1950: Die westlichen Alliierten lockern das Besatzungsstatut und gestatten der Bundesrepublik neben der Gründung des BKA auch die Einrichtung eines Inlandsgeheimdienstes. Noch im selben Monat verabschiedet der Bundestag das längst vorbereitete Bundesverfassungsschutzgesetz (BVerfSchG).

→ Seit 1950 öffentliche Berichterstattung über den VS

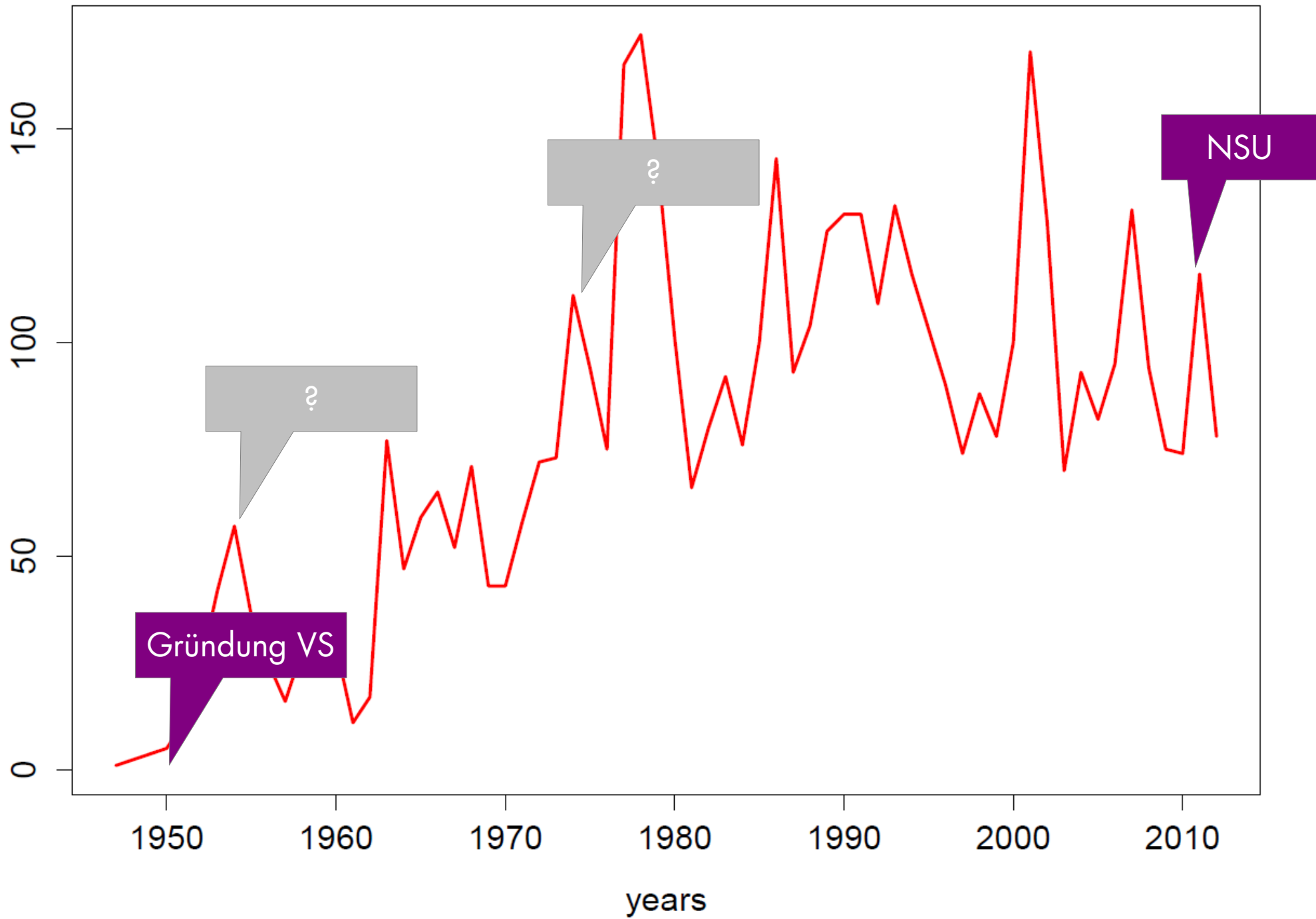


	Zeitraum	#Doks	Tokens
Spiegel	1950 - 2011	2.912	
Zeit	1950 - 2011	2.166	
		5.078	2.817.607

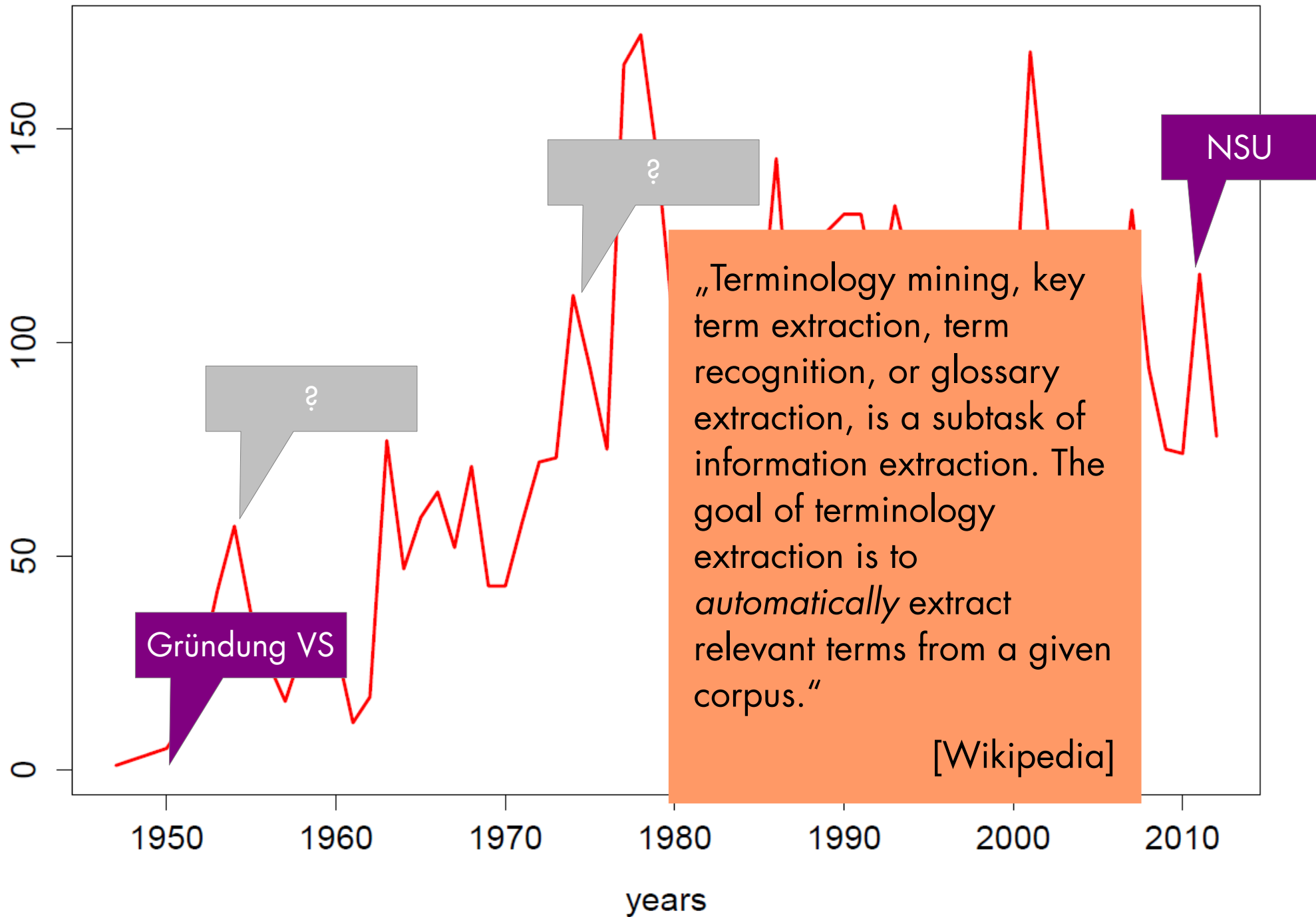
vs documents



vs documents



vs documents



Key term extraction (Frequency)



