

MIKE KESTEMONT (UNIVERSITY OF ANTWERP)  
GUNTHER MARTENS (GHENT UNIVERSITY)  
THORSTEN RIES (GHENT UNIVERSITY)

A CHALLENGE FOR STYLOMETRY AND  
AUTHORSHIP ATTRIBUTION METHODS:  
GOETHE'S CONTRIBUTIONS TO THE  
*FRANKFURTER GELEHRTE ANZEIGEN*  
1772/73



Frankfurter  
gelehrte  
Anzeigen.

vom Jahr 1772.



Frankfurt am Main  
bey den Eichenbergischen Erben

# The Problem

- J. W. Goethe was **editor-in-chief** of the *Frankfurter Gelehrte Anzeigen* 1772 (J. C. Deinert bought the *Frankfurter Gelehrtenzeitung* in 1771)
- *Frankfurter gelehrte Anzeigen* (FgA) became a **flagship of the “Sturm und Drang”** movement in this year because of its editors **Goethe, Johann Heinrich Merck** and **Johann Georg Schlosser** and contributors such as **Johann Gottfried Herder**.
- Goethe most likely also wrote for the FgA in **1773**.
- The “Rezensionen” (i.e. articles) of the FgA were **published anonymously** and have often been **redacted** by the editors, some have even been written **collaboratively**.

# The Problem

- **900 pages of anonymous journal text**, authorship of a lot of the nearly **400 Rezensionen** by around **40 contributors** of 1-7pp length is unclear or authorship attribution relies on shaky hermeutic arguments.
- It is in many cases unclear which have been **penned by Goethe**, which ones were heavily **redacted** by him, and which ones have been **collaborative “protocol reviews”**.
- His self-attribution of some FgA-Rezensionen in his self-edited edition of his works is regarded as **unreliable** (1772/73: 35).
- The majority of texts have **not been tested systematically** at all.

# The Project

- **Computational stylometrics and authorship attribution**
  - Burrows's *Delta*
  - Mike Kestemont (stylo, imposters method).
- **Check all Rezensionen in the FgA 1772/73** with computational stylometric methods to verify whether Goethe wrote them or not.
- **The advantages**
  - **statistical method to detect the stylistic footprint**
    - **tested on large corpora and was trained on large corpora by the authors.**
    - Imposters Method: **very good accuracy**
- We hope to **attribute new texts** to Goethe, correct previous **false positives**, a **new scientific foundation** to previous correct authorship attributions (**poorly tested, small data basis**).

## Early Approaches

- **Max Morris** and **Hermann Bräuning-Octavio** endowed large parts of their academic life to this authorship attribution question, gathering all philological evidence, producing several 800 pp thick monographs. **Otto Trieloff** and **Wilhelm Scherer** joined the conversation.
- **Rather vague notions of style and thematic preference, hermeneutic arguments, recurrence of opinions and topics, individual spelling characteristics (“warrlich”, “Shäckespear”)**. Prose rhythm (Karl Marbe, 1904, 1912, without success).
- Statistical, stylometrical, stylistic approaches have been tried at **small scale, on small samples with a very limited basis and methodological foundation.**

## Early Approaches

Specific for Goethe?

“Schäckespear”

καὶ νεκροῦστέσ εὐνοῖα --  
(berührt nicht Verwesung sein Fleisch; nagt nicht ein Wurm an ihm; denn für ihn sorgen die seligen Götter auch nach dem Tode.)

## Danzig.

Cymbelline, ein Trauerspiel, nach einem von **Schäckespear** erfundnen Stoffe.

Der Verfasser, da er sich, laut dem Vorbericht, nach einer schweren Krankheit aller ermüdenden Arbeiten enthalten musste, beschäftigte sich mit Schäckespears Werken. Das hätten wir ihm nun gleich sagen wollen, war für einen Konvaleszenten keine Lektüre. Wer an dem Leben, das durch Schäckespears Stücke glüht, Theil nehmen will, muß an Leib und Seele gesund seyn. Da bedauereten nun der Herr Verfasser aus innigem Gefühl einer fühlen, schwächlichen, kritischen Sittigkeit, die viele incongruités, durch die (wie der treffliche Johnson ad hoc drama gleichfalls bemerkt hat) many just sentiments, und einige Schönheiten, zu

# Statistical and Stylistic Approaches

- **Bräuning-Octavio 1966: set of ‘typical features’ of Goethe’s style;** language **rhythm and melody, favourite expressions, rhetorical features** such as (vague) specifics of exclamation, questions, address, double negation, accumulation and enumeration, anaphora, parenthesis, typical Rezensionen beginning, Goethe’s grammar during the “Werther Periode”, sentences omitting verb, parallelisms, inversion, emphatic sentence endings, latin quotes etc.
- **The results - beyond the direct philological proof found - remained vague.**
- But Bräuning-Octavio already worked on a **prototype of stylometrics**, as his private archive collection in the Archive of the Technische Universität Darmstadt shows

# In Bräuning-Octavio's archive

“Statistik der Füllwörter in den FGA”

*Statistics of the expletives in the FgA*

Bräuning-Octavio, Hermann  
(1889-1977)

Dr.phil., St.R., Schriftsteller

1. Nachlaß, zu s. Lebzeiten (1972) an die Stadt Darmstadt u. (kleiner Teil) an die Bibl. verkauft. Das Eigentum d. Stadt als Dauerleihgabe bei uns deponiert.  
(Vertragl. Unterlagen wie auch grobes Verzeichnis s. Extraordner H 200)

Fach 24, 1-4

24,1: Repertorium Wittich-Archiv. Ahnentafel Claus Wittich Vorfahren Wittich. - Ahnentafel Justus von Liebig. Buchdruck in Ffm, Gießen u.a., besonders 16. - 18. Materialien (Akten, Briefe u.a. in Abschr.) zu Ludw. Landgraf Christian, Luise von Weimar, Melchior Grimm, Landgräfin Caroline, Ludwig IX. u. F.K. von Moser. - Aktenabschr. betr. Legationssekretär Bast Paris. Materialien zu Arbeiten über G. Chr. Lichten Friedrich v. Bodenstadt, Schillers einteiligen Wallstein. - Materialien u. Msc. einer Arbeit über Herd Hemsterhuis (ungedruckt) u. Herder-Merck-Shaftesbur (ungedruckt). Msc. einer Arbeit über Mercks französische Übersetzung von Sömmers Schrift "Über die körperliche Verschiedenheit des Negers vom Europäer", 1785. - Msc. Vortrag über verbotene Literatur in der Bibliothek der Landgräfin, erweitert abgedr. im Hess. Jahrb. f. d. Landesgeschichte, Bd 16 (1966).

