



**HAL**  
open science

## Rethinking intertextuality through a word-space and social network approach – the case of Cassiodorus

Johannes Bjerva, Raf Praet

► **To cite this version:**

Johannes Bjerva, Raf Praet. Rethinking intertextuality through a word-space and social network approach – the case of Cassiodorus. 2016. hal-01279833

**HAL Id: hal-01279833**

**<https://hal.science/hal-01279833>**

Preprint submitted on 1 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Rethinking intertextuality through a word-space and social network approach – the case of Cassiodorus

Johannes Bjerva<sup>1</sup> and Raf Praet<sup>2</sup>

<sup>1</sup>Center for Language and Cognition Groningen

<sup>2</sup>Groningen Research Institute for the Study of Culture

<sup>1,2</sup>University of Groningen, The Netherlands

Corresponding author: Johannes Bjerva, [j.bjerva@rug.nl](mailto:j.bjerva@rug.nl)

### Abstract

Continuous space representations of words are currently at the core of many state-of-the-art approaches in computational linguistics. The distributional hypothesis, summarised as: ‘You shall know a word by the company it keeps’ [Firth, 1957] is the basis of many such methods. In this paper we use this type of representation, which has seen little to no use in digital humanities, to rethink the concept of intertextuality. We present and use an alternative conceptual concept of intertextuality to ascertain how different persons are portrayed in a late antique letter collection, the *Variae* of Cassiodorus (ca. 485–585 AD). We combine this approach with the well-explored method of network analysis.

‘The study of intertextuality is the study of a certain kind of relation between texts: One text quotes another or others.’ [Edmunds, 2001]. Until recently, intertextuality has been pictured as an interaction between different texts, which has been restricted to the surface forms. We want to transcend this rather limited, one-dimensional concept of intertextuality by using high-dimensional word representations which effectively abstract away from such surface forms. Instead of conceptualising, e.g. Vergil, as the sum of his transmitted oeuvre, we represent him both as a node in a network, and a vector in high-dimensional space. In this way we overcome the border between text and historical person; a border which impedes the ascertaining of the intertextual impact of authors which are partially or not at all preserved. We create word-space representations based on the letters in the *Variae*, using methods based on distributional semantics [Mikolov et al., 2013a, Levy et al., 2015].

In antiquity, the editing and publication of letter collections was a fundamental tool for literary and cultural self-representation. Late antiquity witnessed the zenith of this practice with the publication of several such collections, both in Latin and in Greek. The *Variae* of Cassiodorus are an excellent example of this practice of self-representation [Bjornlie, 2012, Gillett, 1998]. In this paper, we will represent Cassiodorus, his contemporaries, and influential authors of the literary canon, such as Vergil, in one and the same network. This form of visualisation can generate a more nuanced view on how Cassiodorus constructs a cultural profile for himself and his peers. Indeed, the letters of Cassiodorus act as a meeting ground in which both the contemporaries of Cassiodorus, as well as the authors who shaped the intellectual outlook of Cassiodorus and his peers, interact with each other.

### Keywords

intertextuality, word embeddings, late antiquity, social networks, cassiodorus

## I INTRODUCTION

In his masterpiece the *Dialogues of the Dead*, the author of the Roman imperial period, Lucian of Samosata (ca. 125 – after 180 AD) [MacLeod, 1961] presents the underworld as a platform which is liberated from the limitations imposed on mortal humans by time and space; mythological characters, historical figures, philosophers and literary authors who were otherwise separated in time and space in their own lifetime, are seen conversing with each other, often resulting in humorous contradictions and confrontations.

Lucian's underworld is arguably a sound metaphor for the mechanism of intertextuality, or, in general, the presence of an author X in the works of author Y. Not only is the presence of author X in author Y's works measured by textual references and similarities; it is also measurable by how this author appears as a character or concept in the works of author Y. Through his metaphor of the underworld, Lucian presents us with a conceptual approach to intertextuality which is liberated from the narrowing level of the text. As such, his 'conceptual' take on intertextuality does not only serve to trace the presence in author Y of otherwise partially or non-preserved authors. It is also useful for ascertaining the presence of an author in works which were meant for a broad and socially diversified audience; whereas in antiquity only a small intellectual and political elite had the resources to become closely acquainted with the works of a specific author X to acknowledge a reference to this author in the works of author Y, everyone who only has a slight grasp of the cultural value of author X will recognise the reference to this author when this author is mentioned or conceptualised. In the underworld of Lucian, one does not have to be acquainted intimately with the philosophy and writings of, for instance, Diogenes the Cynic – writings which are regrettably lost today – to recognise his 'presence' in Lucian's dialogues; Diogenes simply walks by and starts conversing with his fellow deceased.

The *Variae* [Fridh and Halporn, 1973]<sup>1</sup> of Cassiodorus are a prime example of a text which is designed to appeal to a diversified audience. Flavius Magnus Aurelius Cassiodorus Senator (c. 485 – c. 585) served under the Ostrogothic king Theodoric and his successors until the collapse of the kingdom under the Byzantine armies (535 – c. 540). After the toppling of the realm of the Ostrogoths, Cassiodorus stayed in Constantinople (c. 540 – 554) where he most possibly compiled and composed his collection of official letters, the *Variae* [Bjornlie, 2012, 19-26], which he wrote in the previous years on behalf of king Theodoric, his successors, or on his own account as praetorian prefect.<sup>2</sup>

The multi-layered composition history of the collection accounts for its different intended audiences – and the different levels of perception of intertextuality. On the one hand, the official letters of the Ostrogothic kings carried the weight of law and were addressed to the populace of the realm to communicate the will of the king and his government – most possibly they were also read aloud at different occasions. In this case, it was unlikely that the audience grasped any of Cassiodorus' references to, for instance, Cicero, unless he was explicitly mentioned – what Cassiodorus also does not fail to do in several instances.<sup>3</sup> On the other hand, the *Variae* were also intended as a collection of showpieces to advertise the rhetorical and intellectual capacities of Cassiodorus to his competitors [Gillett, 1998, Jouanaud, 1993] – and his possible new employers.<sup>4</sup> Indeed, Cassiodorus seasons his letters with learned allusions and digressions,

---

<sup>1</sup>A new edition with commentary is being prepared by Giardina et al. [2015].

<sup>2</sup>An overview of the life and works of Cassiodorus can be found in [O'Donnell, 1979].

<sup>3</sup>Namely in letters I.3, VI.5, VIII.12, VIII.19 and the preface to books XI-XII.

<sup>4</sup>For an analysis of how literary skill was a prized means to gain prestige and secure the support of a wealthy and powerful patron in 6th-century Constantinople, see Rapp [2005].

depending on the addressees of the letters; he has a marked tendency to reserve his learnedness for Romans in general and such intellectuals as Boethius and Symmachus in particular, whereas audiences of Goths are usually deprived of – or saved from – his references to classical literature.<sup>5</sup> Furthermore, Cassiodorus compiled his *Variae* in Constantinople after losing his job as administrator in the Gothic state. Most possibly he intended his *Variae* as an advertisement for a future career in the Byzantine state – and therefore it is not excludable that he polished up his collected letters by adding additional learned lore. This second intended audience of the *Variae*, elite politicians and intellectuals in Ostrogothic Italy (for the single letters) and Constantinople (for the collection), is well capable of recognising textual references to, for example, Cicero, without having to be reminded of Cicero with a clear mention of his name in the text.

The intricate ways in which Cassiodorus communicated his references to literary authors and historical figures in his *Variae* to his audiences demands an approach to intertextuality, or, in other words, an approach to the presence of authors in the works of Cassiodorus, which transcends the basic plane of the text, and abstracts away from surface forms. A first attempt to ‘emancipate’ the presence of the author in a text from the level of the text was made by Müller [1991]. He introduces the concept of ‘interfigurality’, namely, the intertextual relation between a character/figure in text A with another figure in text B. Although this concept is a useful first step in detaching the presence of an author in a text from the level of the text, we propose to go a step further in this paper, by placing literary authors influencing the *Variae* of Cassiodorus on the same level of Cassiodorus’ contemporaries and acquaintances mentioned by him in the *Variae*. By subjecting literary authors influencing Cassiodorus side by side with contemporaries and historical characters to the same methods of digital analysis, we aspire to contribute to the development of an open approach to intertextuality. As in Lucian’s underworld, both literary authors and contemporaries of Cassiodorus will be seen as engaging with each other in dialogue on the common platform of the *Variae*.