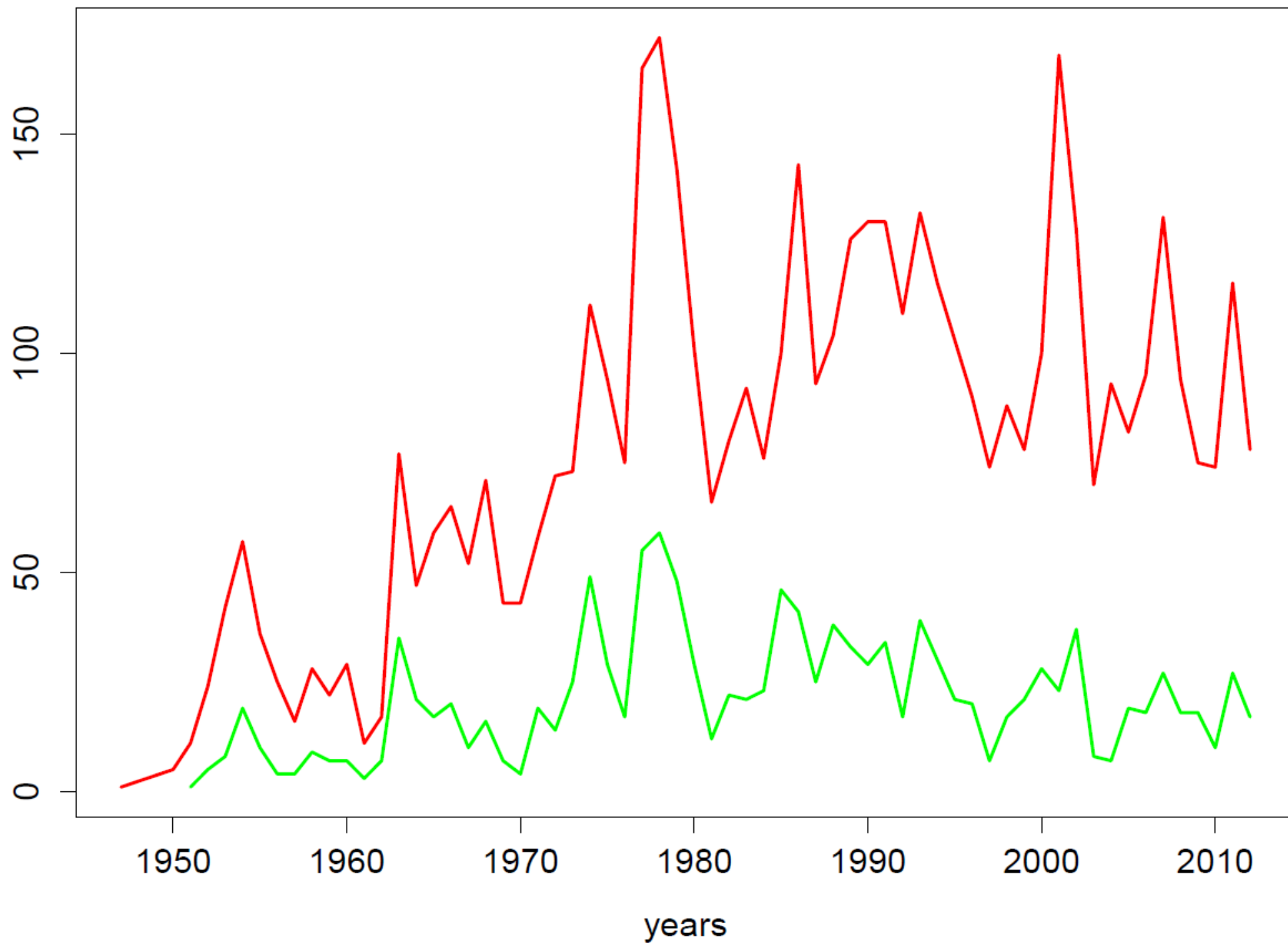
Key term extraction (reference corpus)



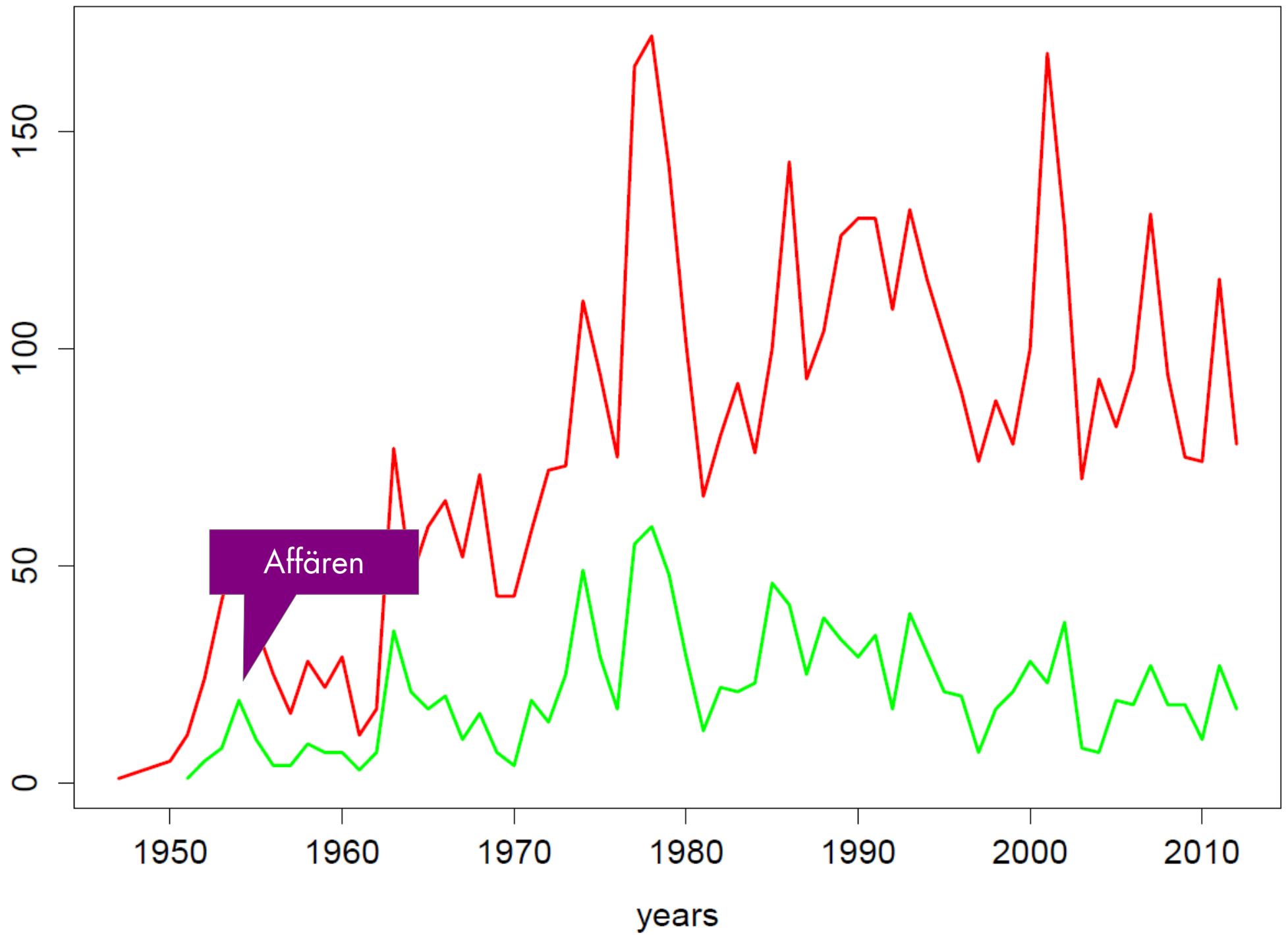
Key term extraction (regular expressions)

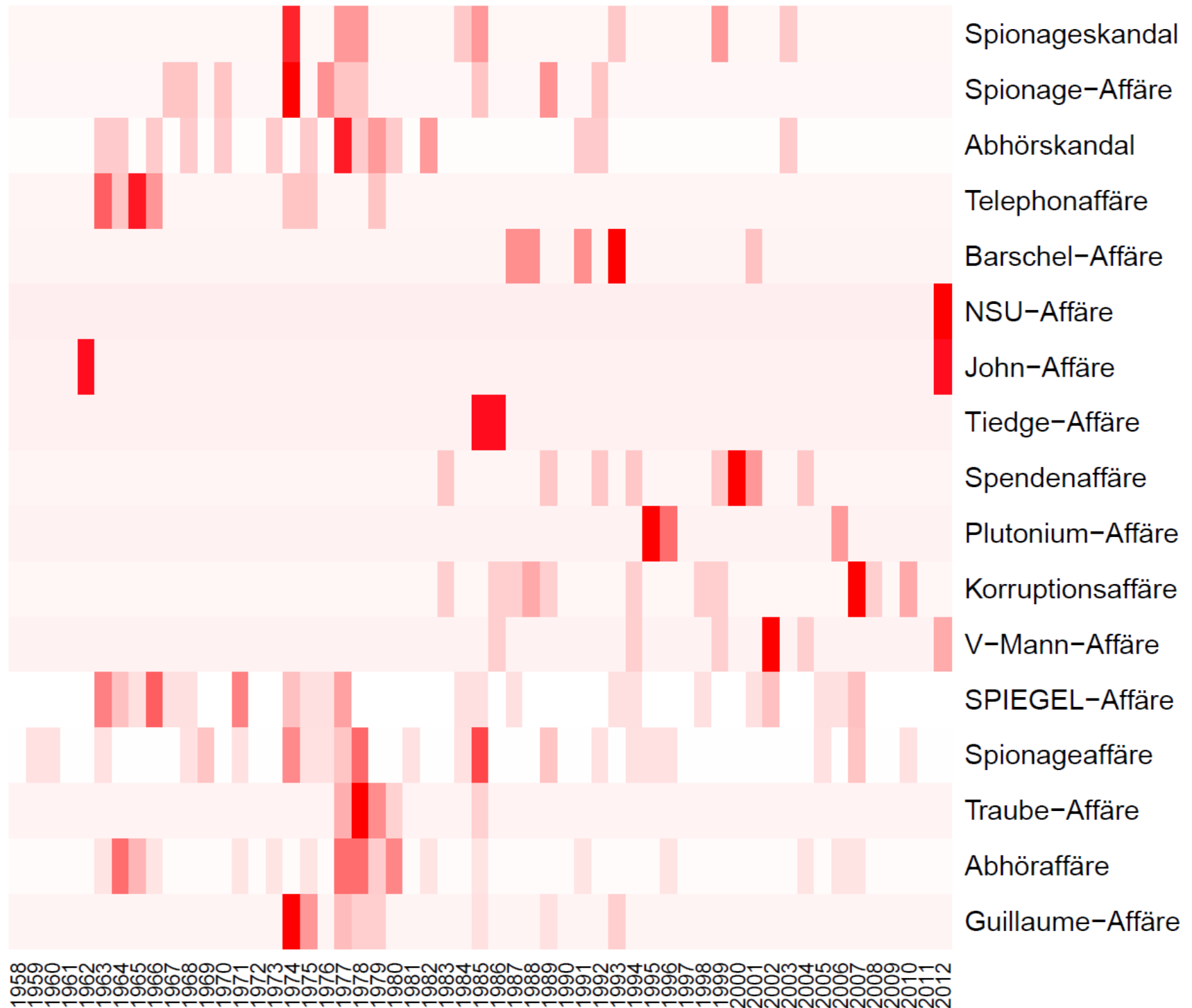
- "skandal | affäre | ^panne":
 - Affäre, Skandal, Panne, skandalös, SPIEGEL-Affäre, Abhör-affäre, Spionageaffäre, Guillaume-Affäre, Abhör-Affäre, Staatsaffäre, Abhörskandal, Korruptionsaffäre, Plutonium-Affäre, Spionage-Affäre, Spionageskandal, Spendenaffäre, Traube-Affäre, V-Mann-Affäre, Spiegel-Affäre, Telephonaffäre, Barschel-Affäre, Flick-Affäre, Parteispendenaffäre, -Affäre, skandalträchtig, Geheimdienstskandal, Justizskandal, Agentenaffäre, Korruptionsskandal, U-Boot-Affäre, Telephon-Affäre, Fibag-Affäre, Kieler Affäre, Finanzaffäre, Telephon-Abhör-affäre, Liebesaffäre, Kießling-Affäre, Verfassungsschutz-Affäre, Polizeiaffäre, skandalträchtigen, Spitzelaffäre, Spielbankenaffäre, Lauschaaffäre Traube, CDU-Spendenaffäre, Telephonabhör-Affäre, Geheimdienst-Affäre, Pannenserie, Geheimdienst-Affären, Bestechungsskandal, Steiner-Affäre, Spielbanken-Affäre, Panzeraffäre, Skandalfigur, Abhör-Skandal, BND-Affäre, Spitzel-Affäre, Skandalminister, Bespitzelungsaffäre, BND-Skandal, Geheimdienstaffäre, Skandalös, Watergate-Affäre, Parteispenden-Affäre, Steueraffäre, Spendenskandal, Bauskandal, skandalumwittert, Politskandal, Amigo-Affäre, Bankenaffäre, Polizeiskandal, MAD-Affäre, V-Mann-Skandal...