24,2: Statistik der Füllwörter in den FGA 1772 mit Vorarbeiten ohne die notwendige (kostspielige) elektronische Aufbereitung. Dazu ein Karteikasten.

Goethes voritalienische Bibliothek (Vorarbeiten u. Vorschläge zu einer notwendigen Ergänzungspublikation) Katalog von Herders Bibliothek. 1804. Xerokopie. Vorschläge u. Vorarbeiten zu einer Herausgabe in der Art wie der Katalog der Bibliothek der Großen Landgräfin (AGB, Ffm 1965).

24,3: Katalog der Herzogin Caroline von Pfalz-Zweibrücken Xerokopie. - Katalog der Bibliothek Ludwigs IX., Xerokopie nebst Materialien.

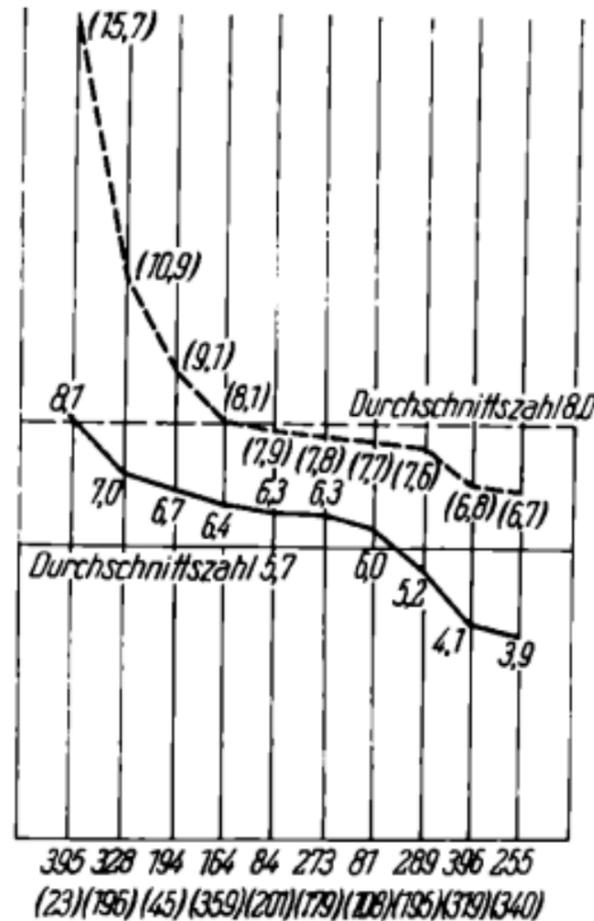
## Statistical and Stylistic Approaches

- **Joachim Thiele** (*Verfahren der statistischen Ästhetik*) **1966**: Untersuchung der Goethe zugeschriebenen Rezensionen in den Frankfurter Gelehrten Anzeigen mit Hilfe **einfacher Textcharakteristiken**, in: *Studia Linguistica* 20 (1966), 83–85.
- **Herbert Sparmann 1970: Häufigkeitsuntersuchungen**, ein Hilfsmittel für den Vergleich von Texten und für die Feststellung der Verfasserschaft. *STUF - Language Typology and Universals* 1970, 227-231.
- **Karin Haenelt 1984**: Die Verfasser der Frankfurter Gelehrten Anzeigen von 1772. Ermittlung von Kriterien zu ihrer Unterscheidung durch **maschinelle Stilanalyse**, in: *Euphorion* 78.4 (1984), 368–382.

# Statistical and Stylistic Approaches

Graphische Darstellung auf Grund der beiden Tabellen

(auf der Abszisse die Rezensionen, auf der Ordinate die Anzahl des bestimmten Artikels, auf je 10 Zeilen bezogen, die Zahlen für MERCK in Klammer; ausgezogene Linie GOETHE, gestrichelte Linie MERCK)



- **Herbert Sparmann 1970:** Tried to distinguish Goethe from Merck by the frequency of the use of the definite article, finding Merck uses the **definite article** 40% more frequent than Goethe.
- **Very small corpus**, taken from FgA!

# Statistical and Stylistic Approaches

- **Karin Haenelt 1984:** Die Verfasser der Frankfurter Gelehrten Anzeigen von 1772. Ermittlung von Kriterien zu ihrer Unterscheidung durch **maschinelle Stilanalyse**, in: *Euphorion* 78.4 (1984), 368–382.
- profile categorising **frequency of word function**: nouns, adjectives, and lexicon variation; analysis of words in 1st, 2nd, last position in the sentence.
- First **computational approach!** Using LDVLIB by R. Drewek, an early textstatistical processor about which you hardly find anything but mentions in books online.
- Very **small corpus**, taken from the FgA!

# FgA Challenges for Computational and Stylometric Approaches

- **Corpus acquisition:** OCR – German Fraktur of the 18<sup>th</sup> century, specific training of engines needed. Consequently, our corpus was partly ‘dirty’.
- **Length** of the Rezensionen varies between 1 page to 7 pages, the shortness of samples may be a problem for authorship attribution.
- Goethe has - in his role of an editor-in-chief - certainly **redacted** some or many Rezensionen by others.
- Corpus: the co-editors have **not written as much as Goethe**
- Corpus: Goethe’s **style might have changed** over the years (from a literary perspective for sure, from a statistical perspective, we don’t know)

# FgA Test Case(s)

- We took four examples from the FgA for a test drive (**blind test**):

