

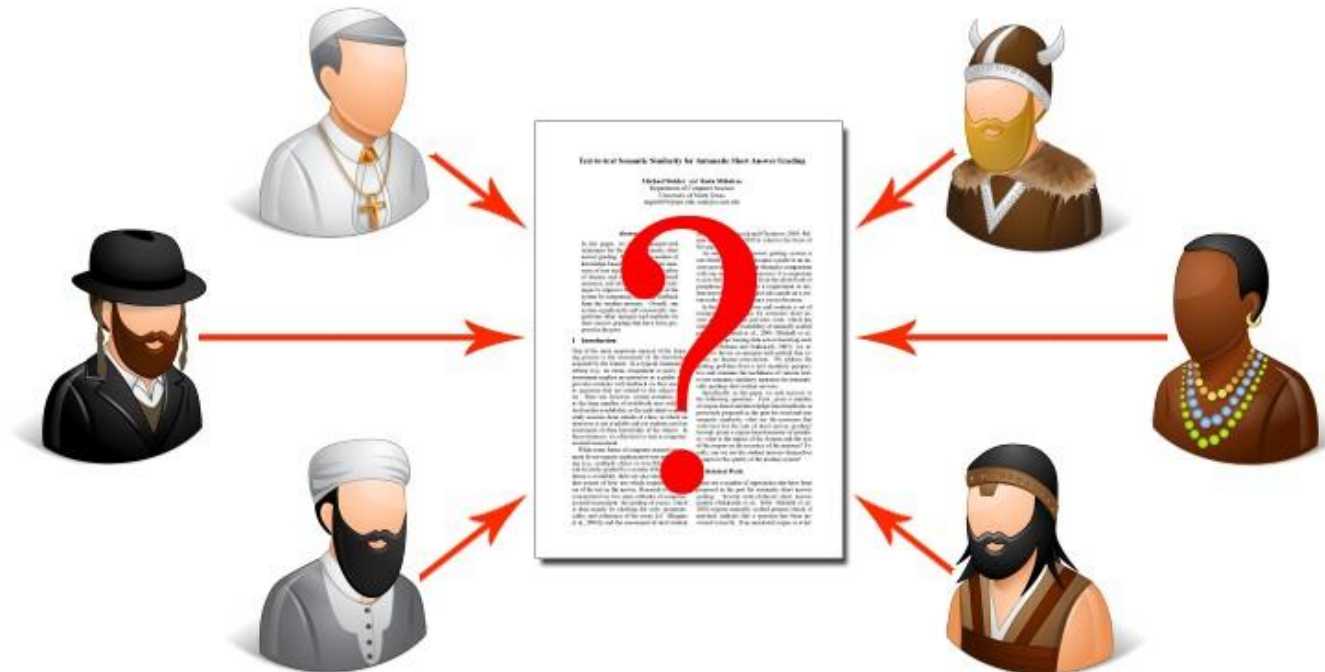
# Authorship Attribution



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

## English for Computer Science II

Oren Halvani, Wu Ding



# Overview

- Motivation
- How does it work?
- Take-Home-Messages
- Discussion
- References



# Motivation

- In this day and age there is an incredible amount of information worldwide

Quote [1]:

*"The Amount of Digital Information Reached 281 Exabytes  
(281 Billion Gigabytes)"*.

- Rough estimate:  $\sim 85\%$  of these are available in a textual representation



# Motivation

- As far as we know there are  $\sim 7.000.000.000$  humans on earth
- From this it follows that there must be many (different) authors, who have produced these textual information !
- Sometimes it is not clear **which** author wrote a specific text
- In order to determine an author of an unknown text, so-called "authorship attribution methods" might be helpful
- These methods offer a wide range of applications...



# Motivation

## Applications:

- Plagiarism detection (remember the Guttenberg Affair ?)
- Forensic evidence in court (verifying confessions)
- Unmasking pseudonymous authors (e.g. in terror extremist blogs)
- Finding additional material of the same author
- Categorization of texts by authors (e.g. in unstructured text collections)
- And many more...