vs documents

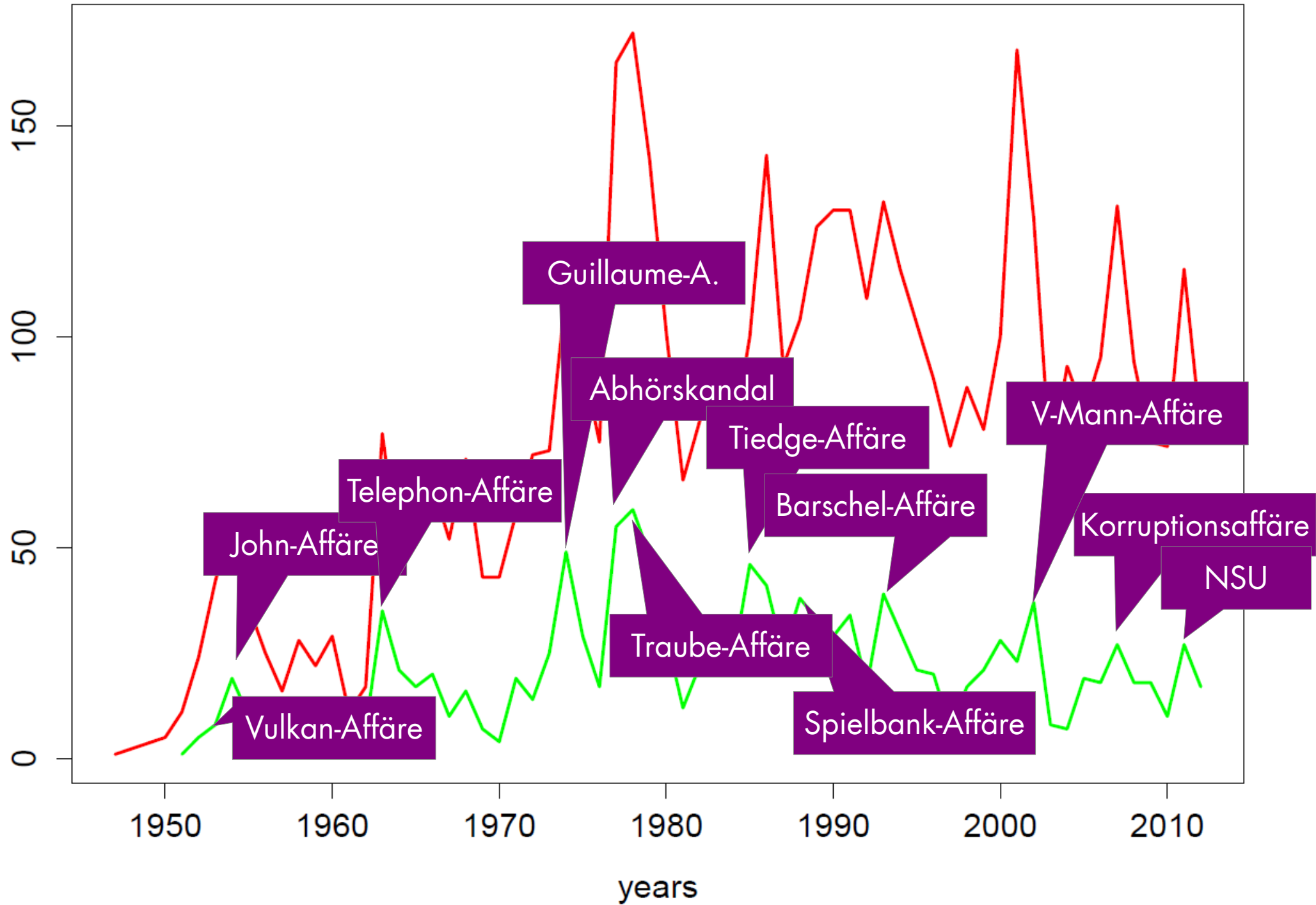


vs documents

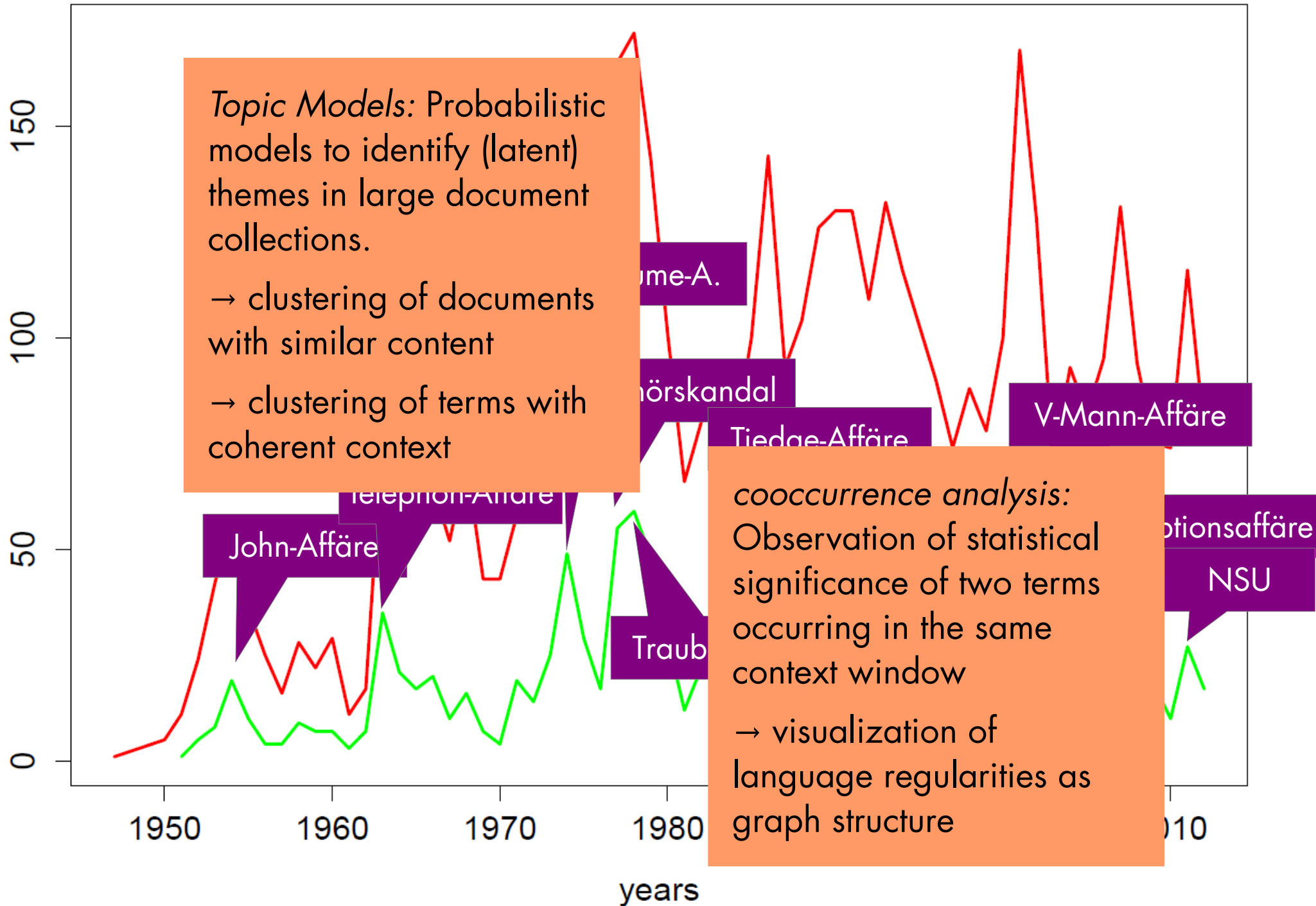


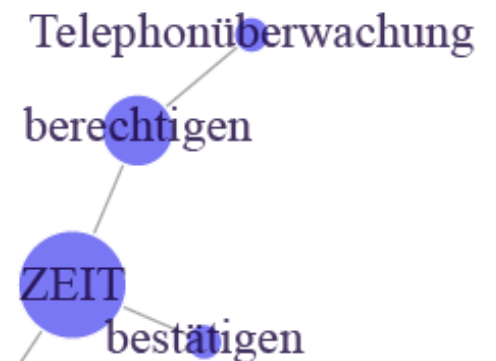
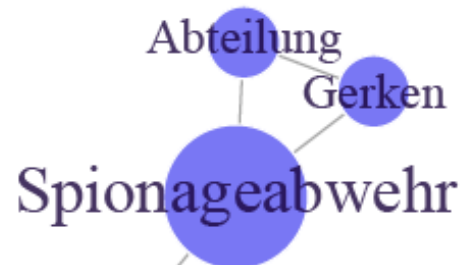
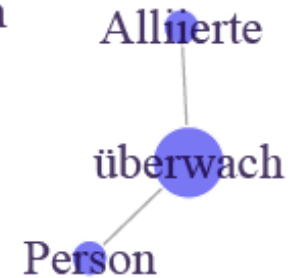
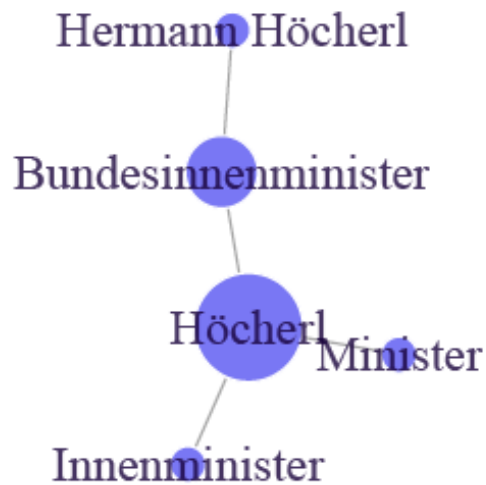
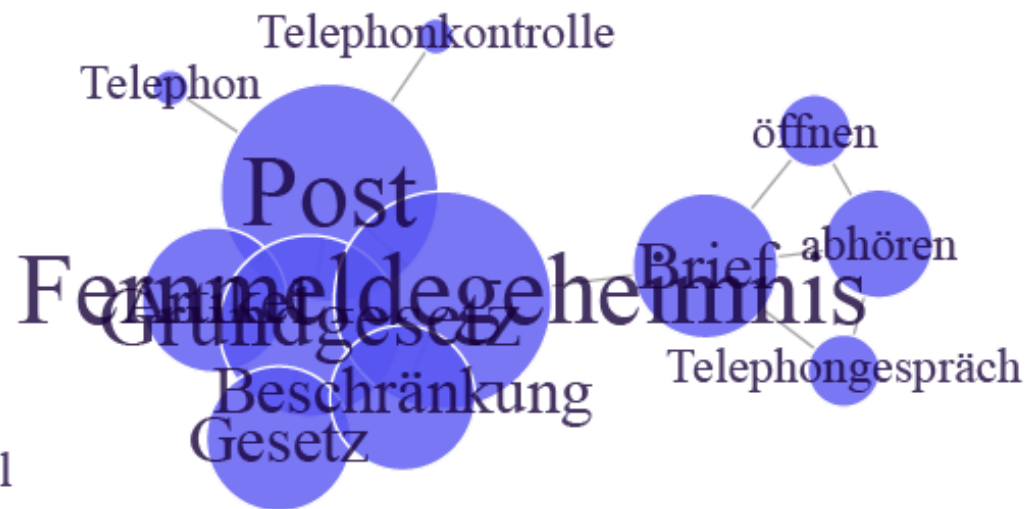