	Title	Length in pp	Goethe self-attributed?	Haenelt	Kestemont, Martens, Ries.
1	Cymbelline, ein Trauerspiel, nach einem von Schäckespear erfundnen Stoffe.	2	Yes		
2	Empfindsame Reisen durch Deutschland von S. 2ter Theil. Bey Zimmermann ...	4	Yes		
3	Essays on song-writing: with a collection of such English Songs, as are most eminent for poetical merit. [...]	5	No		
4	Die schönen Künste in ihrem Ursprung, ihrer wahren Natur und besten Anwendung, betrachtet von J. G. Sulzer. 1772	7	Yes		

# Authorship attribution

Anonymous document X



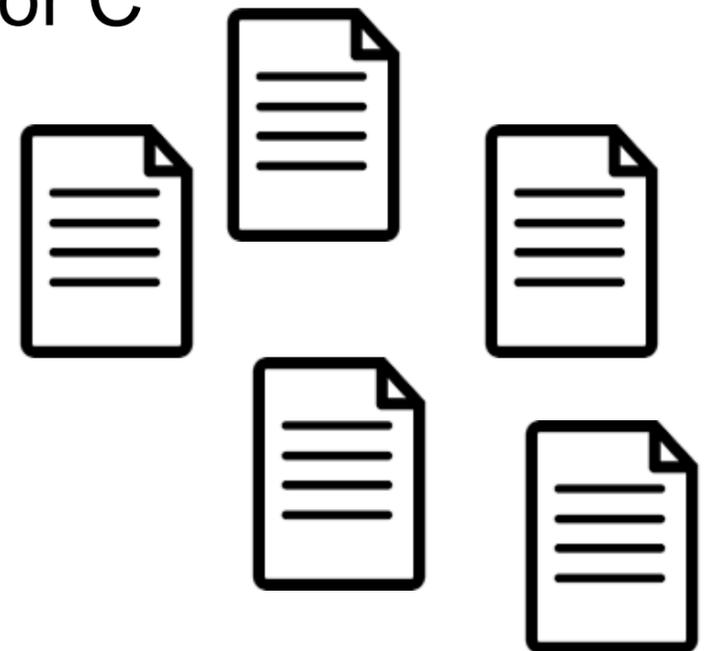
Author A



Author B



Author C



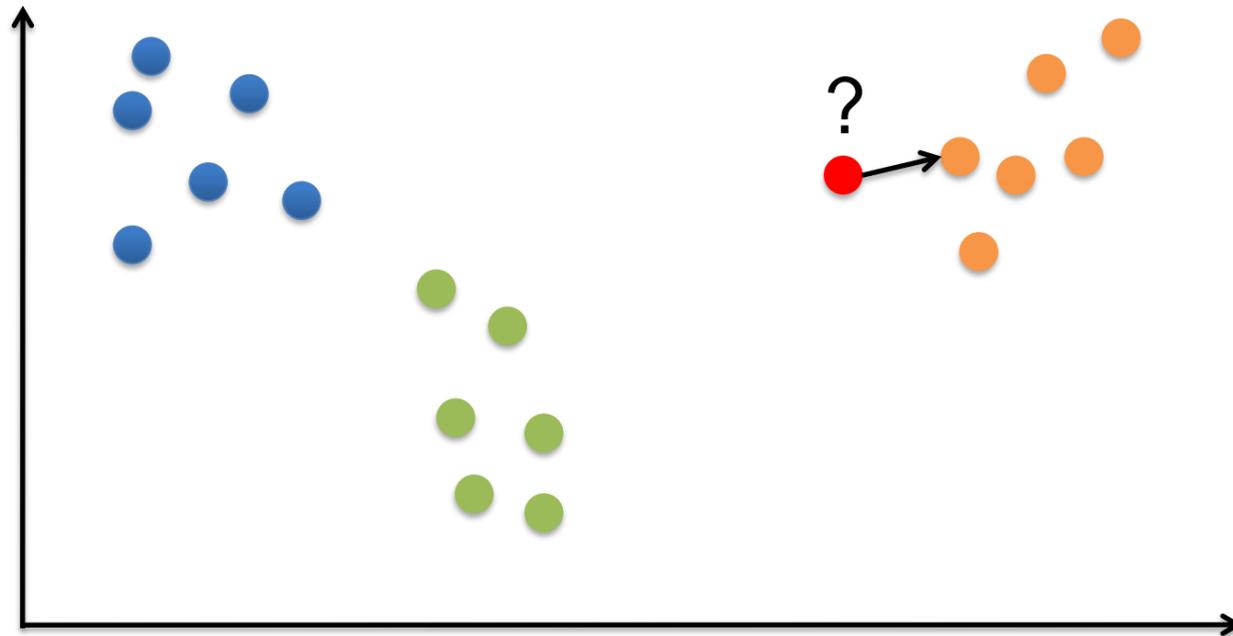
[Stamatatos 2009]