We define conceptual intertextuality as instances in which high associations are found between literary authors, historical persons and concepts within a document, or the entire oeuvre of an author. We approach this conceptual intertextuality using state-of-the-art methods from distributional semantics, as such models allow for effective abstraction from the surface forms of the texts at hand. Word embeddings are one example of such models. They have seen much recent use within computational linguistics, although usage within digital humanities appears to be limited. Recent work by Koopman et al. [2015] employs vector representations to calculate similarities between entities such as authors and journals in an article database. Bjerva and Praet [2015] use word embeddings trained on around 1.4 billion tokens of latin text in order to find associations between historical characters and concepts relevant to the period of late antiquity. Usage of word embeddings in the humanities is further discussed by Tahmasebi et al. [2015], who suggest that they could be useful for comparing word vectors trained on different epochs of time, thus revealing changes in usage of words across time. This idea is (indirectly) picked up by Zhang et al. [2015], who looked at the development of terms across time in relatively modern texts, e.g. finding that the term *iPod* is a modern analogue to the word *walkman* from some decades earlier.

---

<sup>5</sup>[Barnish, 2001, 367]: ‘The most learned letters, moreover, tend to be directed to Romans of known learning, like Boethius, while biblical allusions tend to occur in letters to, or on behalf of, men of known religious interests, like Theodahad. There is, then, some attempt at adaptation to the audience (. . .)’.

## 1.1 Aims of this work

This paper aims to ascertain whether a combination of digital methods, such as a heat map and network analysis, can contribute to a nuanced assessment of the representation of Cassiodorus, his colleagues and contemporary political leaders, through historical and literary references in the *Variae*. We hope to both contribute to the understanding of Cassiodorus' ways of (self-) presentation in the *Variae* and to contribute to the redefinition of the concept of intertextuality in terms of a multifarious 'presence' of an author in a literary work.

Furthermore, we aim to investigate to what extent word embeddings can be used as a tool to aid an analysis which traditionally would be purely qualitative in nature. In doing this analysis, we also wish to learn to what extent word embeddings are a suitable tool for research where deep interpretation of the model itself is needed.

This paper contributes to the related work done by Bjerva and Praet [2015] in the assessment of the construction of ideological images of late antique characters through their associations with key ideological values. Regarding intertextuality, recent work has been done on the automatic detection of intertextuality in, for example, poetry [Coffee et al., 2013]. The application of digital tools to intertextuality has resulted in a redefinition of the concept of intertextuality altogether, to which, as already mentioned, this paper also contributes. For example, Ghiban and Truşan-Matu [2013] already approached intertextuality as a network, and Małajowicz [2013] explored the merits of stylometry for the assessment of intertextual ties.

## II DATA

### 2.1 The *Variae*

#### 2.1.1 History

The *Variae* are a collection of state letters ordered in twelve books [Fridh and Halporn, 1973, Zecchini et al., 2014]. After a general preface, books I-V (with 46, 41, 53, 51 and 44 letters respectively) were written on behalf of king Theodoric (454-526 AD). Books VI and VII (25 and 47 letters) are anonymised model letters of appointment for the various offices in the Ostrogothic government, which are also called *formulae*. Books VIII and IX (33 and 25 letters) were written on behalf of Theodoric's successor Athalaric (516-534 AD), and book X (35 letters) was written on behalf of the last kings, queens and regents of the Ostrogothic dynasty in Italy (Amalasuetha, Theodahad, Gudelina and Vitiges). After a second preface, books XI and XII (40 and 28 letters) were written on Cassiodorus' own behalf as praetorian prefect.

The collection as we have it was subjected to two phases of composition. In a first phase, Cassiodorus wrote the single letters of the first ten books on behalf of his Ostrogothic masters, when he served in several offices of the Gothic state (ca. 507-532 AD). Cassiodorus wrote the letters of the last two books in the capacity of praetorian prefect (533-540 AD).

The second phase consisted of the collection, selection, ordering and rewriting of the different letters to form the *Variae* as a whole. As already mentioned, Cassiodorus most possibly redacted his text between 540 and the mid-540's in the city of Constantinople. He added two prefaces to the volume and closed the collection with a thirteenth book, a treatise with the title *On the Soul* – which we will not treat in this paper.<sup>6</sup>

---

<sup>6</sup>The *Variae* and the treatise *On the Soul* are admirably analyzed as two parts of one composition, which engages in a close dialogue with Augustine's *Confessions* by Halporn et al. [2004, p. 19-22]. The same unity of both works has been stressed by Barnish et al. [1992, xxiv-xxv], who places them within the project of providing the Roman secular officials with an adapted ideology of service.

Apart from the addition of the prefaces and book XIII, there are indications that Cassiodorus made many changes to the letters on a micro-level in his redaction.<sup>7</sup> For example, in books VI and VII Cassiodorus systematically omits names to give the letters a more exemplary character as model letters or *formulae* [Gillett, 1998, p. 46]. There are also indications he added extra bibliographical references.<sup>8</sup>

Table 1: Overview of the letters in the *Variae* with letter counts per book.

Book	Notes	Number of letters
I	on behalf of king Theoderic	46
II	—  —	41
III	—  —	53
IV	—  —	51
V	—  —	44
VI	<i>formulae</i>	25
VII	—  —	47
VIII	on behalf of king Athalaric	33
IX	—  —	25
X	on behalf of the last regents	35
XI	on Cassiodorus' own behalf	40
XII	—  —	28
<b>Total</b>		468

In total, the *Variae* contains 468 letters, and a total of approximately 120,000 tokens. Each letter is ‘tagged’ with the writer and recipient of the letter, which is a prerequisite for the method we use to construct a network of letter interactions. In this paper, the *Variae* are used as a source of historical figures and concepts. Additionally, it is used as raw text to train word embeddings.

### 2.1.2 Pre-processing

In order to facilitate the training of the distributional semantic model used in this study, we employ some standard pre-processing steps. We convert all text to lower-case, remove punctuation, and replace all numerals (arabic and roman) with a placeholder token *\_NUM\_*.

<sup>7</sup>Barnish et al. [1992, xviii] summarises the principles underlying the composition of the *Variae*. Portions of official files are included, approximately undisturbed, at the center of each book. The beginning and end of each book, on the other hand, are determined by more literary criteria. Diplomatic showpieces are set in front or conclude each book. These ordering principles left Cassiodorus with ample opportunity to rework portions of his correspondence for various reasons. Giardina [1993, p. 69-70], for instance, has already showed how one can single out several passages in the collection which were added or reworked during the second phase to fit a specific Cassiodorean purpose. The singling-out of these purpose-specific passages can even amount to a relative chronology of composition. Bjornlie [2009, p. 149]: ‘Signs of heavy revision and adaptation appear throughout the *Variae*, including two extensive prefaces, the deletion of epistolary protocols, and the inclusion of two books of *formulae*.’ Bjornlie [2009, p. 144] uses these signs of revision to argue for a reworking of the state letters out of political motives.

<sup>8</sup>We have one instance in which Cassiodorus provides an insight into his didactic rewriting of the *Variae*. The arrival of an expert on how to find water in letter III.53, gives an admirable pretext for Cassiodorus to elaborate on the different signs which point to the existence of water. At the beginning of paragraph 4, Cassiodorus gives a bibliographical reference which seems incomplete: ‘Hanc sapientiam sequentibus pulchre tradiderunt apud Graecos ille, apud Latinos Marcellus’. Apparently Cassiodorus had to check on the Greek writings on the art of finding water, which he eventually failed to do.



### III METHOD

#### 3.1 Distributed word representations

The core of the method used in this paper is based on the freely available *word2vec* tool, which can be used to create high quality word embeddings based on a large corpus of text [Mikolov et al., 2013a].<sup>9</sup> Such methods are frequently based on the theory of distributional semantics, which can be summarised by the distributional hypothesis: ‘You shall know a word by the company it keeps’ [Firth, 1957]. The result of running the *word2vec* tool on a given collection of texts is a collection of high-dimensional vectors, representing the words in this space. Such word representations thus map words from a text to a space, which has the desirable property that words which are somehow related to each other are close to each other in this space. For instance, colours (‘red’, ‘green’, ‘blue’, etc.) tend to be close to each other in the models generated by many distributional semantics methods. Conversely, unrelated words will be far away from each other in this space.