# Motivation

- Authorship Attribution is an extensively researched topic
- Researchers claim: task is not far from being solved

(...for some scenarios)

Quote [2]:

*"Trying to classify an unseen text as being written by one of two or of a few authors is a relatively simple task, which in most cases can be solved with high reliability and accuracies over 95%".*



# Motivation

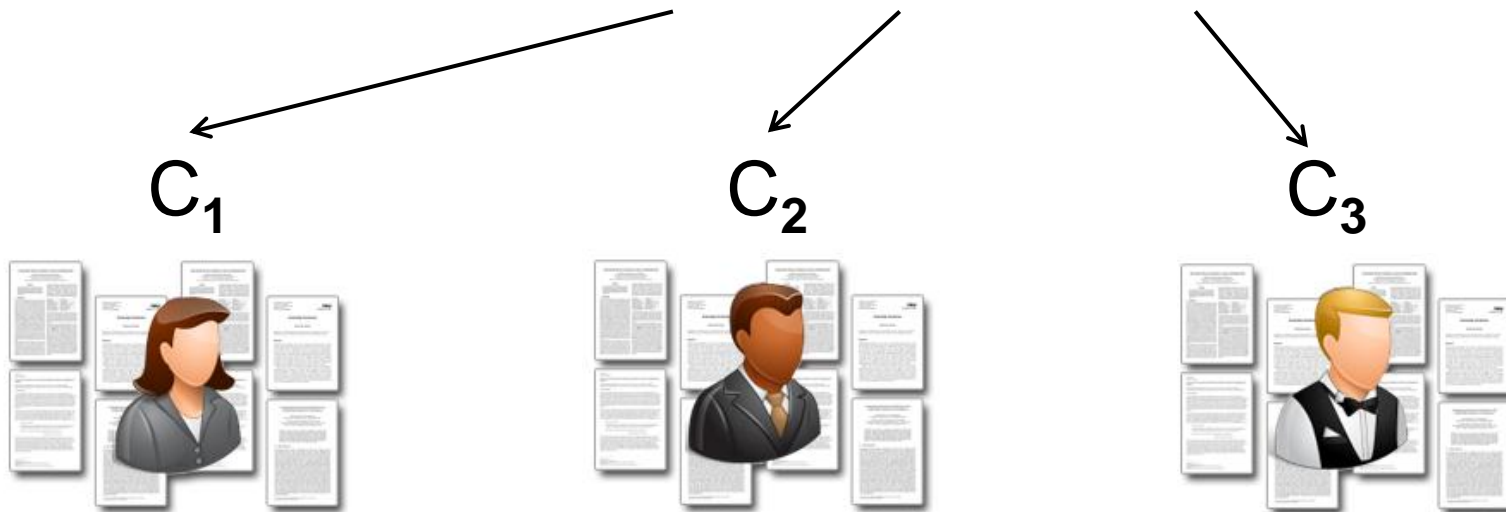
- But, how is it actually possible to recognize **who** wrote a given text?
- Dozens of techniques have been proposed to answer this question...
- Due to a lack of time, we will focus only on one 😊