# Represent documents in bag-of-words table

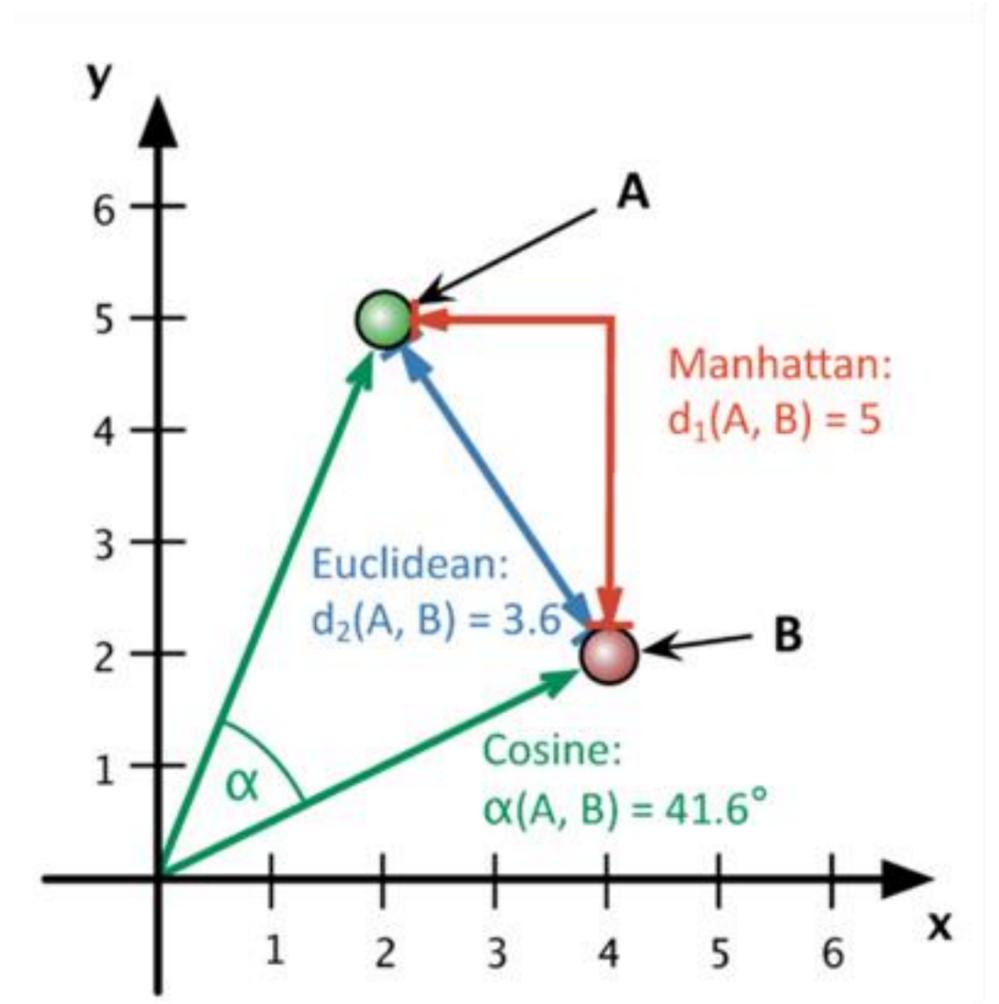
[Burrows 2002; Argamon 2008; Evert et al. 2017; Sebastiani 2002]

	Item 1	Item 2	Item 3	Item 4
Document 1	10	0	12	3
Document 2	2	11	3	0
Document 3	7	8	8	9
Document 4	12	0	1	3