vs documents

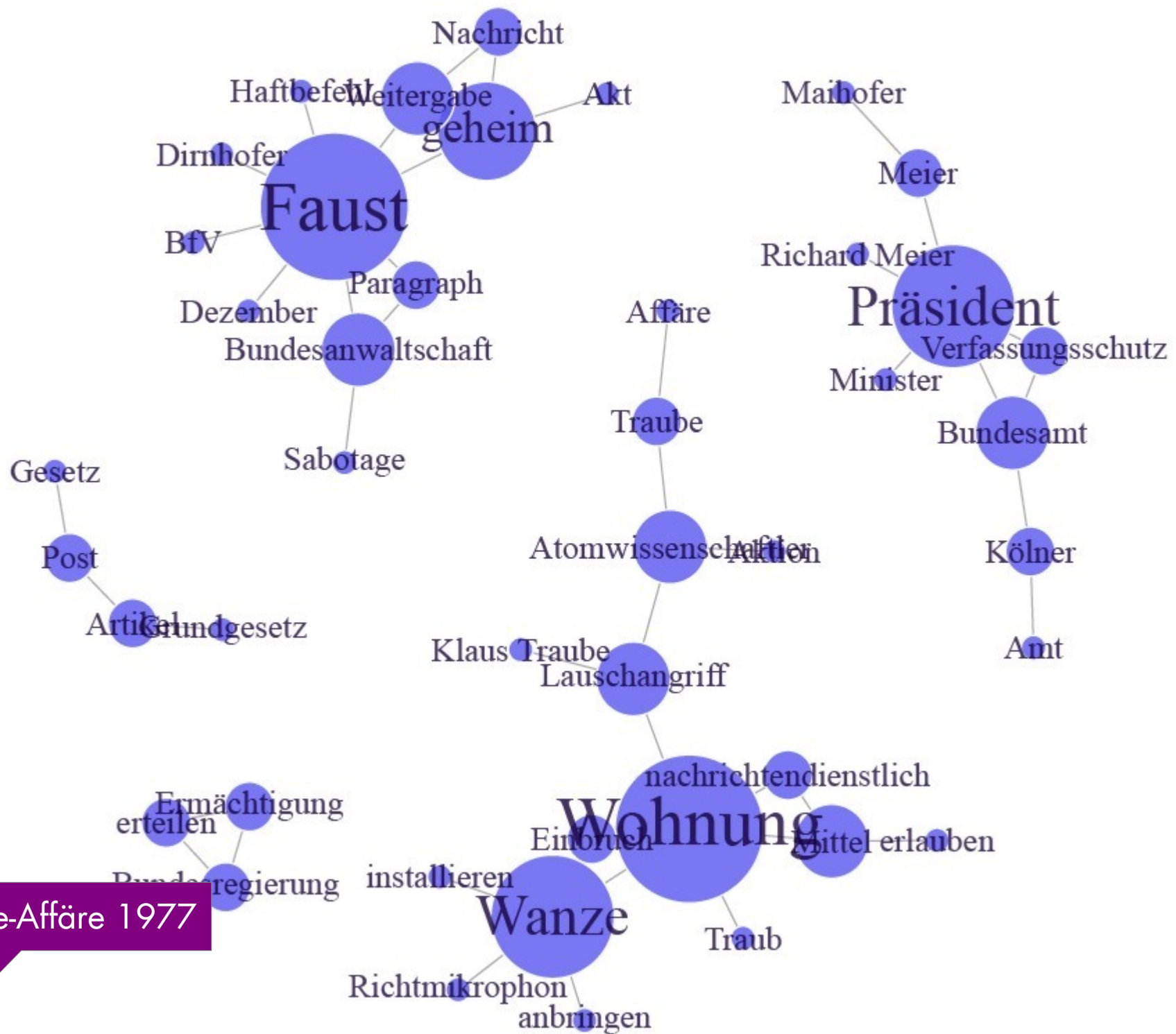


vs documents

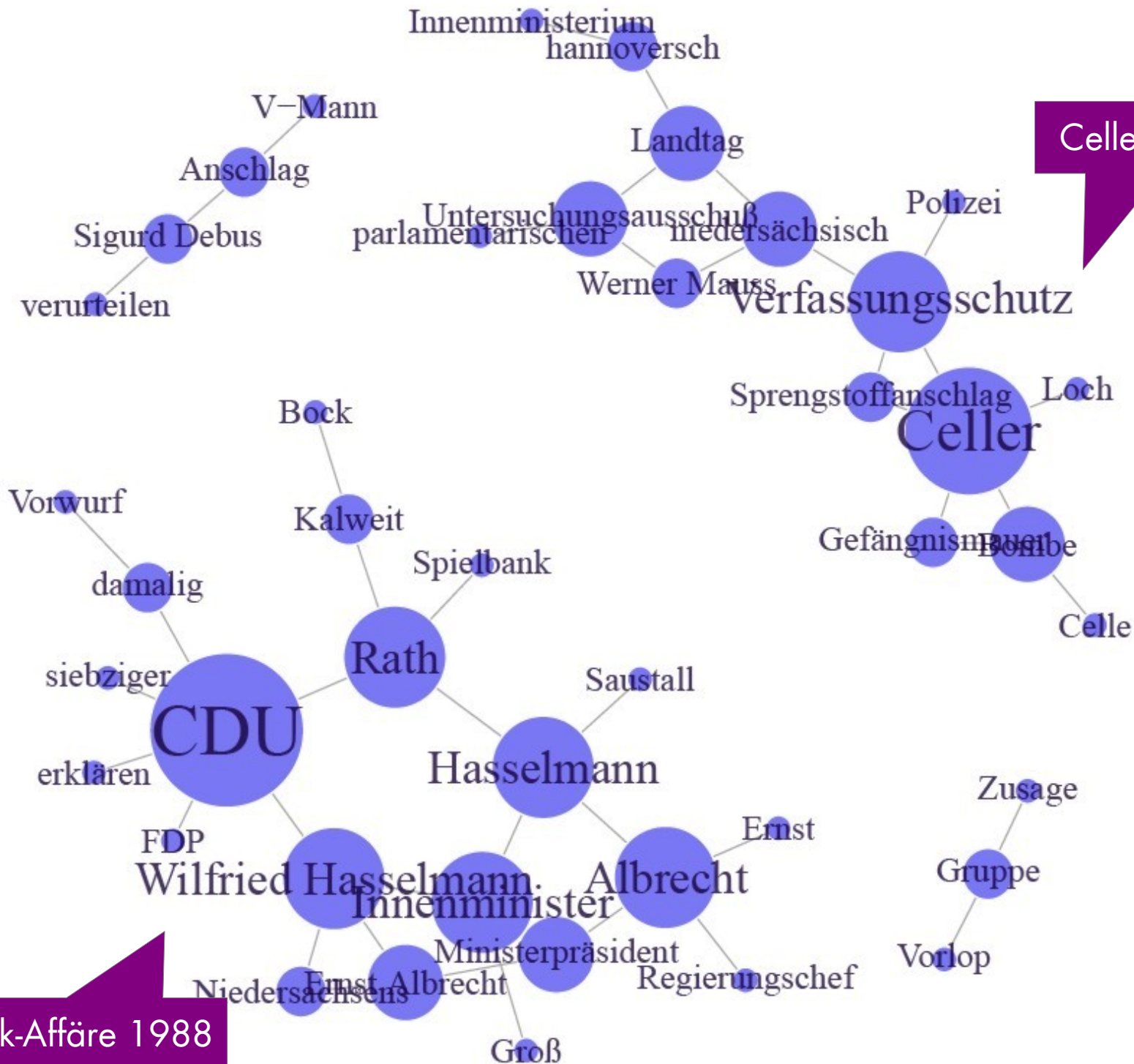




Telephon-Affäre 1963

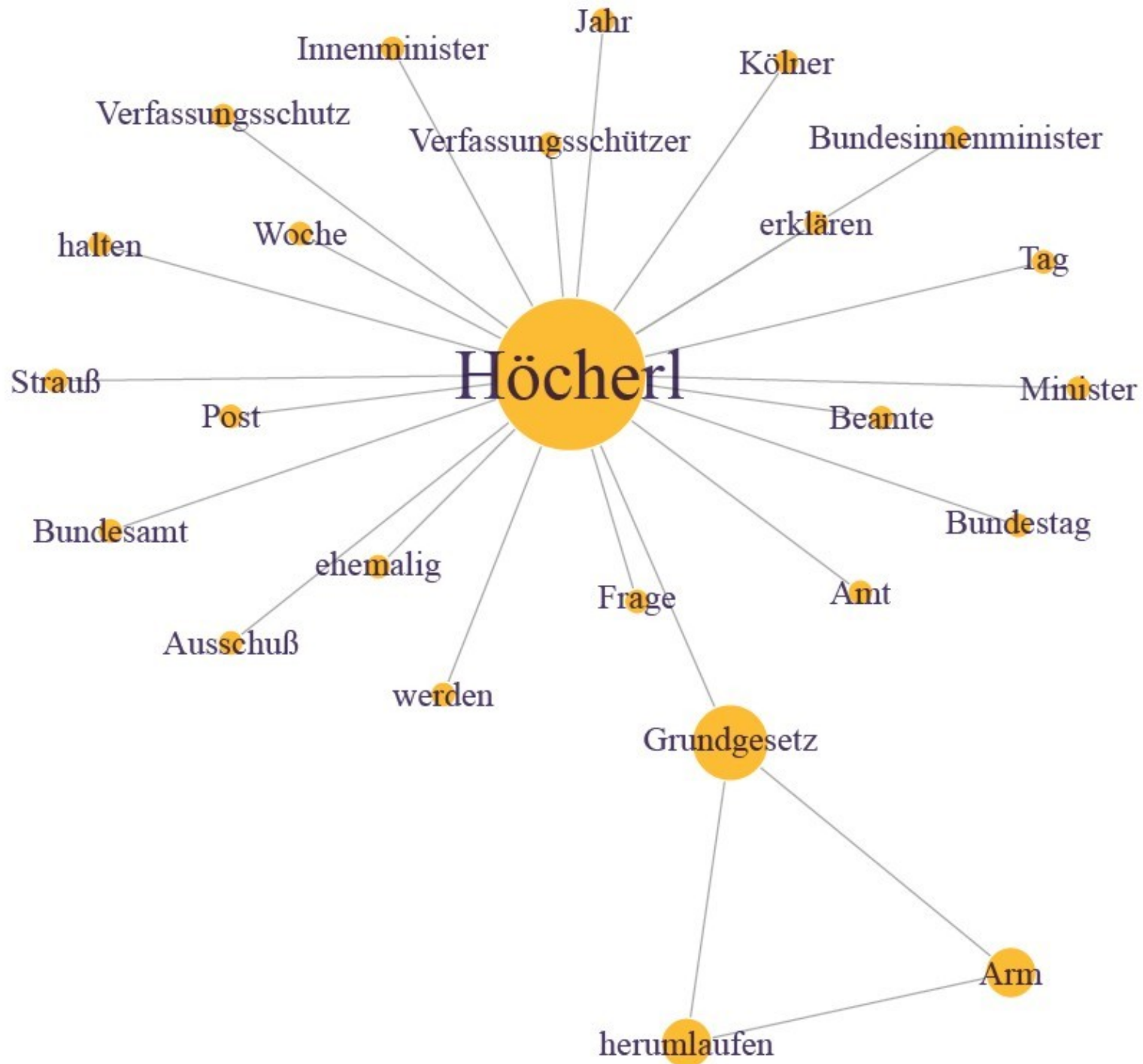


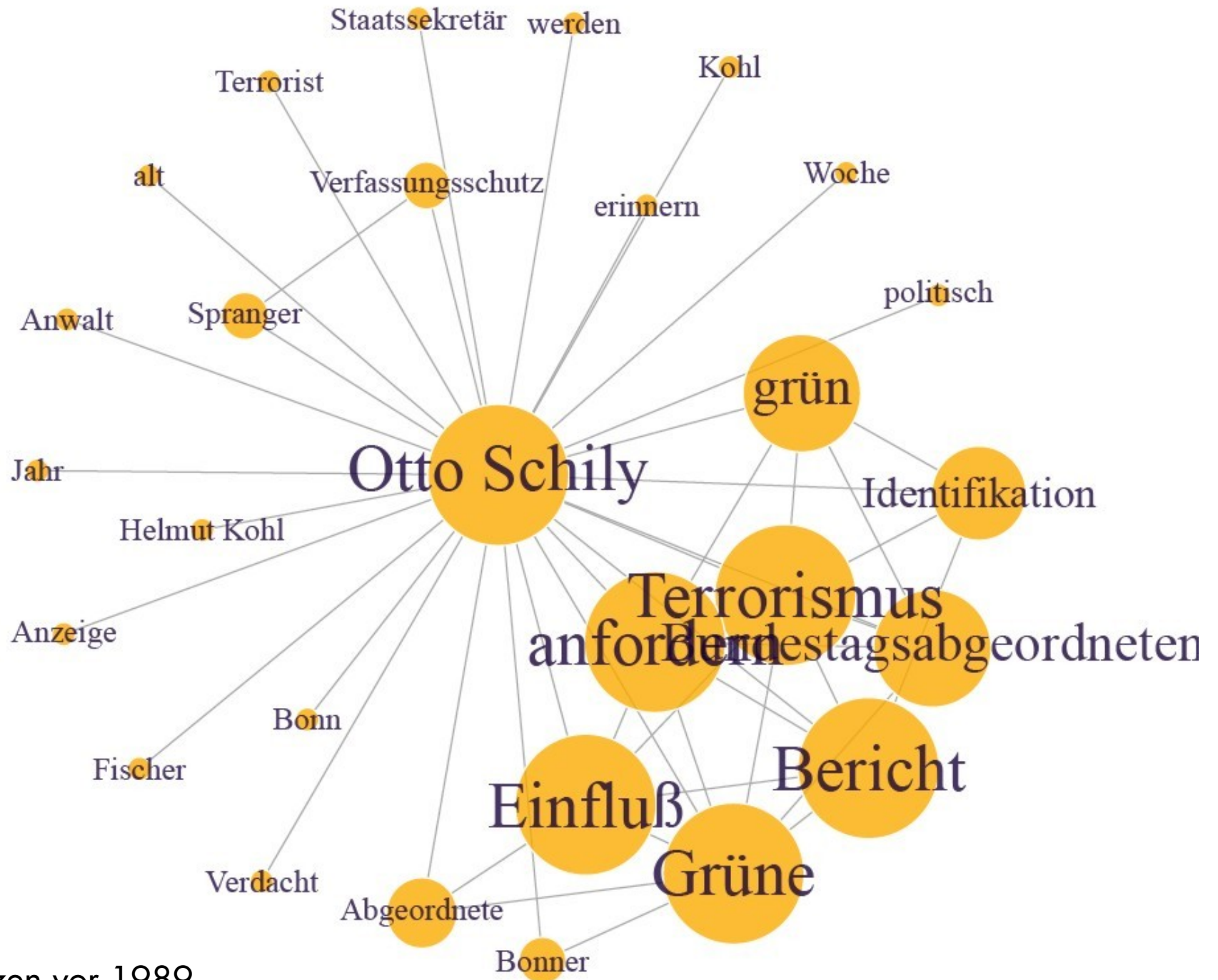
Traube-Affäre 1977



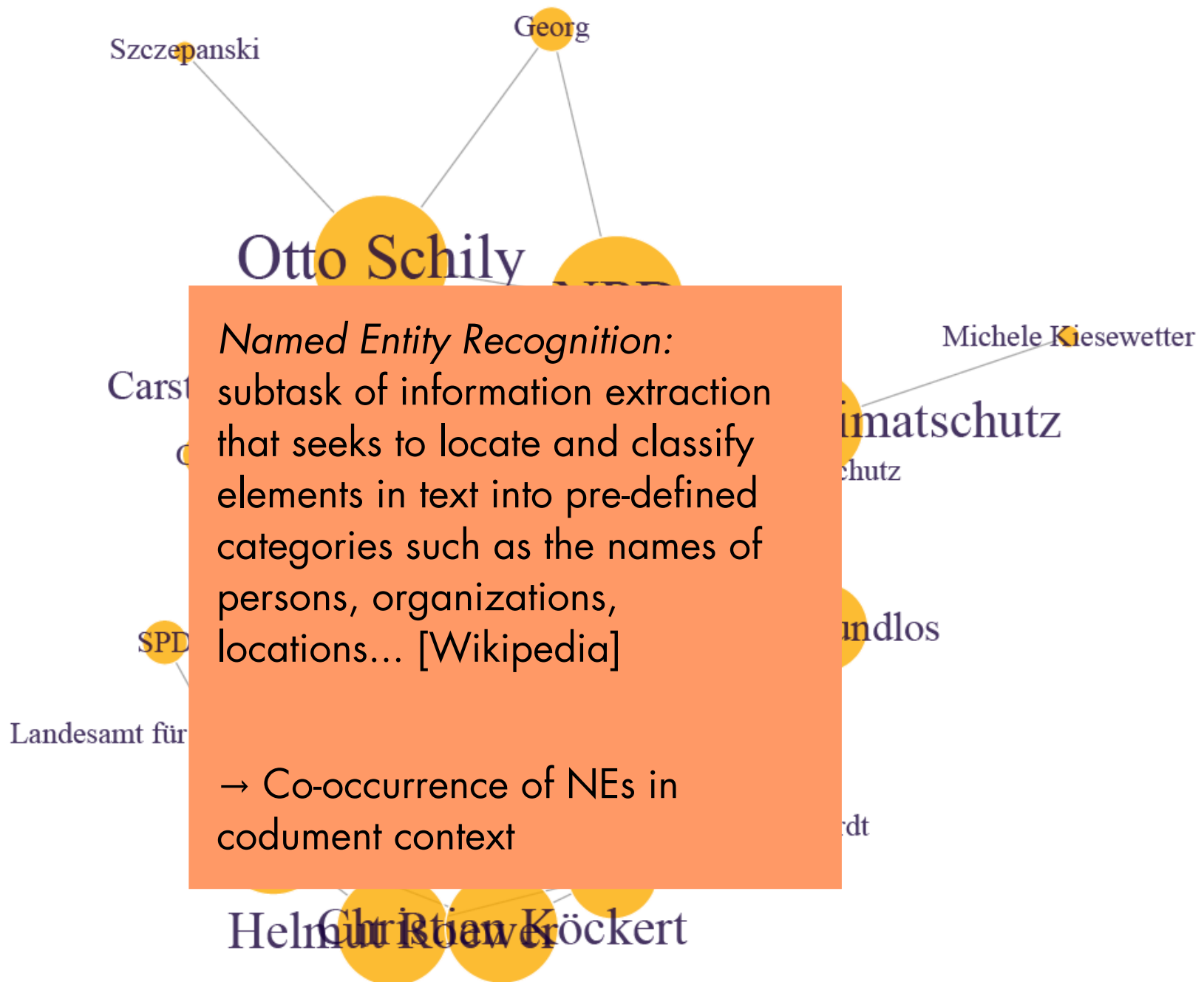
Celler Loch 1986

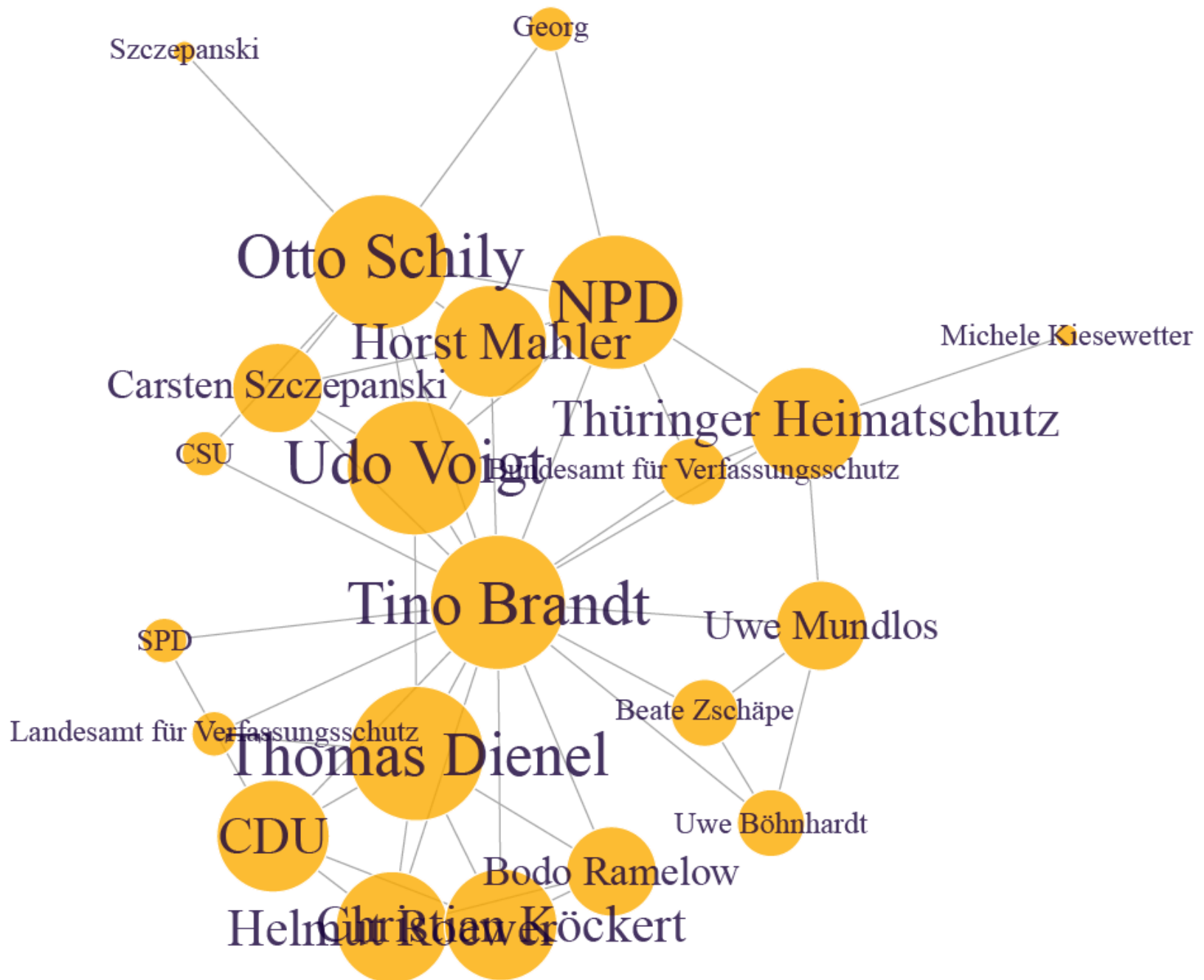
Spielbank-Affäre 1988

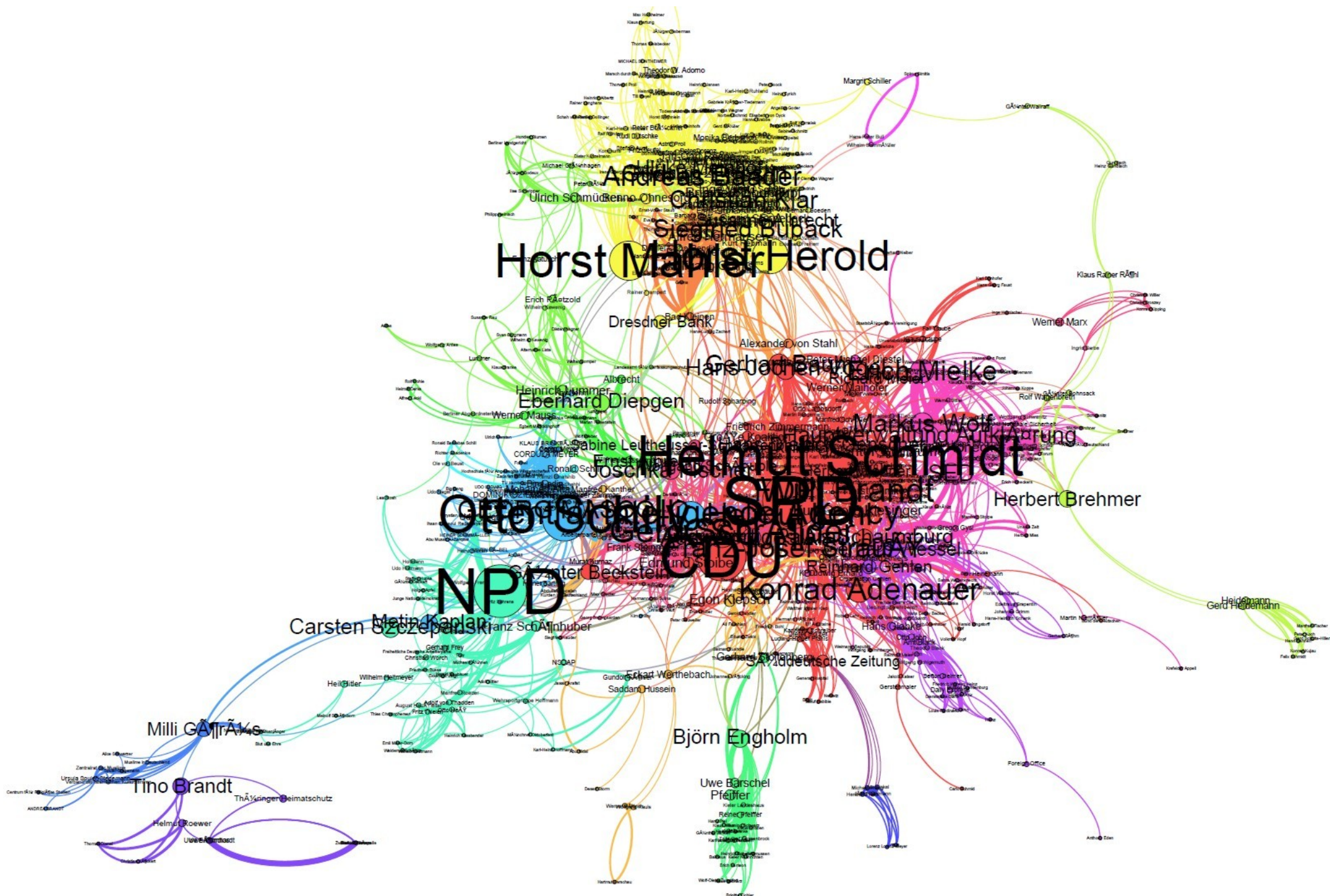




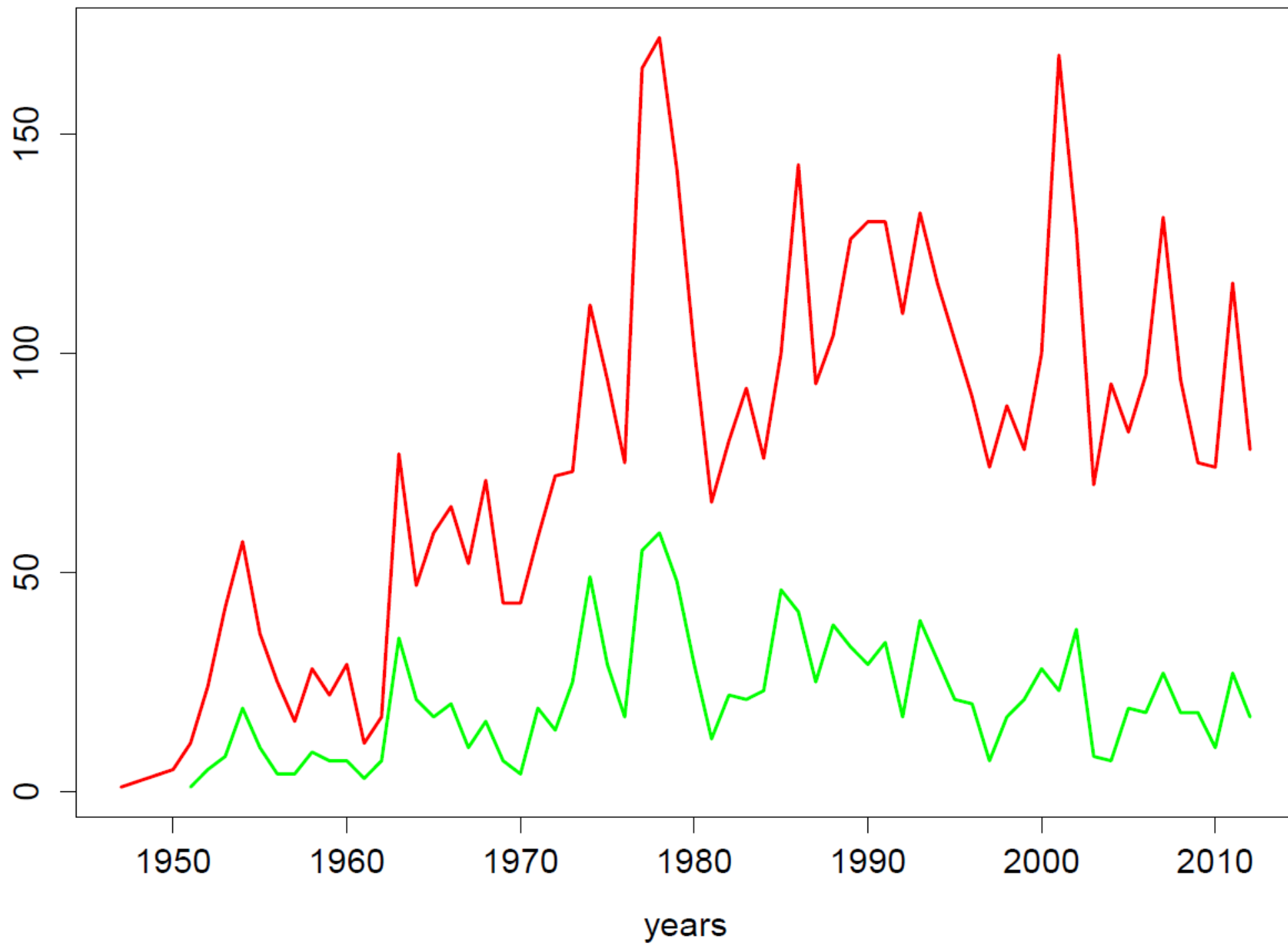
Kookkurrenzen vor 1989







vs documents



Reform

„Als Frühwarnsystem hat das BfV damit mehrfach kläglich versagt“ (NSU Untersuchungsausschuss)

- Markus Ulbig (Sächs. Innenminister) bei der Vorstellung des Berichts der „Expertenkommission zur Evaluierung des Sächsischen Landesamtes für Verfassungsschutz“ am 20. Februar 2013
 - „Wir wollen einen Philosophiewechsel beim Verfassungsschutz. Unser Verfassungsschutz soll modern und serviceorientiert sein. Der Gesamtkomplex NSU hat uns allen gezeigt, dass es Verbesserungsbedarf im Verfassungsschutz in Deutschland gibt.“
- Ulbig am 25. Juni 2013 anlässlich der Vorstellung des sächsischen VS-Berichtes 2012
 - „Guter Verfassungsschutz ist das Fernlicht der Sicherheitsbehörden – da wollen wir hin.“
- Ulbig am 5. Mai 2014 anlässlich der Vorstellung des sächsischen VS-Berichtes 2013
 - „Der neue Bericht ist Beweis für den erfolgreichen Philosophiewechsel im Verfassungsschutz: Das Landesamt für Verfassungsschutz ist ein Fernlicht für eine wehrhafte Demokratie.“