# How does it work?

## Profile-based approach ( proposed in: [5] )

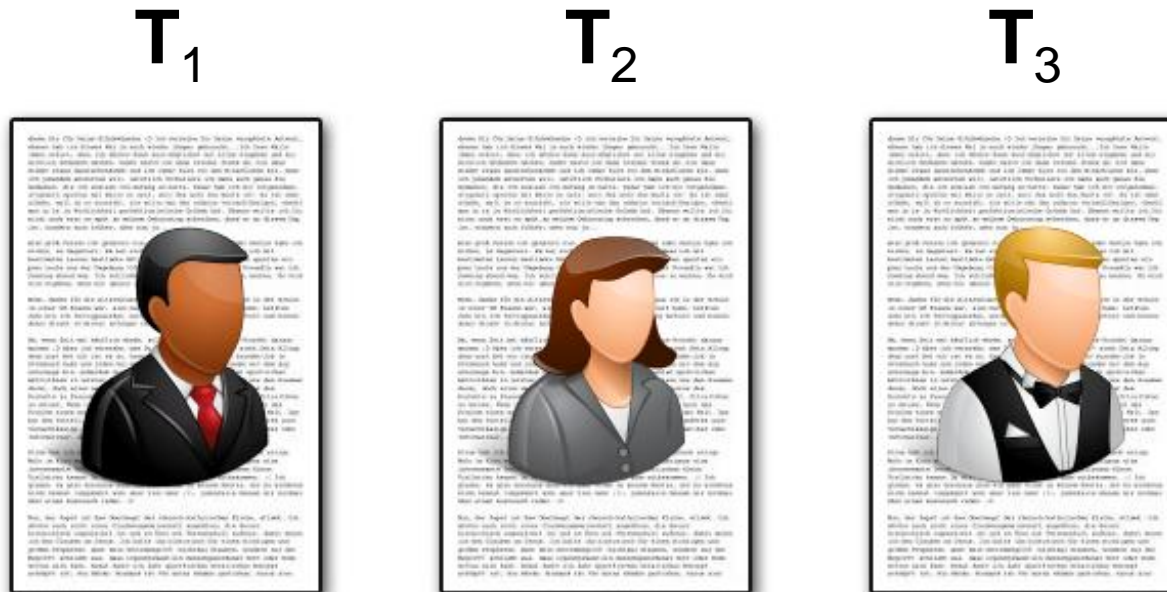
- Assume we have a document collection:  $\mathbf{D} = \{ D_1, D_2, D_3, \dots \}$
- That have been produced by several authors:  $\mathbf{A} = \{ A_1, A_2, A_3, \dots \}$
- Merge both into clusters:  $\mathbf{C} = \{ (A_1, D_3), (A_2, D_4, D_5), (A_3, D_7, D_8, D_{12}), \dots \}$



# How does it work?

## Profile-based approach ( proposed in: [5] )

- Build for each document cluster  $C_i$  one big textfile  $T_i$



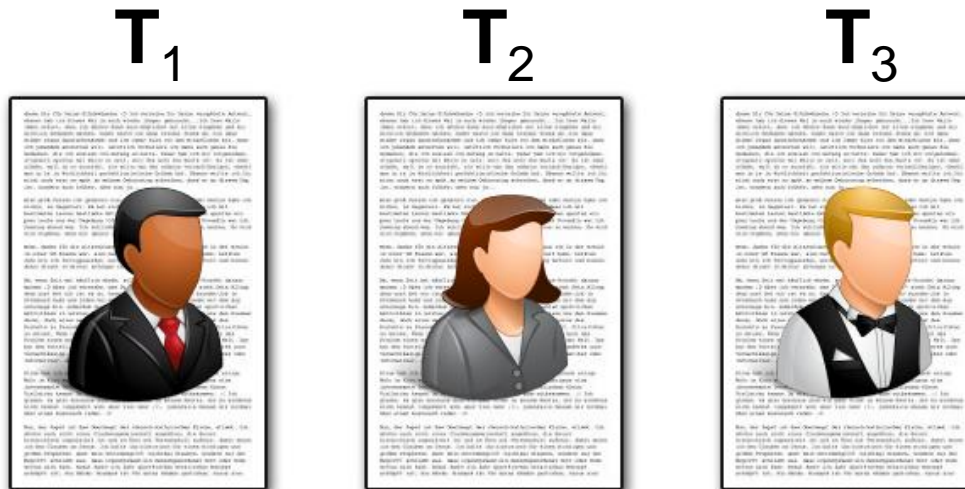
# How does it work?

## Profile-based approach:

- What is the reason we should do that?
- Imagine we collect from an author  $A_1$  documents like:  
private e-Mails, scientific papers, blogs, reports, ...
- Merging these texts will lead to an abstraction of style variation
- Specific style patterns of  $A_1$  remain in  $T_i$  (the big textfile)
- We need a model to find these patterns...