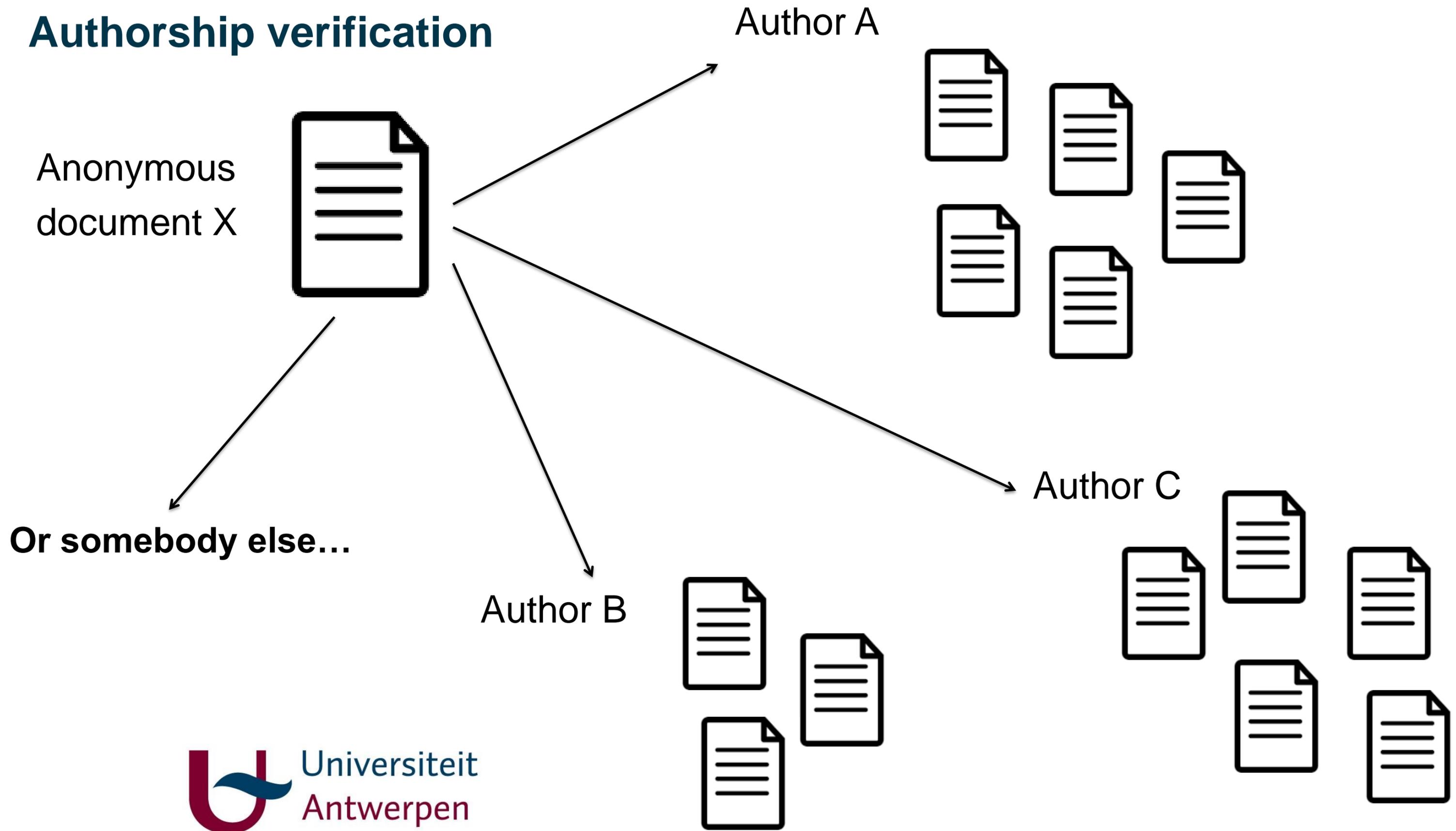
→ “Vocabulary” of features



Find ‘nearest neighbor’ using a distance metric



# Authorship verification



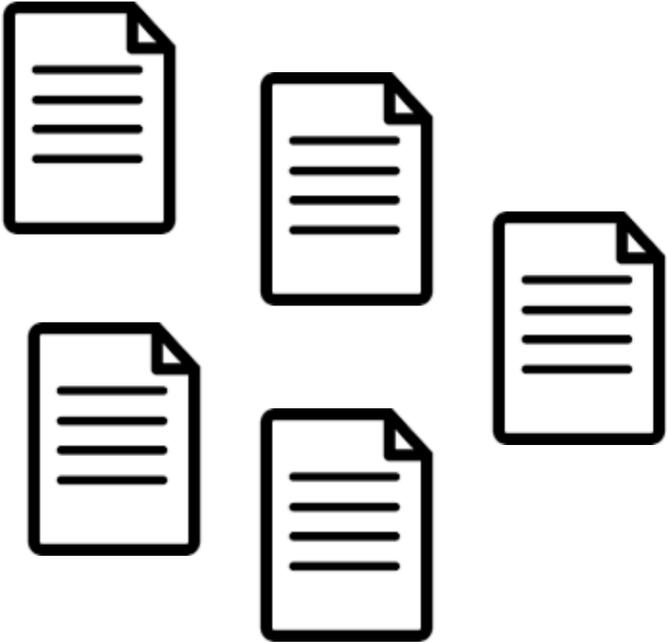
# Authorship verification

Anonymous document X



yes / no

Author A



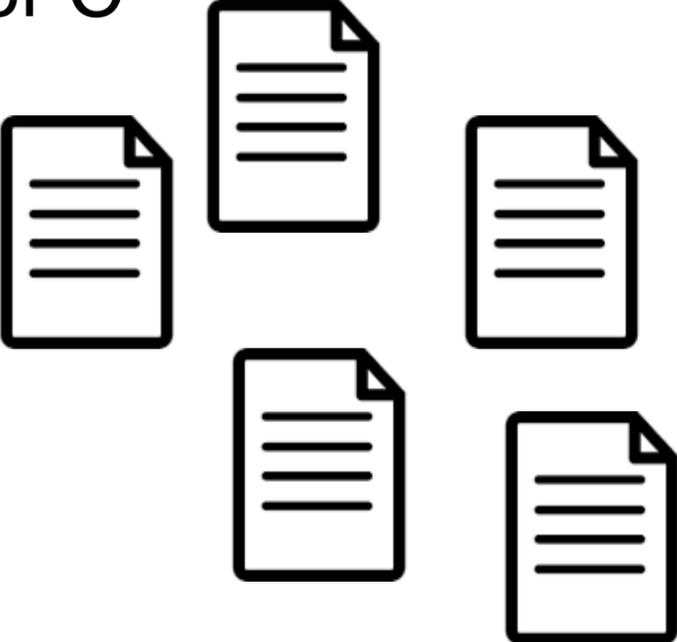
yes / no

yes / no

Author B



Author C





**Repeat e.g. 100 times**

*Random selection  
e.g. 50%*

Vocabulary

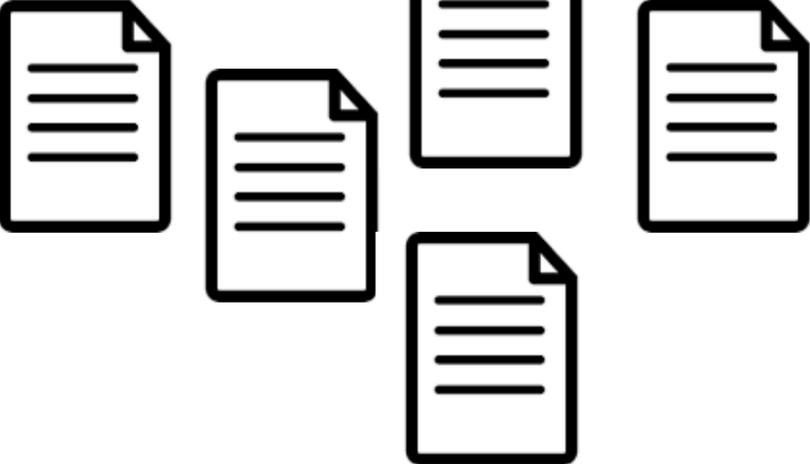


Anonymous document X



Author A

*Closest?*



Imposters "pool"