Reform

„Als Frühwarnsystem hat das BfV damit mehrfach kläglich versagt“ (NSU Untersuchungsausschuss)

- Markus Ulbig (Sächs. Innenminister) bei der Vorstellung des Berichts der „Expertenkommission zur Evaluierung des Sächsischen Landesamtes für Verfassungsschutz“ am 20. Februar 2013
 - „Wir wollen einen Philosophiewechsel beim Verfassungsschutz. Unser Verfassungsschutz soll modern und serviceorientiert sein. Der Gesamtkomplex NSU hat uns allen gezeigt, dass es Verbesserungsbedarf im Verfassungsschutz in Deutschland gibt.“
- Ulbig am 25. Juni 2013 anlässlich der Vorstellung des sächsischen VS-Berichtes 2012
 - „Guter Verfassungsschutz ist das Fernlicht der Sicherheitsbehörden – da wollen wir hin.“
- Ulbig am 5. Mai 2014 anlässlich der Vorstellung des sächsischen VS-Berichtes 2013
 - „Der neue Bericht ist Beweis für den erfolgreichen Philosophiewechsel im Verfassungsschutz: Das Landesamt für Verfassungsschutz ist ein Fernlicht für eine wehrhafte Demokratie.“

Wird jetzt also alles gut?

Reform

- „In dieser Lage hat der Bundesinnenminister ein paar richtungsweisende Reformen in Gang gesetzt. Die Dienste sollen sich auf reale Gefahren konzentrieren, nicht auf Ideen. Sie sollen zur Zusammenarbeit gezwungen werden [...]. Und schließlich soll das Parlament mehr Einblick bekommen“