In so-called ‘count-based’ methods of obtaining distributed word representations, each dimension (originally) represents the occurrence of a given word within a given context window. More modern approaches, however, are based on predicting a given word given a context, and thus each dimension represents some latent semantic feature. These ‘prediction-based’ methods have been found to generally outperform the count-based methods [Baroni et al., 2014]. Later research has, however, dispelled these seemingly magical properties, and found that they can all be attributed to hyper-parameters present in, e.g., the *word2vec* tool [Levy et al., 2015]. Nonetheless, *word2vec* remains a useful tool, and is remarkably efficient when it comes to learning such distributed word representations.

Most often when training embeddings, one uses a training corpus consisting of millions or billions of tokens. Since our training corpus is relatively small, parameter selection for *word2vec* must be done with this in mind. The standard *word2vec* settings include a token cut-off frequency, meaning that tokens occurring less than five times are excluded from the model. In the case of the *Variiae*, several of the historical characters in which we are interested only occur a handful of times (i.e.,  $< 5$ ). Because of this, we remove the cut-off criterium, and allow vectors to be generated for all words. Considering the relatively small size of the corpus, and the resulting low amount of types (approximately 24,000), this does not lead to an unmanageably large amount of vectors, which would likely be the case if training on, e.g., the enormous corpus of Latin texts by Bamman and Smith [2012], as was done by Bjerva and Praet [2015].

*Word2vec* offers two training models: a continuous bag-of-words model (cbow) and a skip-gram model. These two models have many similarities, but differ fundamentally in what their training objectives are. While a cbow model learns word representations by attempting to maximise the probability of predicting the correct word given a context, the skip-gram model does the inverse. That is to say, the skip-gram model maximises the probability of predicting the correct context, given the current word [Mikolov et al., 2013a]. We use the skip-gram model for our experiments. This choice was made considering the fact that the skip-gram model generally performs better than the cbow model on smaller datasets, and is better for infrequent words [Mikolov et al., 2013a].

Subsampling of high frequent words is useful, e.g., in order to obtain speedup when training word embeddings [Mikolov et al., 2013b]. Considering the relatively small size of our training corpus, we do not use subsampling.

---

<sup>9</sup><https://code.google.com/archive/p/word2vec/>



Figure 1: Illustration of the embeddings obtained with our model, visualised with t-SNE. The words shown are *amore* (love), *honesta* (honesty), *fide* (faith) and *inimica* (enemy).

We further use a context window size of 6 words. The embeddings are trained on the corpus over the course of 200 epochs. We set the vector dimensionality to 100. The word representations obtained after applying this method can be visualised using an algorithm for visualising high-dimensional data, such as t-SNE [Van der Maaten and Hinton, 2008]. Using t-SNE, the 100-dimensional representations we have are reduced to two for illustrative purposes. Figure 1 shows an excerpt of one region of this reduced space, with some related words close to each other (*love*, *honesty*, *faith*) and, interestingly, relatively close to the word *enemy*.

### 3.2 Heatmap generation

There exists a large number of different methods for calculating distances between vectors. For many tasks within NLP, task-specific distance measures have been found to increase system performance (see, e.g., Kotlerman et al. [2010], Lenci and Benotto [2012]). However, within a relatively simple framework such as the one used in this work, we have chosen to use the cosine similarity measure. This measure is both computationally and intuitively simple, and is frequently used in state-of-the-art applications in NLP (see, e.g., Mikolov et al. [2013c]).

We calculate the relatedness between (PERSON, CONCEPT) and (PERSON, PERSON) pairings as follows: For each concept we amass a set of vectors based on the related Latin words. For each person we use the vector representation in our model based on all occurring conjugations and variants of the person's name. We then find the cosine distance between the centroid of all vector representations of each pairing. We take this distance to be a measure of the relationship between a person of interest and the concept in question.

In order to open up for an interpretation of the models output, these (PERSON, CONCEPT) and (PERSON, PERSON) relations are illustrated using heatmaps. In doing so, cosine similarities between each pair of centroids is mapped to a colour range, which allows for analysis by a historian. We use the plotly library to generate these heatmaps.<sup>10</sup>

### 3.3 Network generation

We employ network analysis in order to place the correspondence between characters in the *Variae* into perspective. We first build a co-occurrence matrix using the letters' tags of (SENDER, RECIPIENT) pairs. This matrix is then converted into a graph using the Fruchterman-Reingold layout algorithm [Fruchterman and Reingold, 1991], in the implementation found in the Graph-Tool library [Peixoto, 2014].

In order to enrich the representation obtained in the networks further, we combine this with our heatmap approach. By doing this, we obtain networks which both embody the correspondence patterns found, as well as the associations found between correspondents and certain characters.

<sup>10</sup><https://plot.ly>



### 3.4 Method of analysis

Table 2: Persons of interest used in the study, along with personal details.

Name	Status	Lifetime
Agapetus	pope	? – 536
Alaricus	Visigothic king	c. 370 – 410
Anastasius	Byzantine emperor	c. 431 – 518
Athalaricus	Ostrogothic king of Italy	516 – 534
Boethius	scholar, Ostrogothic official	c. 480 – 524
Cassiodorus	scholar, Ostrogothic official	c. 485 – c. 585
Iustinianus	Byzantine emperor	c. 482 – 565
Liberius	Ostrogothic/Roman official	c. 465 – c. 554
Odoacer	barbarian general, king of Italy	433 – 494
Symmachus	mecenas, Ostrogothic official	? – 526
Theodahadus	Ostrogothic king of Italy	c. 480 – 536
Theoderic	Ostrogothic king of Italy	454 – 526
Theodora	Byzantine empress	c. 500 – 548

Table 3: Relevant concepts used in the study, with examples of related Latin words.

Concept	Words			
Modernity	Modernus	Novus	Novitas	...
Romanness	Romuleus	Quirites	Latialis	...
Greekness	Graecus	Graeculus	Atticus	...
Gothness	Gothus	Hamalus	Gothicus	...
Antiquity	Vetus	Antiquitas	Senex	...
Liberty	Libertas	Libertatus	Liber	...

We will ascertain how Cassiodorus, his colleagues and contemporary actors in the political patchwork of late antiquity (Table 2) are associated in the *Variae* with six ideological concepts central to political and cultural profiling in late antiquity (Table 3). In order to do this, we will ascertain the role important Latin authors, Greek authors, and historical characters (Table 4) mentioned in the *Variae* play in the association between a person and an ideological value.

We consider to be significant: 1) very strong associations, 2) very weak associations and 3) great differences in the measure of association. If a certain association is equally weak or strong, we consider this to be insignificant. In the case of a strong (negative or positive) association between person X and ideological concept A, this could also mean that person X is in the same strong measure (either positively or negatively) associated in reverse with ideological concept  $-A$ . This could be the case when a specific ideological concept is opposed to another ideological concept. We assume the following pairs possibly to be opposed: Modernity-Antiquity, Romanness-Greekness, Romanness-Gothness, Greekness-Gothness.

### 3.5 Interpretation of embeddings

The core of the methodology employed in this paper is based on the usage of word embeddings. The algorithm used to compute our embeddings is a distributional one, meaning that the meaning (and representation) of words is gathered from the contexts in which they appear. Since we use texts from only one author (i.e. Cassiodorus) to generate our embeddings, the

<sup>11</sup>[Giardina et al., 2015, p.304]

<sup>12</sup>[Halporn et al., 2004, p.175 n.11]

Table 4: Selection of characters mentioned in the *Variae*, along with personal details: Latin Authors, Greek Authors, Historical Characters