*Random selection  
e.g. 100 imposters*



Vocabulary

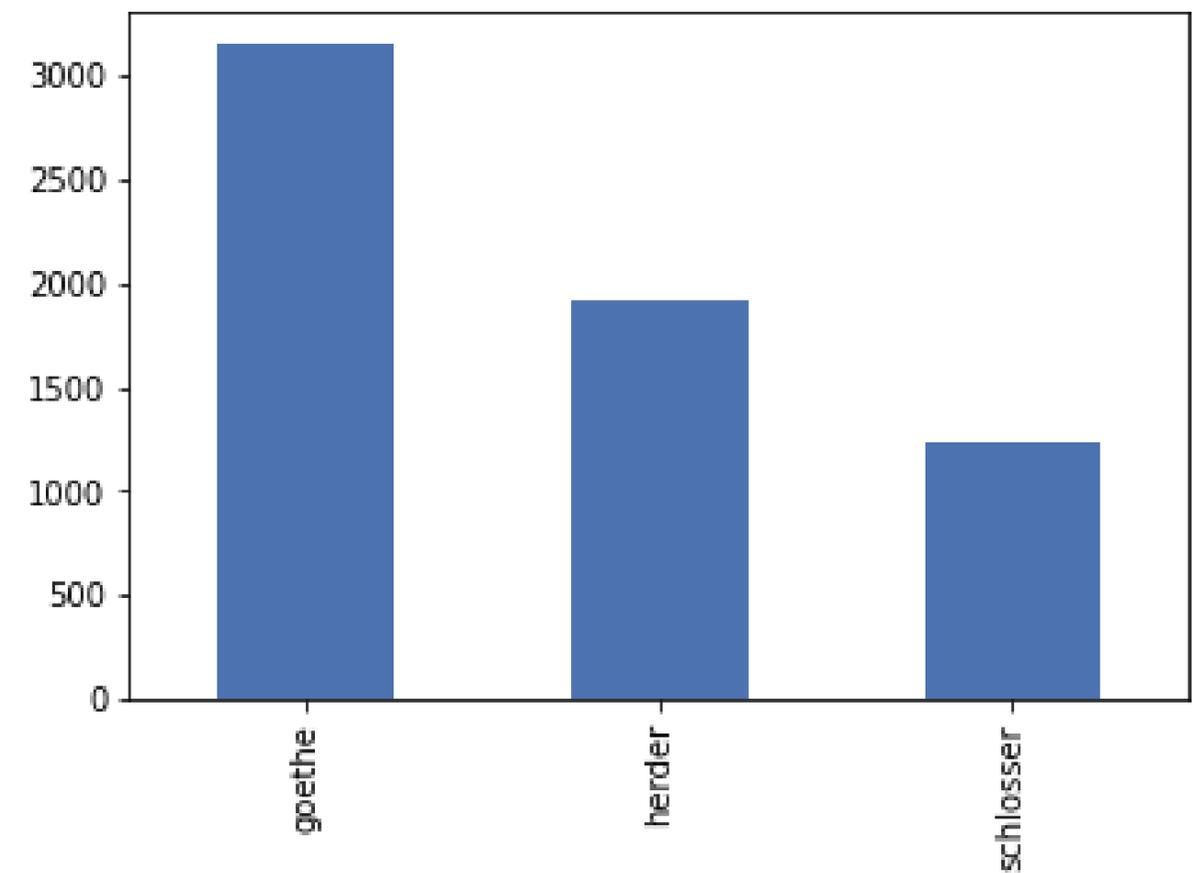
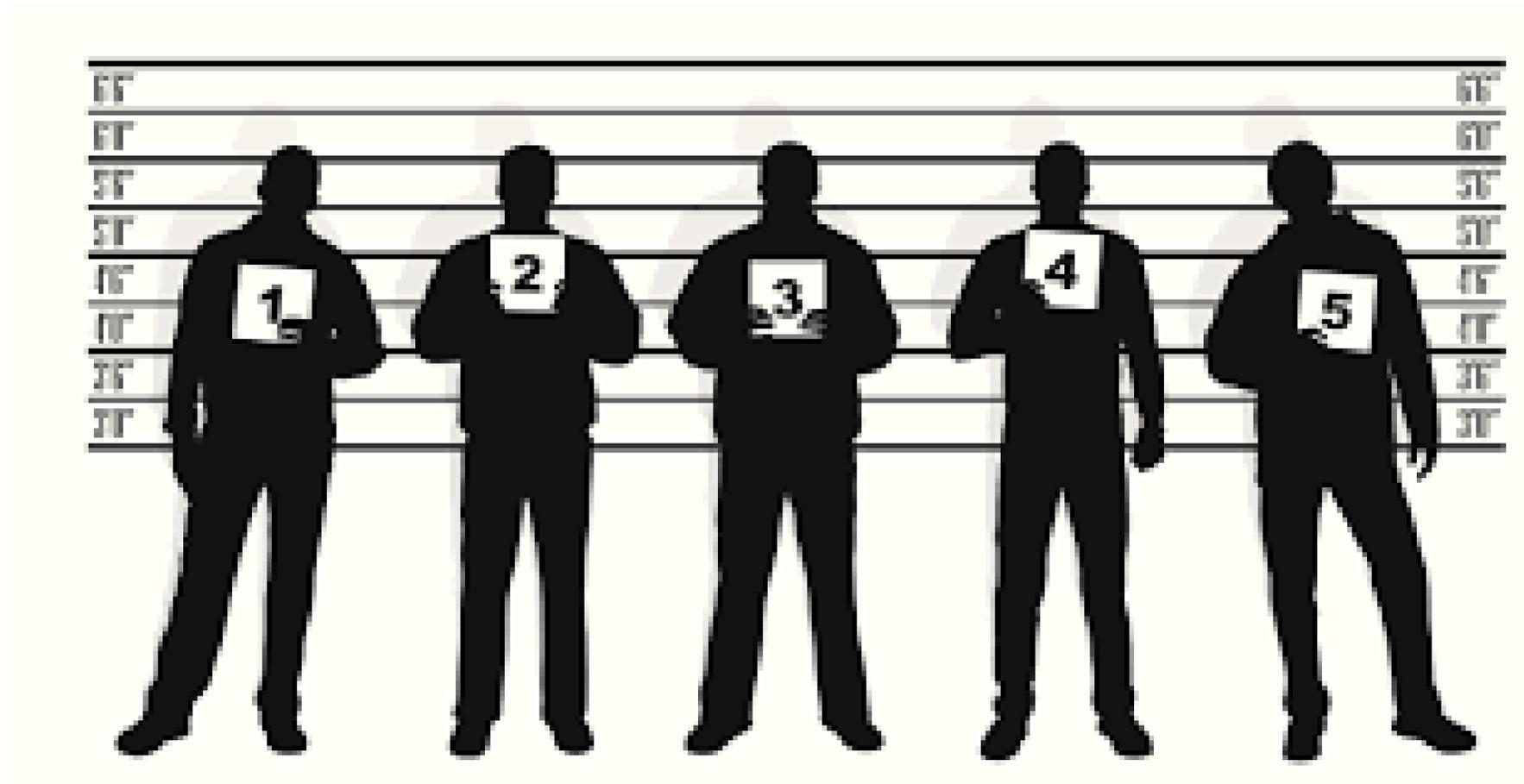
	Item 1	Item 2	Item 3	Item 4
Document 1	10	0	12	3
Document 2	2	11	3	0
Document 3	7	8	8	9
Document 4	12	0	1	3

Documents

- Bootstrapped or stochastic process:  $n$  samples in two dimensions
- Single verification score: e.g. 15/100 vs 87/100
- Apply threshold: e.g.  $\geq 25 \rightarrow$  attribute
- Intuition: documents by same author are similar across random samples from the vocabulary and more similar than other random selections of texts
- Good results in competitions (e.g. PAN)