- „Reform beim Verfassungsschutz Nach der Reform des Bundesnachrichtendienstes treibt Innenstaatssekretär August Hanning nun auch den Umbau des Bundesamts für Verfassungsschutz (BfV) voran. Das Amt solle "zum zentralen Inlands-Nachrichtendienst-Kompetenzzentrum" ausgebaut werden“
- „Seit [... dem Vorfall] ist bereits einiges an Verbesserungen eingeleitet worden, was die Zahl der unvermeidlichen Pannen verhindern müßte. Vor allem gibt es [...] eine zentrale EDV-Kartei der Nachrichtendienste, die das frühere System wesentlich verbessert“
- „Der Bundestag hat bereits zur Zeit der Großen Koalition für eine solche Reform Vorarbeiten geleistet. Ein Ausschuß aller drei Parteien [...] hat gefordert, die Bundesregierung solle einen koordinierenden Beauftragten für die drei Nachrichtendienste ernennen“
- „Lücke will außerdem die Richtlinien überarbeiten und ergänzen, mit denen die Zusammenarbeit der Verfassungsschutzämter von Bund und Ländern, des militärischen Abschirmdienstes, des Bundesnachrichtendienstes, der Polizei und der Staatsanwaltschaften geregelt werden soll.“
- „Er schlug ihnen vor, eine allgemeine Reorganisation des Systems der Verfassungsschutzämter des Bundes und der Länder vorzunehmen, und zwar in der Form, daß der Bund auch die Befugnis erhält, die Personalpolitik bei den Verfassungsschutzämtern der Länder entscheidend zu beeinflussen.“

Reform