# How does it work?




Attribution  
Modell

Most likely  
author



# How does it work?

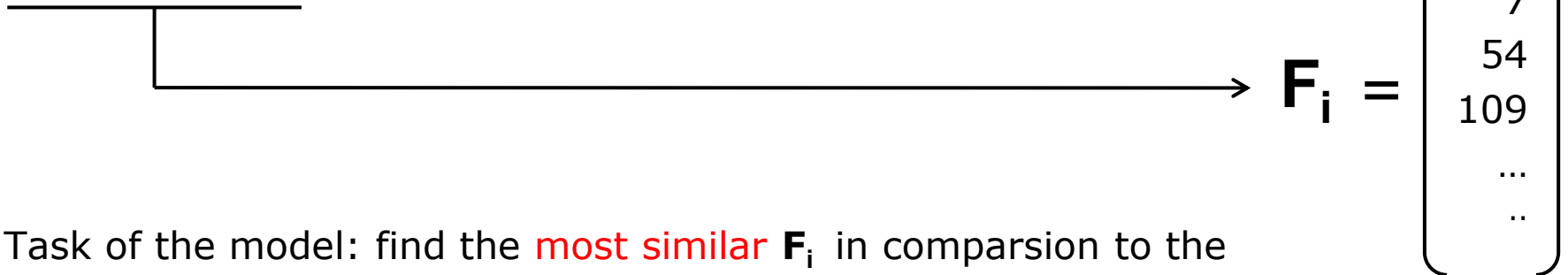
- Sounds simple? Let's have a look behind the scenes...
- If we want to discriminate authors (represented through  $\mathbf{T}_i$ ) we first must understand how to distinguish their style
- Bad news → there is no definition for style 😞
- However, style can be approximated through a combination of various **features**

- 
- vocabulary richness
  - average word/sentence length
  - number of specific symbols (-.:,#?!'&)
  - number of: adjectives, nouns, verbs, ...
  - ...

# How does it work?

- What we have so far:  $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3, \dots$

- Applying feature extraction on the  $\mathbf{T}_i$  will result in the so-called feature vectors:

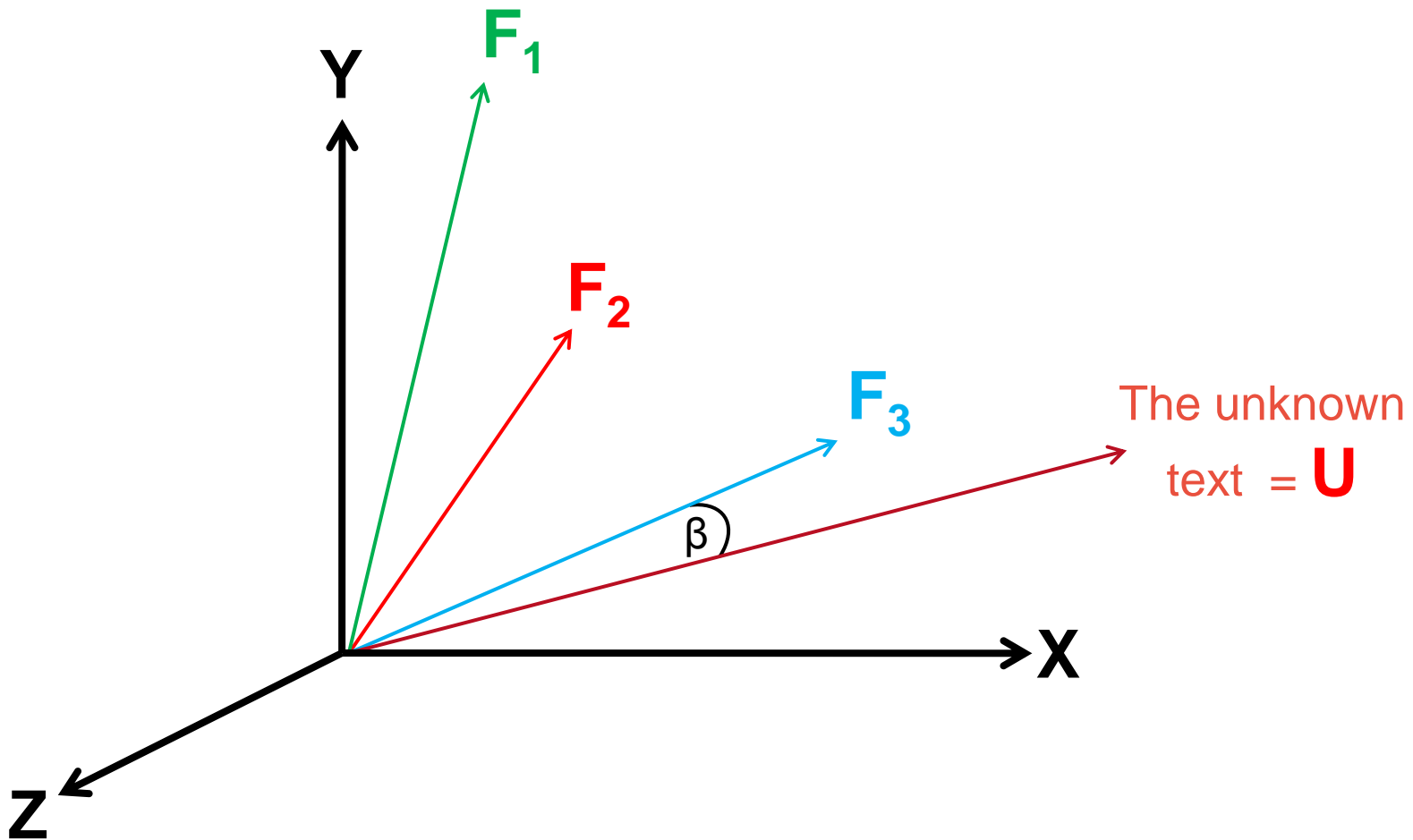


- Task of the model: find the **most similar**  $\mathbf{F}_i$  in comparison to the features of the unknown text document

- **most similar** = shortest distance (e.g. in a vector space...)



# Recap: Linear Algebra



# Recap: Linear Algebra

- The angle  $\beta$  represents similarity between 2 vectors
- Can be computed as follows:

$$\cos(\beta) = \frac{F_i * U}{\| F_i \| * \| U \|}$$

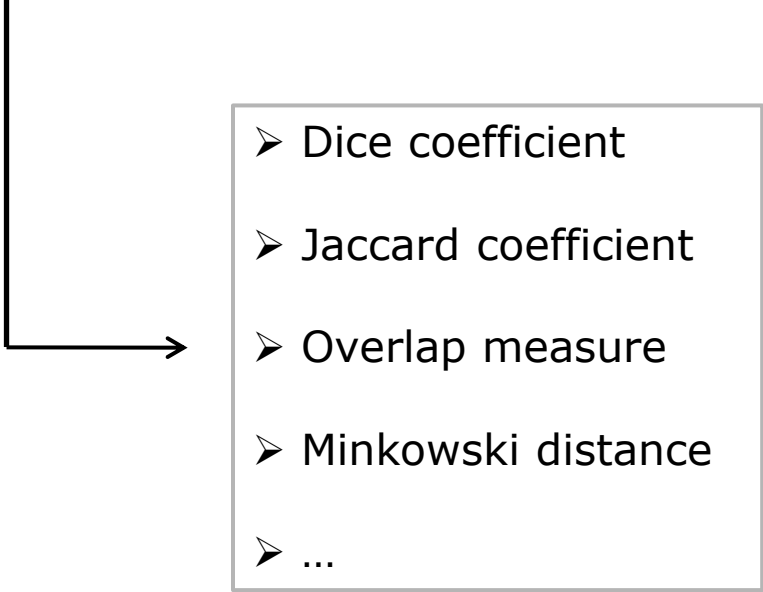
- Standardized resulting number is between: [ 0 ; 1 ]

completely different

absolutely similar

# Recap: Linear Algebra

- Besides the cosine similarity many other “**metrics**” are typically used

- 
- Dice coefficient
  - Jaccard coefficient
  - Overlap measure
  - Minkowski distance
  - ...

- All of these metrics share the same idea:

→ Figure out if two vectors correlate with each other !