Category	Name	Status	Lifetime
<b>Latin</b>	Cicero	politician, rhetorician, writer	106 – 43 BC
	Horace	poet	65 – 8 BC
	Juvenal	poet	End 1st c. – beginning 2nd c. AD
	Marcellus	author of technical treatises	? <sup>11</sup>
	Pliny the Elder	army commander, scientist	23 – 79 AD
	Pliny the Younger	lawyer, magistrate, writer	61 – 113 AD
	Tacitus	senator, historian	ca. 56 – after 117 AD
	Terentian	grammarian	fl. End 2nd c. AD
	Vergil	poet	70 – 19 BC
<b>Greek</b>	Archimedes	scientist, engineer	ca.287 – ca.212 BC
	Aristotle	philosopher, scientist	384 – 322 BC
	Euclid	mathematician	fl. ca. 300 BC
	Helenus	grammarian	? <sup>12</sup>
	Hero of Alexandria	mathematician, engineer	ca. 10 – ca. 70 AD
	Homer	poet	?
	Nicomachus of Gerasa	mathematician	c. 60 – c. 120 AD
	Metrobius	scientist, engineer	?
	Plato	philosopher	428/427 or 424/423 – 348/347 BC
	Ptolemy of Alexandria	scientist, mathematician	c. 100 – c.170 AD
	Pythagoras	philosopher, mathematician	c. 570 – c. 495 BC
<b>Historical</b>	Cato the Elder	senator, historian, moralist	234 – 149 BC
	Cato the Younger	politician, rhetorician, moralist	95 – 46 BC
	Metellus	various generals and politicians	republican period
	Pompey	army commander, politician	106 – 48 BC
	Trajan	Roman emperor	53 – 117 AD
	Valentinian I	Roman emperor	321 – 375 AD
	Valentinian II	Roman emperor	371 – 392 AD
	Valentinian III	Western Roman emperor	419 – 455 AD

representations can be interpreted to embody how Cassiodorus himself presented matters. In detail, we interpret this to capture both explicit and implicit, as well as conscious and subconscious similarities and associations between, e.g., persons and concepts in his texts. Using this interpretation, we can use the measured similarities between certain persons as an indication of Cassiodorus' view on these people. Furthermore, we can measure the similarity between historical characters (e.g. Cicero), and investigate the similarity between him and the letters of Cassiodorus themselves. In doing so, we abstract away from the surface forms of these letters, which are used for the detection of concrete intertextuality, and observe a form of conceptual intertextuality. This conceptual intertextuality is thus not found in explicit citations, but in the relatedness between a text and a different author.

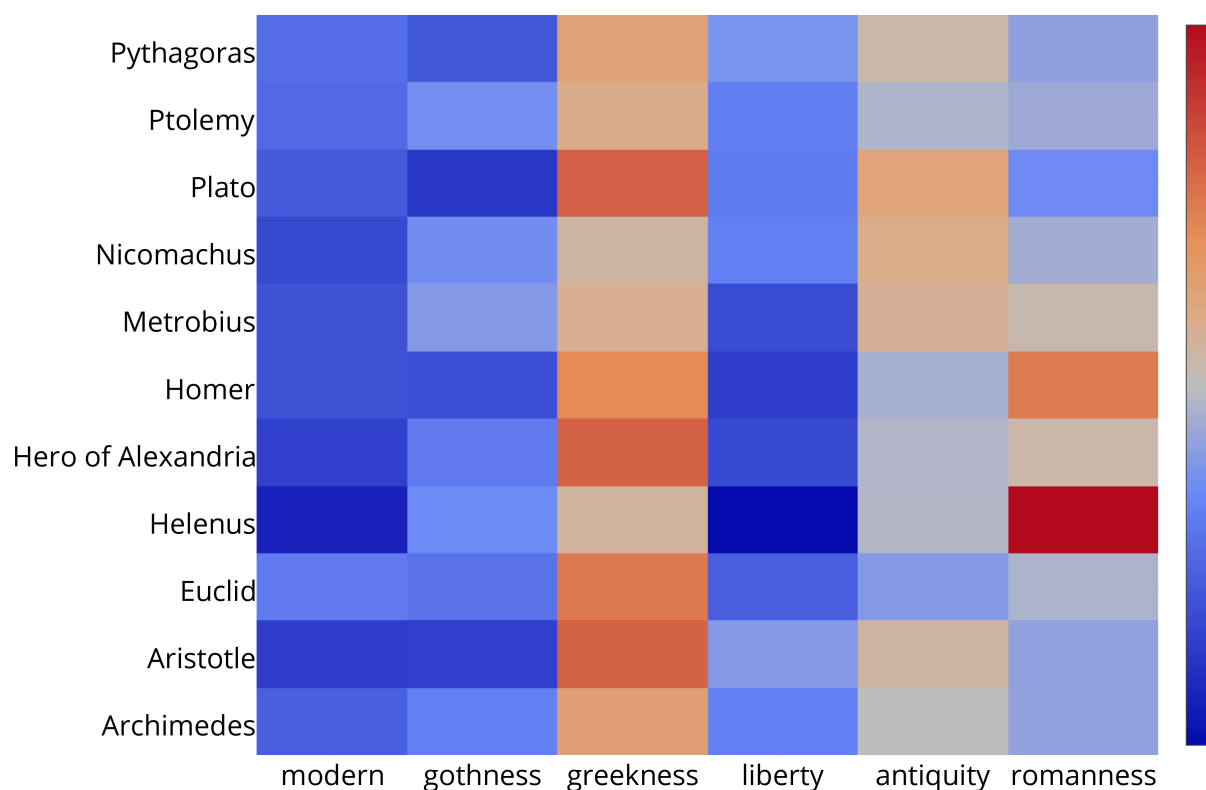


Figure 2: Associations between Greek authors and our selection of ideological concepts

## IV ANALYSIS AND DISCUSSION

### 4.1 Heatmap analysis: Historical and literary resonances in the political image building in Cassiodorus' *Variae*

A preliminary digital analysis of the association in Latin texts between six key ideological concepts in late antiquity on the one hand, and two sets of contemporaries of Cassiodorus mentioned in his *Variae* on the other hand, has shown two basic tendencies [Bjerva and Praet, 2015]. First, late antiquity saw the gradual assimilation of so called 'barbarian' elements in the landscape of late antique ideology. Whereas the first barbarian generals and kings Odoacer and Alaric were notoriously associated with their 'Gothness', this association dramatically declines under the dynasty of the Ostrogoths, who were known for their intensive profiling as Roman rulers. Second, the reverse tendency can be observed as regards the Roman emperors from the east-Roman empire; in comparison with his imperial predecessors, the emperor Justinian is in high measure dissociated from the Roman legacy [Bjerva and Praet, 2015].

With the help of a digital analysis of the presence of canonical authors and historical characters mentioned by Cassiodorus in the *Variae*, we can ascertain in two steps whether literary and historical resonances are used by Cassiodorus in the *Variae* to contribute to or differentiate from these two tendencies of association. In order to do this, we compiled three sets of characters of ideological significance mentioned in the *Variae*, namely Latin canonical authors, Greek canonical authors and historical characters (Table 4).

First we analyzed to what degree these characters are associated with the six key ideological concepts in the *Variae*. The results can be found in Figures 2, 3, and 4. In general, these results corroborate the general assumptions, both in late antiquity and in scholarly literature, which characters are emblematic embodiments of which ideological value. For instance, among Latin

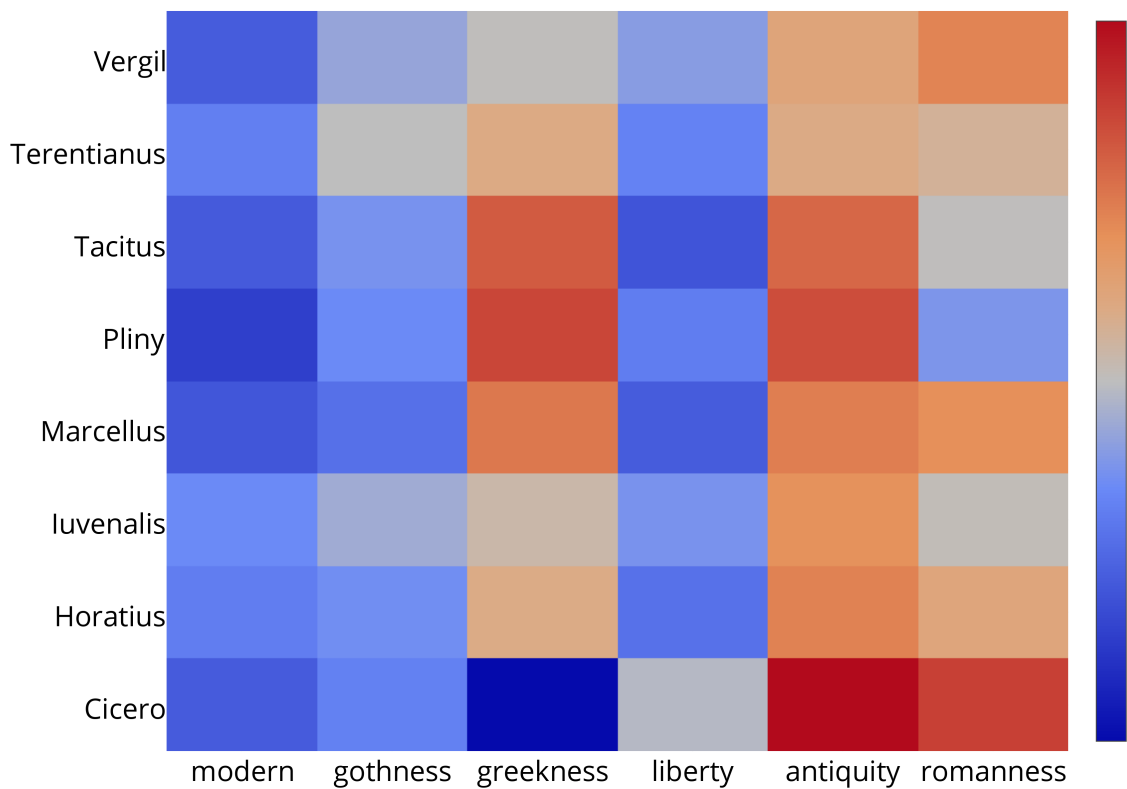


Figure 3: Associations between Latin authors and our selection of ideological concepts