2012

- „In dieser Lage hat Bundesinnenminister ein paar richtungsweisende Reformen in Gang gesetzt. Die Dienste sollen sich auf reale Gefahren konzentrieren, nicht auf Ideen. Sie sollen zur Zusammenarbeit gezwungen werden [...]. Und schließlich soll das Parlament mehr Einblick bekommen“

2007

- „Reform beim Verfassungsschutz: Nach der Reform des Bundesnachrichtendienstes treibt Innenstaatssekretär August Hanning nun auch den Umbau des Bundesamts für Verfassungsschutz (BfV) voran. Das Amt solle "zum zentralen Inlands-Nachrichtendienst-Kompetenzzentrum" ausgebaut werden“

1974

- „Seit [... dem Vorjahr] ist bereits einiges an Verbesserungen eingeleitet worden, was die Zahl der unvermeidlichen Pannen verhindern müßte. Vor allem gibt es [...] eine zentrale EDV-Kartei der Nachrichtendienste, die das frühere System wesentlich verbessert“

1974

- „Der Bundestag hat in der Zeit der Großen Koalition für eine solche Reform Vorarbeiten geleistet. Ein Ausschuß aller drei Parteien [...] hat gefordert, die Bundesregierung solle einen koordinierenden Beauftragten für die drei Nachrichtendienste ernennen“

1966

- „Lücke will außer Reichweite überarbeiten und ergänzen, mit denen die Zusammenarbeit der Verfassungsschutzämter von Bund und Ländern, des militärischen Abschirmdienstes, des Bundesnachrichtendienstes, der Polizei und der Staatsanwaltschaften geregelt werden soll.“