# Take-Home-Messages

## Authorship Attribution:

...is no longer a utopian vision !

...is useful in many scenarios beyond Computer Science !

...is something that could affect you too (hopefully not)



# Thanks for your attention !

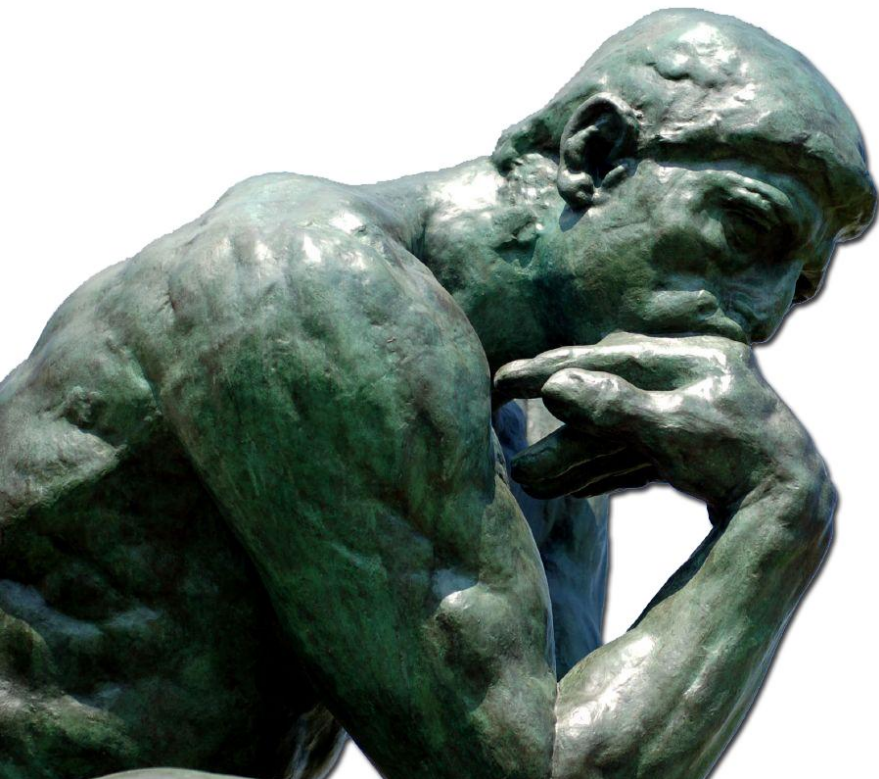


TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



[3]





*Questions...?*

[4]



# Your turn ;-)

- A couple of features have been mentioned during the presentation, can you think of additional features?
- Do you think that symbol-related features (number of hyphens, commas, etc.) are useful to discriminate the style of authors?
- Imagine we have 1000 features to train the model, how should we handle less useful features without discarding them?
- Name at least one application where you could use Authorship Attribution for your own purpose



# References

- [1] **“The Amount of Digital Information”**,  
<http://www.infoniac.com/hi-tech/amount-digital-information-reached-281-exabytes.html>  
[Accessed on: 09.06.2011]
- [2] **“Authorship Attribution and Verification with Many Authors and Limited Data”**,  
Kim Luyckx and Walter Daelemans, CNTS Language Technology Group,  
University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium.  
<http://www.aclweb.org/anthology/C/C08/C08-1065.pdf>  
[Accessed on: 09.06.2011]
- [3] **“CARTOON NETWORK and the logo are trademarks of and © 2011  
Cartoon Network. A Time Warner Company. All Rights Reserved.”**



# References

[4] **“Questions picture: photo of "Le Penseur”**,

A bronze sculpture made by [Auguste Rodin](#), held in the [Musée Rodin](#) in Paris, France.

[5] **“A Survey of Modern Authorship Attribution Methods”**,

Efstathios Stamatatos, Dept. of Information and Communication Systems Eng.  
University of the Aegean, Karlovassi, Samos – 83200, Greece .

<http://www.icsd.aegean.gr/lecturers/stamatatos/papers/survey.pdf>

[Accessed on: 09.06.2011]