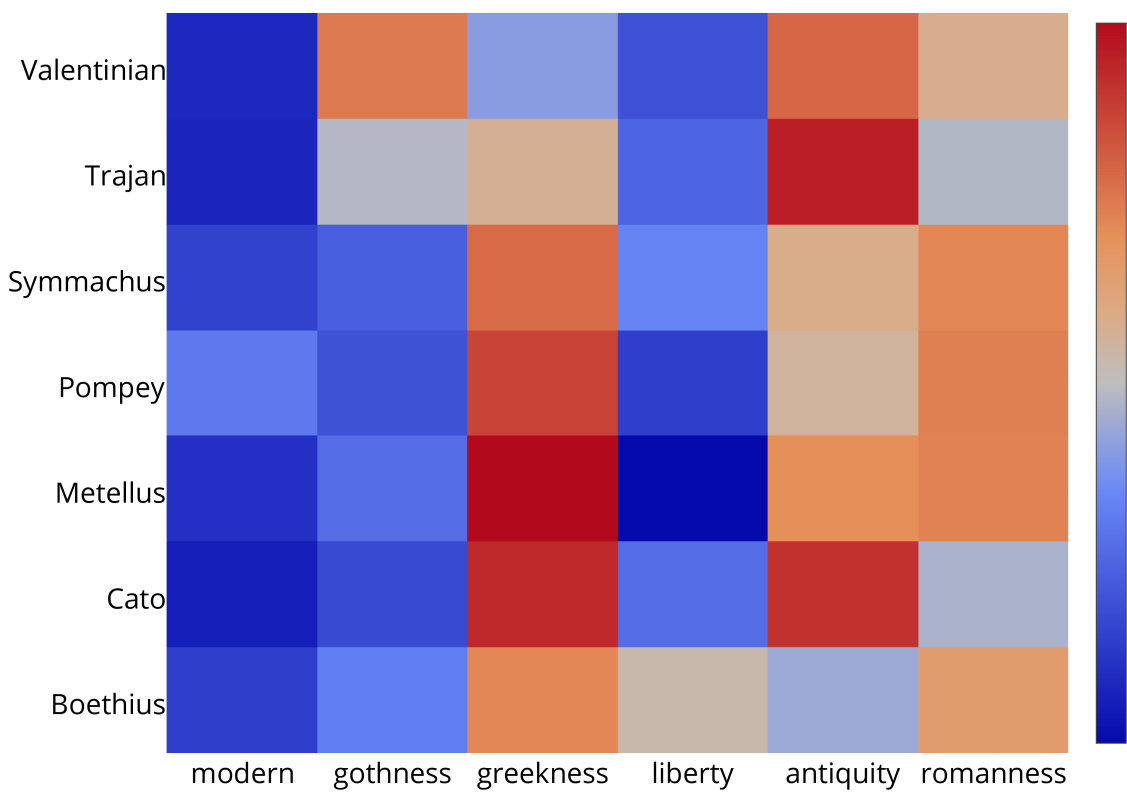


Figure 4: Associations between Historical characters and our selection of ideological concepts

authors, we see Cicero being associated to a high degree with both the concepts antiquity and Romanness, whereas authors such as Vergil, and to a lesser degree, Horace, Marcellus and Terentianus are highly associated with Romanness. Apart from Cicero, also authors such as Pliny (the Elder and the Younger) and Tacitus are, to a lesser degree, associated with venerable antiquity. Among historical characters mentioned in the *Variae*, Cato (the Elder and the Younger), Trajan and Valentinian<sup>13</sup> are highly associated with antiquity. Some peculiarities do however arise in the analysis of the historical characters. Whereas one would assume an exemplary emperor as Trajan to be highly associated with Romanness, he appears neither as a significantly Roman nor as a significantly Greek emperor. As the different Catones from Roman history presented themselves specifically as Romans in opposition to Greek decadence, we could assume a high association between Cato and Romanness. Yet, surprisingly enough, in the case of Cato, there is a high association with Greekness instead of Romanness. We could explain this phenomenon as Cato being to a high measure associated negatively with Greekness.<sup>14</sup>

Second, we ascertain to which measure the contemporaries of Cassiodorus are associated with the above mentioned authors and historical characters as if these authors and characters were ideological concepts. The results can be seen in Figures 5, 6, and 7. Our hypothesis is that, if a person is highly associated with ideological value X, he will likewise in a high degree be associated with historical characters or authors who are also associated with ideological value X. If this is the case, we could infer that intertextual or historical allusions by Cassiodorus do in fact enhance the profiling of the contemporaries of Cassiodorus in his *Variae*.

In order to ascertain this hypothesis, we choose from the three sets of characters the most conspicuously profiled persons, namely Cicero (high association Romanness and antiquity), Vergil (high association Romanness) and Cato, Trajan and Valentinian (high association antiquity). In general, we could say the two tendencies as summarised at the outset of the analysis are also reflected in the associations between the contemporaries of Cassiodorus and the authors/historical characters.

First, whereas the first ‘barbarian’ rulers of Rome, Alaric and most conspicuously Odoacer are dissociated from symbols of Romanness such as Cicero and Cato, the Ostrogothic kings (Theodoric, Athalaric and Theodahad) are in general more intensively associated with these historical characters and the Romanness they convey. Yet with regards to the Ostrogothic kings, the picture becomes also more nuanced; Theodoric is overall best associated with the Roman characters Cato and Cicero, and these associations become less intense as the Ostrogothic dynasty proceeds from Theodoric to his successor Athalaric and from Athalaric to his successor Theodahad. One notable exception is Athalaric’s high association with Cicero, which will be

---

<sup>13</sup>As regards Cato, Pliny and Valentinian, one has to take into account that these names can refer to several historical characters. The two most important Catones were Marcus Porcius Cato and Marcus Porcius Cato Uticensis. Pliny can refer to Gaius Plinius Secundus and Gaius Plinius Caecilius Secundus, whereas there were three late antique emperors under the name Valentinian. Although the appearance of the name can imply the existence of different historical characters, we consider this not to inhibit the analysis as all representatives of the name broadly convey the same ideological associations; this likeness in ideological representation was enhanced by the historical characters themselves. For instance, Cato the Younger meticulously styled himself on his ancestor to cultivate the values of Romanness as opposed to the decadent Greek world, republicanism and moral sternness. Pliny the Younger was closely related to Pliny the Elder and their works are easily associated within the same thrust of imperial intellectualism. As regards the three Valentinians, it was a common late antique practice of emperors to model their stereotype image intensively on predecessors.

<sup>14</sup>The same mechanism of negative association accounts for the fact that Odoacer is highly associated with Romanness, in spite of being the ‘barbarian par excellence’ who deposed the last emperor Romulus Augustulus [Bjerva and Praet, 2015].