1954

- „Er schlug ihnen eine Reorganisation der Verfassungsschutzämter des Bundes und der Länder vorzunehmen, und zwar in der Form, daß der Bund auch die Befugnis erhält, die Personalpolitik bei den Verfassungsschutzämtern der Länder entscheidend zu beeinflussen.“

Reform

2012

- „In dieser Lage hat der Bundesinnenminister ein paar richtungsweisende Reformen in Gang gesetzt. Die Dienste sollen sich auf reale Gefahren konzentrieren, nicht auf Ideen. Sie sollen zur Zusammenarbeit gezwungen werden [...]. Und schließlich soll das Parlament mehr Einblick bekommen“

2007

- „Reform beim Verfassungsschutz Nach der Reform des Bundesnachrichtendienstes treibt

1974

- „Seit [... dem Vorjahr] ist bereits einiges an Verbesserungen eingeleitet worden, was die Zahl vermeidlichen Pannen ern müsste. Vor allem gibt eine zentrale EDV-Kartei Nachrichtendienste, die das System wesentlich ert“

Klassifikation:

Anwendung des überwachten maschinellen Lernens. Anhand von Trainingsdaten (hier: codierte Textstellen) werden unbekannte Daten in Kategorien sortiert

→ semi-automatische Identifikation ähnlicher Textstellen

1974

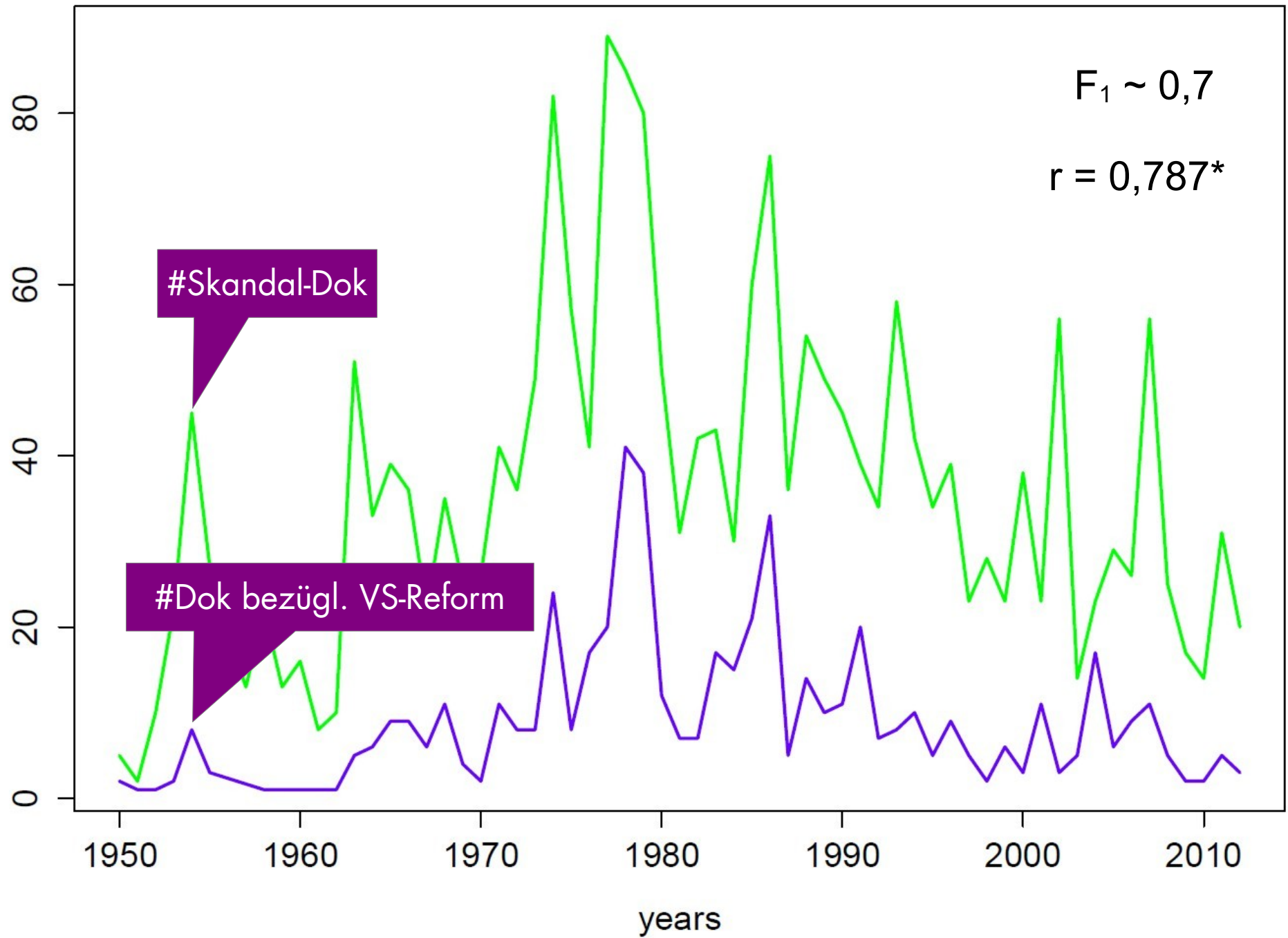
- „Der Bundestag hat in der Zeit der Großen Koalition für eine solche Reform Vorarbeiten geleistet. Ein Ausschuß aller drei Parteien [...] hat gefordert, die Bundesregierung solle einen koordinierenden Beauftragten für die drei Nachrichtendienste ernennen“

1954

geregelt werden soll.“

zug ihrer Organisation der Nachrichtenschutzämter des Bundes Länder vorzunehmen, und der Form, daß der Bund Befugnis erhält, die Politik bei den Nachrichtenschutzämtern der Länder entscheidend zu beeinflussen.“

VS Reform

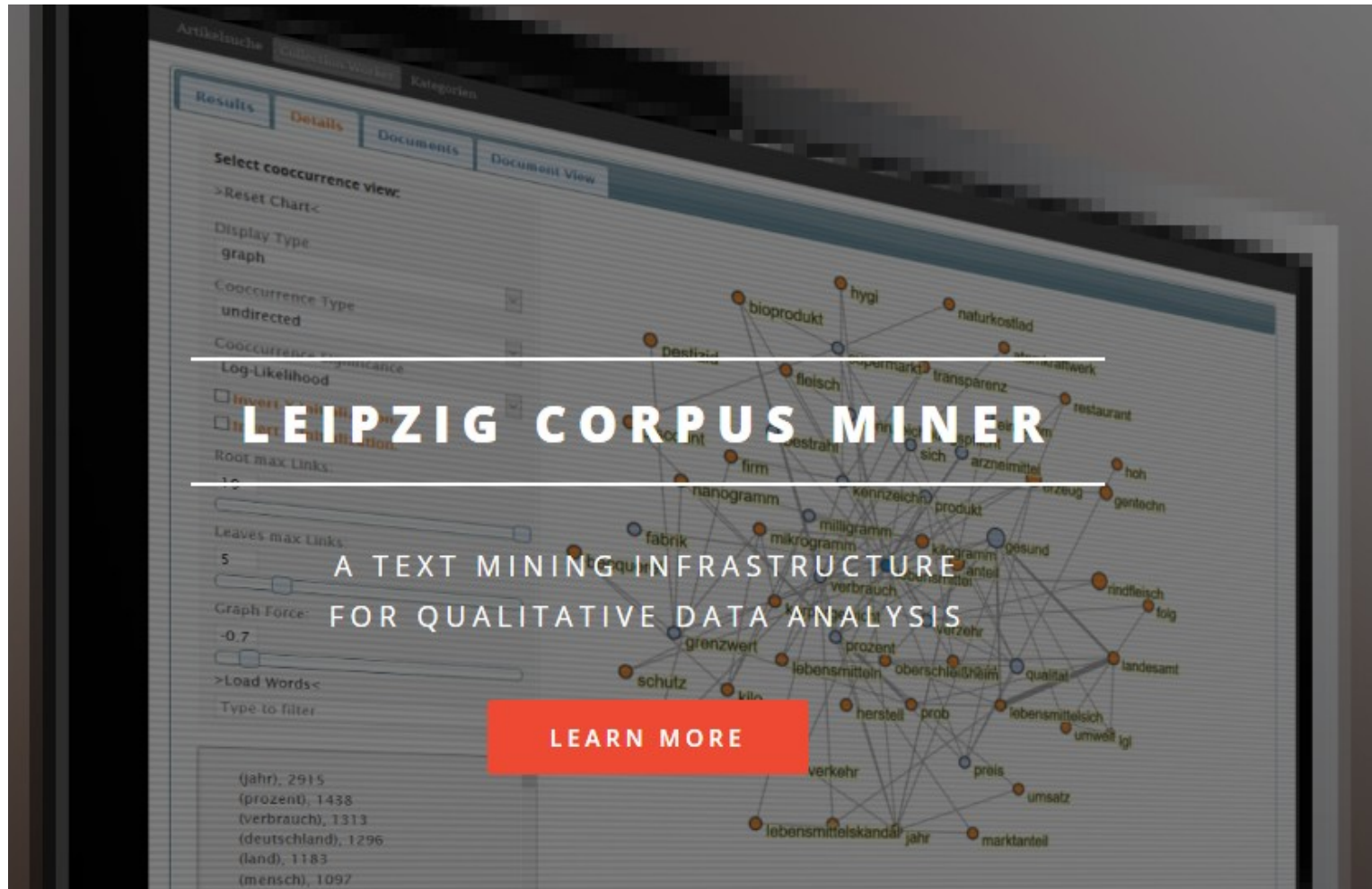


Summary

- Text Mining kann helfen große Dokumentmengen effizient zu explorieren
 - Unterstützung induktiver Forschungsschritte
- Klassifikation / Kategorie-Subsumption
 - Unterstützung deduktiver Forschungsschritte
- Methods
 - Key Terms
 - Kookkurrenzen
 - Topic Modelle
 - Named Entities
 - Klassifikation

Software Demonstrations

Leipzig Corpus Miner



- <http://ilcm.informatik.uni-leipzig.de>