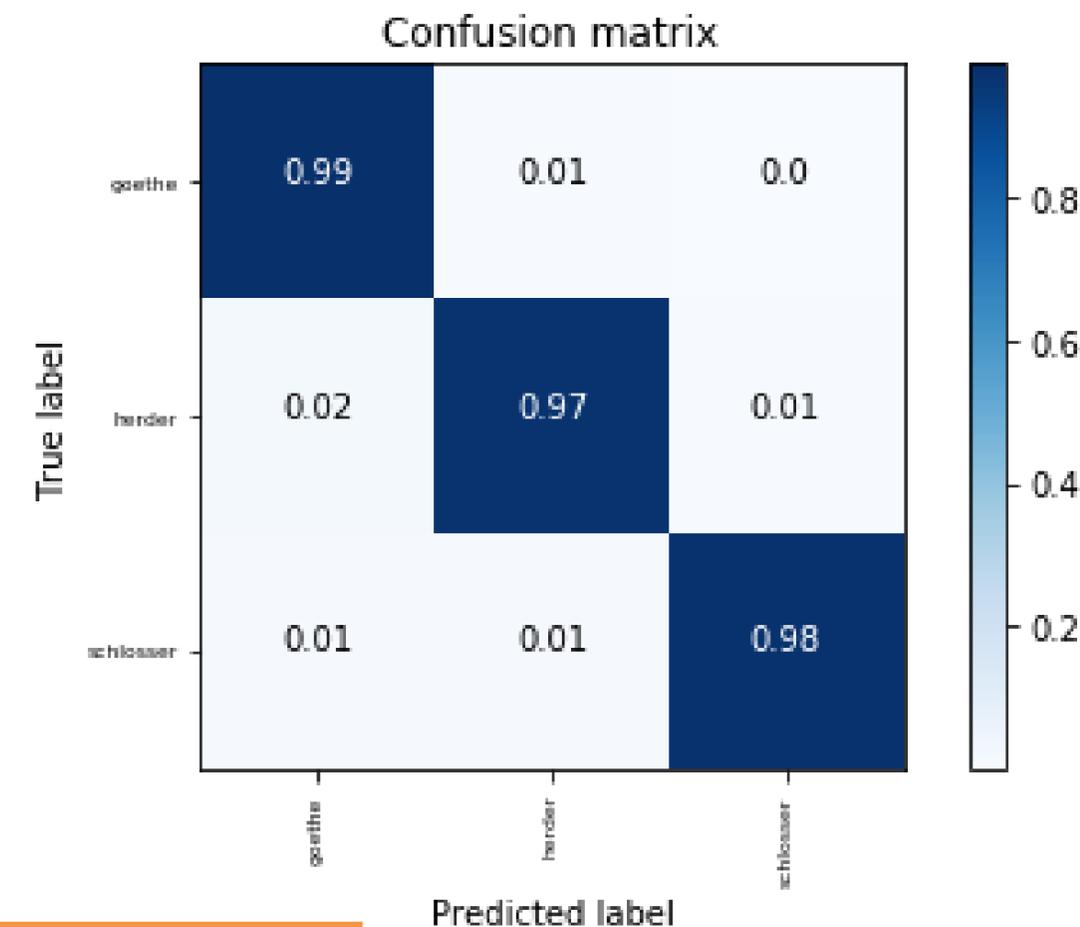
# Imposter selection

- Main difficulty: come with good pool of imposters (cf. police line-up)
- As similar as possible to test and train texts, in terms of genre, date, etc.
- But not too similar either...
- *Rezensionen* from the *Frankfurter Gelehrten Anzeigen* would be ideal, but problematic because all anonymous...
- Restricted to Goethe (target author) vs. Herder & Schlosser (imposters)



# Calibration: development results

- Segment to shortest test sample size (=2,102 words)
- Set apart development set (20% of documents)
- Verify authorship of development set
- Evaluate accuracy (and F1-score) of verifications
- Sampling clearly helps, across both dimensions



	baseline	+ features (50%)	+ features (50%) + imposters (250)
Accuracy	92.48	95.01	<b>98.02</b>
F1-score (macro)	92.34	94.77	<b>97.83</b>