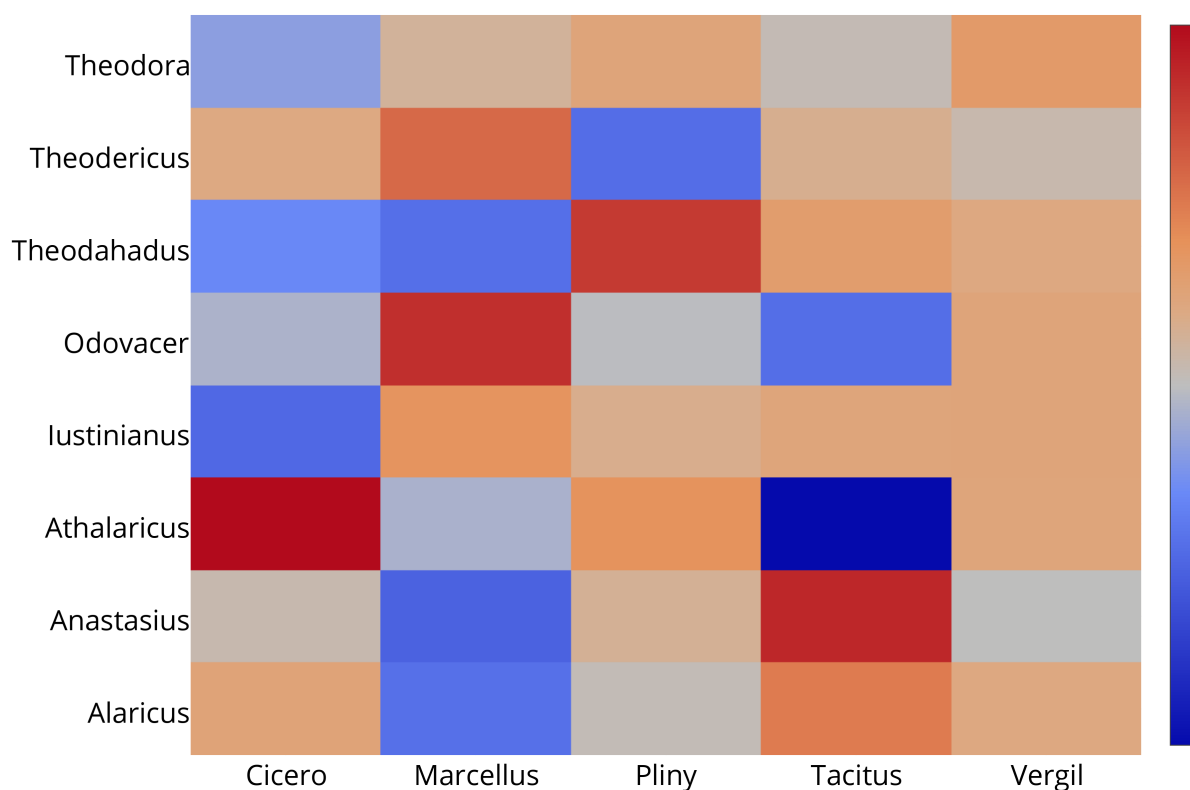


Figure 5: Associations between persons of interest in the *Variae* and Latin authors

explained in section 4.3. Perhaps these declining associations with symbols of the Roman legacy are a symptom of the gradual loss of authority of the Ostrogothic dynasty as a worthy heir to the Roman emperors. Theoderic remained the strongest exponent of this dynasty, which soon after his decease dissipated into wars and conflicts with the Roman senate.

The second tendency, namely the tendency of the eastern Roman emperors losing touch with a Roman image, is also corroborated by the lack of association between Anastasius and Justinian, both eastern Roman emperors on the one hand and Cicero and Cato on the other hand. Considering Vergil and Trajan, we can perceive how these characters are equally associated with all contemporaries of Cassiodorus, be it barbarian generals, Ostrogothic monarchs or eastern Roman emperors. In the case of Vergil, one could assume that the most important poet of Latin literature conveyed a general association which was too general to be cultivated by a specific individual. In comparison, both Cicero and Cato were politicians whose literary activities automatically also conveyed political meaning. In the case of Trajan, we could argue that, by the 6th century, his specific ideological meaning erodes as he becomes part of a stereotypical list of ‘good emperors’.<sup>15</sup>

<sup>15</sup>Trajan was one of the ‘Five good emperors’. Although the term was first coined by Machiavelli in his *Discourses on the First Decade of Titus Livy*, I.10, there were lists of good emperors circulating from late antiquity onwards. For instance, in his *Praise of the Emperor Anastasius*, vv. 45-49, the 6th-century Priscian of Caesarea compares Anastasius favorably to Antoninus Pius, Marcus Aurelius, Nerva, Titus and Trajan. The same mechanism is applied by another contemporary of Cassiodorus, John of Lydia, when he says Justinian emulates Trajan, Augustus, Titus and Marcus Aurelius in his treatise *On the Magistracies of the Roman State*, II.28.



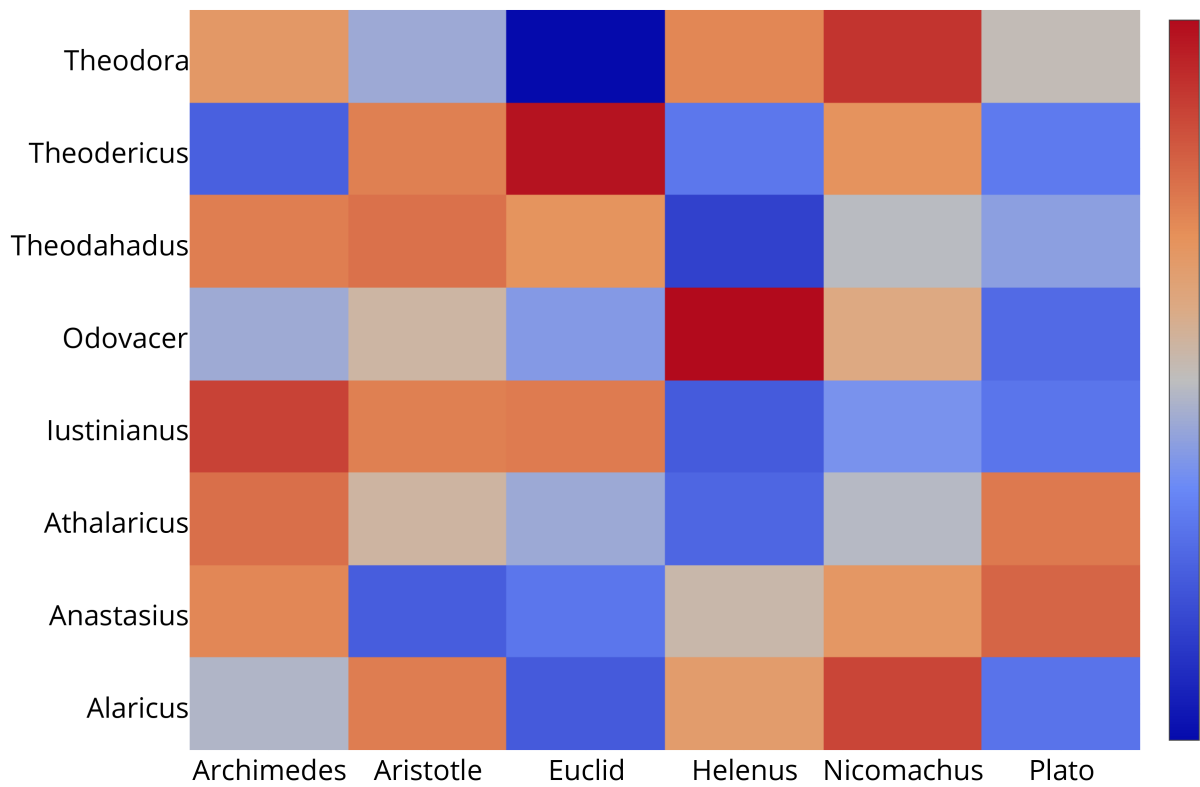


Figure 6: Associations between persons of interest in the Variae and Greek authors

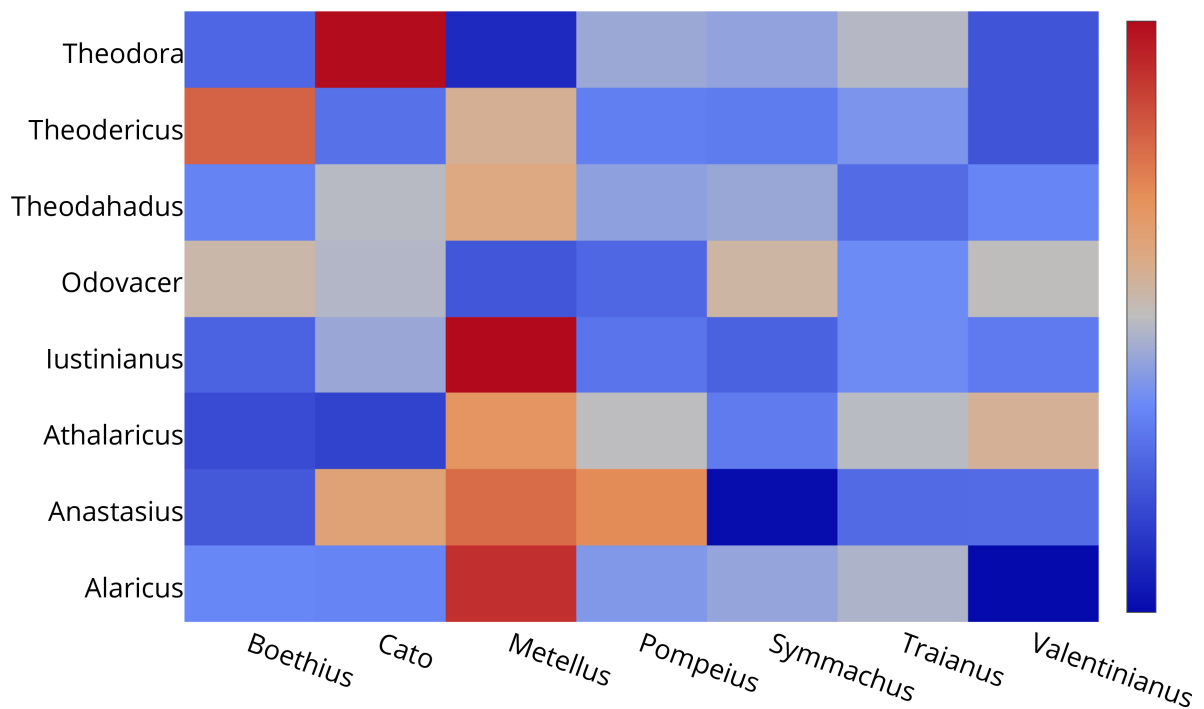


Figure 7: Associations between persons of interest in the Variae and Historical characters

#### 4.2 Heatmap analysis: Intellectual profiling through literary associations: Cassiodorus, Symmachus and Boethius

A preliminary analysis of the association of Cassiodorus and his competitors in the intellectual and political network of the 6th century on the one hand with the six key ideological concepts on the other hand has only shown a general divide. Higher aristocrats such as Symmachus and Boethius were still associated with the Greek cultural sphere of the empire whereas intellectuals and politicians such as Cassiodorus himself and Liberius (ca. 465–ca. 554) already foreshadowed the exclusive focus on Latin in the late antique and early medieval western half of the former Roman empire [Bjerva and Praet, 2015]. In this paper, an analysis of how Cassiodorus and his colleagues were associated with specific Greek and Latin authors in the *Variiae*, such as can be seen in Figures 8, 9, and 10 will allow us a more nuanced view of how Cassiodorus presented his colleagues and himself as late antique intellectuals.

Such an analysis of intellectual profiling through the association with Latin and Greek authors will focus predominantly on the triad Cassiodorus, Symmachus and Boethius. Symmachus and Boethius were central characters of the intellectual life of Italy under the Ostrogoths [Heather, 1993]. Symmachus, the father-in-law of Boethius, was the offshoot of the influential clan of Symmachi. He was as well a patron of literary activity as an author himself. He was involved in the publication of the commentary by Macrobius on Cicero's *The Dream of Scipio*, a philosophical treatise on the universe, and also composed a now lost *Roman History*. Boethius is mostly known for his vast enterprise of translating Greek philosophical works in Latin, aside from composing philosophical treatises himself. Both Boethius and Symmachus eventually fell out with king Theodoric, who had them executed in 524 and 526 respectively. Boethius and Symmachus were of special interest to Cassiodorus; they are addressees of some of his most elaborate letters,<sup>16</sup> and Cassiodorus even tries to establish his kinship with the two aristocrats in his highly tendentious autobiography, the *Ordo Generis Cassiodorum* [O'Donnell, 1979]. There are several possible reasons for Cassiodorus' attempts at associating himself with Boethius and Symmachus; perhaps he wanted negatively to dissociate himself from any complicity in their execution, which upset not only the senatorial elite in the city of Rome, but also the surroundings of the imperial court at Constantinople, where Cassiodorus hoped to regain his position after the conquest of Italy by the Byzantines. In a positive way, the association between Cassiodorus and leading intellectuals under the Ostrogoths served to shed a positive light on Cassiodorus and the Ostrogothic regime he was part of [Bjornlie, 2009, p. 150-152], [Bjornlie, 2012, p. 163-184]. In spite of Cassiodorus' attempts, however, his efforts were not reciprocated; neither Symmachus nor Boethius mention Cassiodorus in their works.

The digital analysis of the image building of Cassiodorus, Symmachus and Boethius in the *Variiae* can shed further light on the relationships between these three intellectuals, which appear to be not that straightforward. When we compare the differences between the three contemporaries as regards their associations with Latin authors, we can see how Cassiodorus is at pains to create for himself an intellectual profile which is notably differentiated from the pair Symmachus-Boethius; those authors which are highly associated with Boethius and Symmachus (namely Pliny and Vergil) are associated to a lesser degree with Cassiodorus an vice versa. Cassiodorus' associations with Cicero, Marcellus and Tacitus do not appear for Boethius and Symmachus.

Furthermore, not only does Cassiodorus profile himself as an intellectual distinct from Symmachus and Boethius, the associations which Cassiodorus does cultivate also imply an emula-

<sup>16</sup>Letters I.10, I.45 and II.40 are addressed to Boethius, whereas letter IV.51 is addressed to Symmachus.