# Test

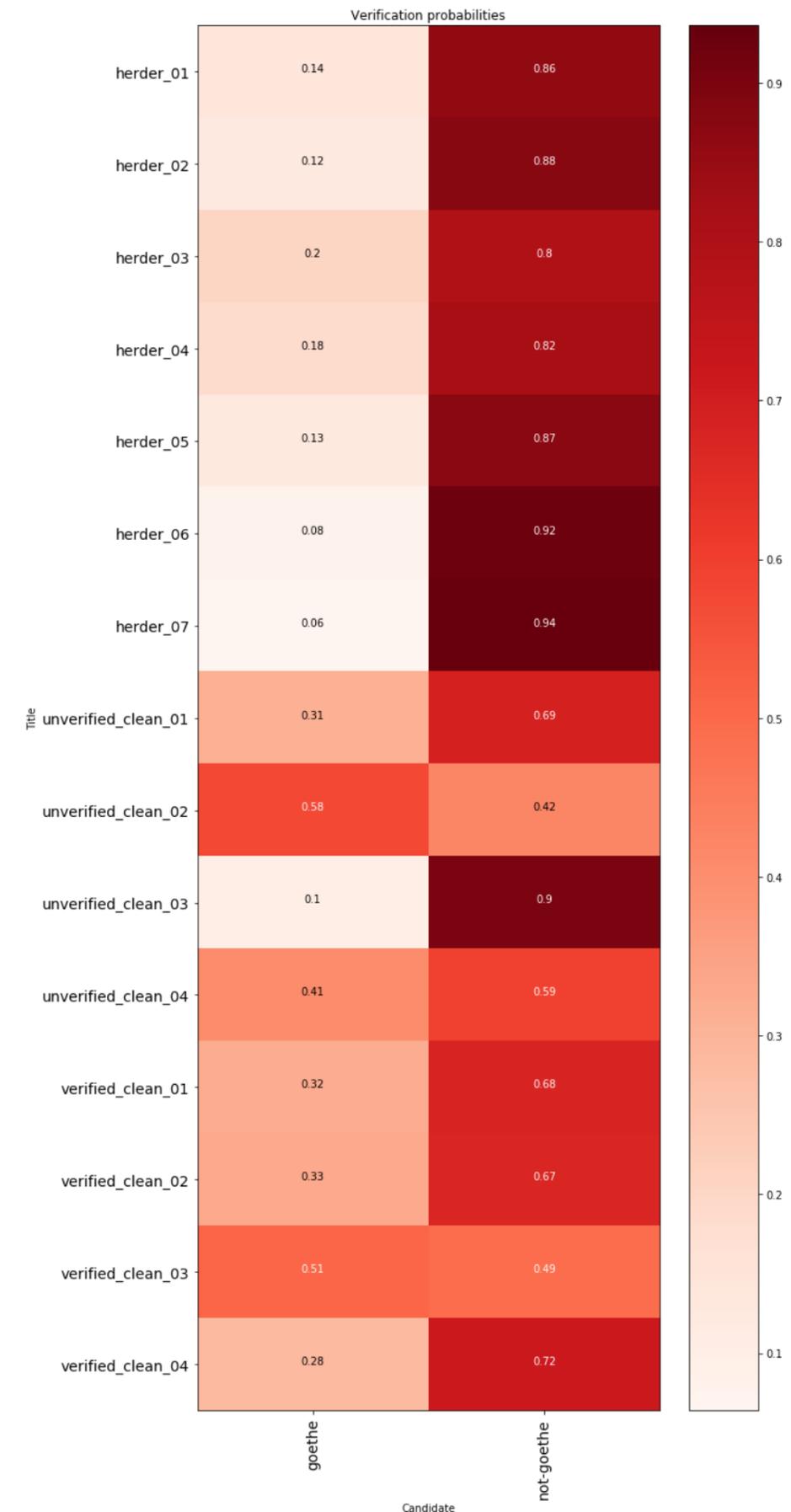
- Apply calibrated system to:
  - verified Herder (7)
  - verified Goethe (4)
  - unverified texts (4)
- Optimal thresholds:
  - $\geq 25$ : Goethe
  - $\leq 16$ : Herder
- Test scores: .51 (solid “yes”) to 0.1 (solid “no”); two in between (.41, .31)
- Send email to Thorsten... (Unbiased)



Herder

Unverified

Goethe



# FgA Test Case(s)

- Summary of test results:

	Title	Haenelt	Kestemont, Martens, Ries	
1	Cymbelline [...]	Highest probability: <b>Herder</b> Features: 2xMerck,2xHerder,1xGoethe	<b>borderline case</b> , wouldn't bet on it, but it does reach the attribution threshold in this experiment.	<b>Same result:</b> it is unclear. It is reasonable to assume that there are features of all three authors in here: collaboration or redaction by multiple authors. The result seems to corroborate some Goethe impact here.
2	Empfindsame Reisen durch Deutschland [...]	Highest probability: <b>Goethe</b> Features: 4xGoethe,1xHerder	really big chance that it's <b>Goethe</b>	<b>Same result.</b> Even clearer here than Haenelt's, where the non-Goethe feature was the vocab composition and distribution, which is one of the most style-determining feature in her matrix.
3	Essays on song-writing: [...]	<b>Herder</b> (Haenelt is sure) Features: 5x Herder	very <b>unlikely that it's Goethe</b>	<b>Same result.</b>
4	Die schönen Künste in ihrem Ursprung, [...]	<b>Goethe</b> (Haenelt is sure) Features: 5x Goethe	<b>reasonable chance</b> that it's <b>Goethe</b>	<b>Almost same result.</b>

THANK YOU FOR YOUR ATTENTION!

len allen diesen Liebhabern des Wahrhaftigen, dieses Buch, sie werden ihre Kennerschaft unglaublich erweitern.

\* \* \*

### Nachrede

statt der versprochenen Vorrede.

Die besondre Aufmerksamkeit, mit der ein geehrtes Publicum bisher diese Blätter begünstigt, läßt uns für die Zukunft eine schmeichelhafte Hoffnung fassen; besonders, da wir uns mit allen Kräften bemühen werden, sie seiner Gewogenheit immer würdiger zu machen.

Man hat bisher verschiedentlich Unzufriedenheit mit unsern Blättern bezeugt; Autoren sowohl als Kritiker, ja sogar das Publikum selbst, haben gewünscht, daß manches anders seyn möchte und könnte, dessen wir uns freylich gerne schuldig geben wollen, wenn uns nicht Unvollkommenheit aller menschlichen Dinge genugsam entschuldigt.

Es ist wahr, es konnten einige Autoren sich über uns beklagen. Die billigste Kritik ist schon Ungerechtigkeit; jeder machts nach Vermögen und Kräften, und findet sein Publikum, wie er einen Buchhändler gefunden hat. Wir hoffen, diese Herren werden damit sich trösten, und die Unbilligkeit verschmerzen, über die sie sich beschweren. Unsre Mitbrüder an der kritischen Innung, hatten ausser dem Handwerksneid, noch einige andere Ursachen, uns öffentlich anzuschreyen und heimlich zu necken. Wir trieben das Handwerk ein Bißchen freyer als sie, und mit mehr Eifer. Die Gleichheit ist in allen Ständen der Grund der Ordnung und des Guten, und der Becker verdient Strafe, der Brezeln backt, wenn er nur Brodt aufstellen sollte, sie mögen übrigens wohl schmecken wem sie wollen.

## Selected references

- Argamon, S., Interpreting Burrows's Delta: Geometric and Probabilistic Foundations, *LLC* (2008).
- Burrows, 'Delta': a Measure of Stylistic Difference and a Guide to Likely Authorship, *LLC* (2002).
- Evert et al., Understanding and explaining Delta measures for authorship attribution, *DSH* [2017].
- Kestemont et al., Authenticating the writings of Julius Caesar. *ESWA* (2016).
- Koppel & Winter, Determining if Two Documents are by the Same Author, *JASIST* (2014).
- Sebastiani, Machine Learning in Automated Text Categorization, *ACM Surveys* (2002).
- Stamatatos, A Survey of Modern Authorship Attribution Methods, *JASIST* (2009).