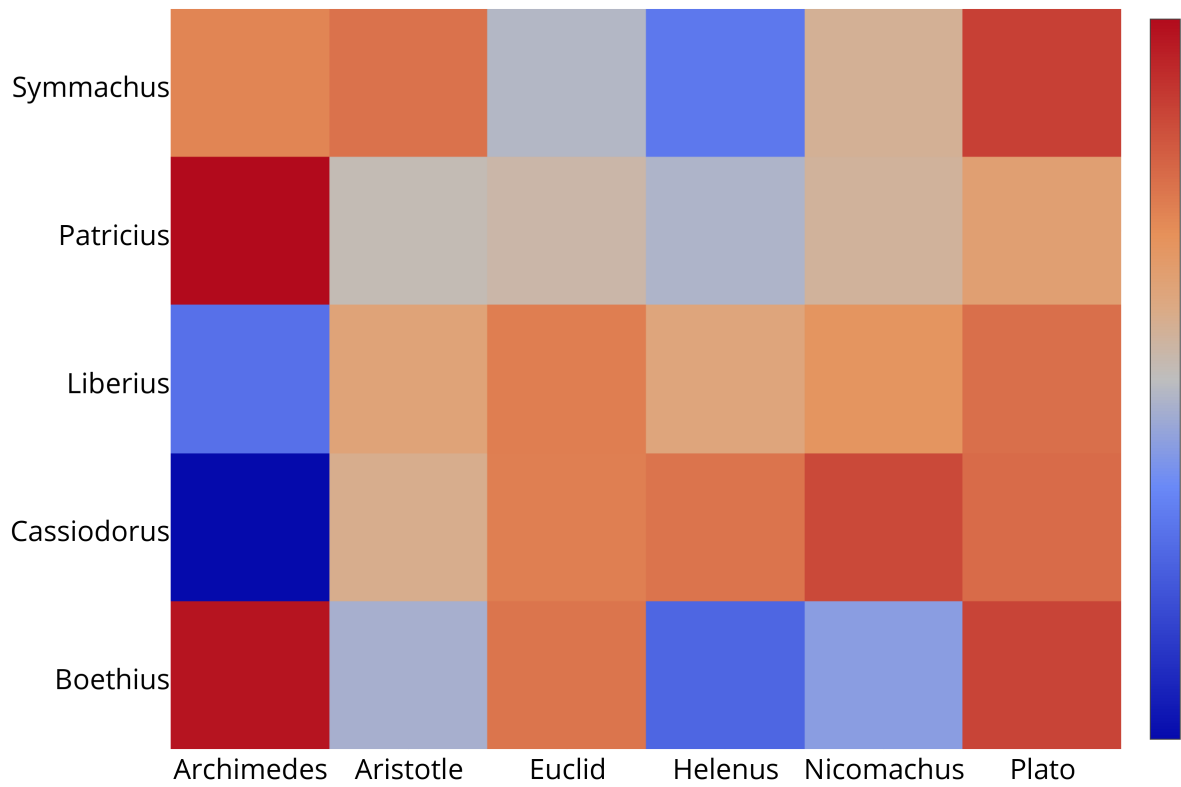


Figure 8: Associations between Cassiodorus and his peers, and Greek authors

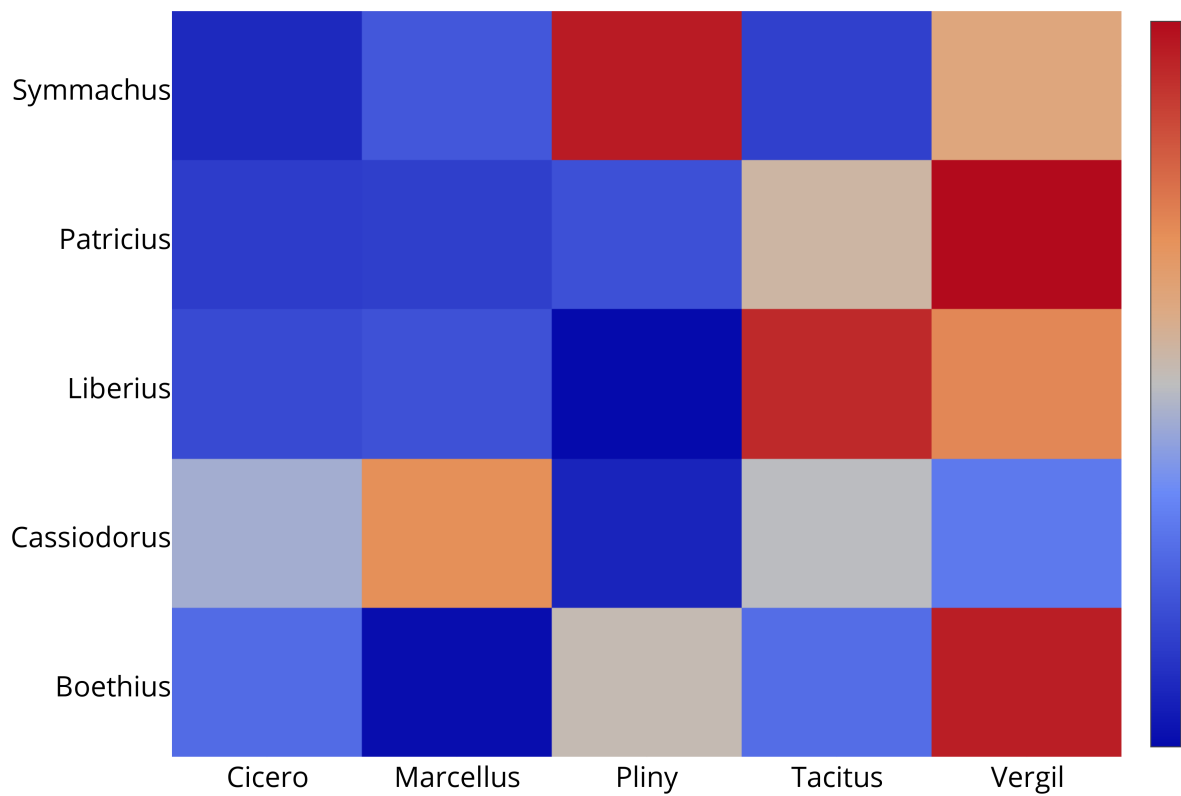


Figure 9: Associations between Cassiodorus and his peers, and Latin authors

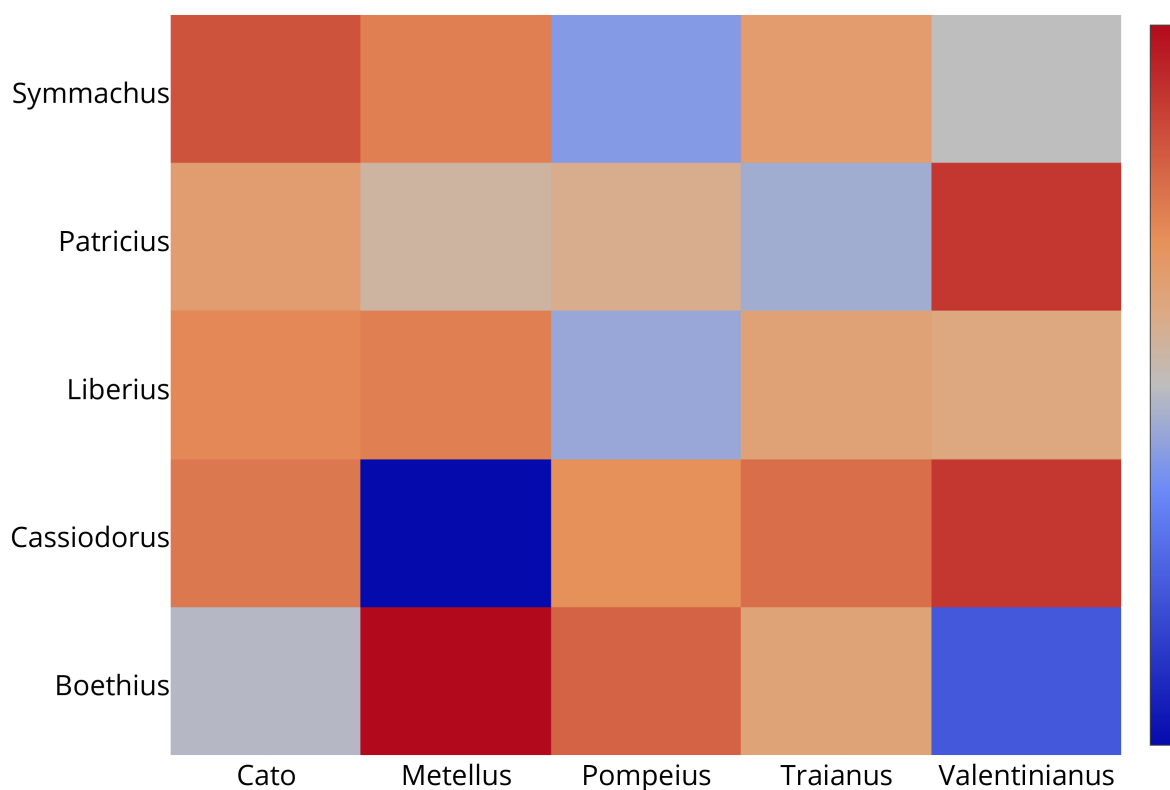


Figure 10: Associations between Cassiodorus and his peers, and Historical characters

tion of the pair. The two aristocrats are highly associated with Virgil, yet this author is also associated in a general, more neutral way with all political players mentioned in our analysis above; apparently, Boethius and Symmachus are only associated as ‘standard intellectuals’ with general emblems of learning such as Virgil. Cassiodorus is, through his associations with the orator Cicero and the historian Tacitus furthermore presented as a better alternative to Symmachus. Symmachus conspicuously lacks the associations with Cicero, in spite of editing a commentary on a treatise by this author, whereas Cassiodorus is associated with Cicero as a symbol of Roman eloquence. Symmachus also lacks the associations with Tacitus, the Roman historian par excellence, in spite of the fact that he derived most of his prestige from his own *Roman History*, which was modeled on Tacitus. Cassiodorus, who also wrote works of historiographical interest, is, however, associated with Tacitus. The implications of these differences in associations are clear; Cassiodorus first presents himself as an intellectual distinct from Boethius and Symmachus, but also as an intellectual superior to them. Cassiodorus does not only appear as a better rhetorician, but also a better historian than Symmachus.

As regards the associations between Cassiodorus, Symmachus, Boethius, and the Greek authors, also in this case Cassiodorus is seen to improve his intellectual profile in comparison with Symmachus and Boethius. Cassiodorus retains a generally high association with all authors which enhances his broad intellectual profile in comparison to Boethius and Symmachus – and in this sense the analysis profoundly nuances the image of a strictly Roman and non-Greek image of Cassiodorus. Boethius and Symmachus lack this general association with all Greek authors, an association which they were presumed to have as parts of the aristocratic elite with interests in both the Greek and Latin cultural spheres of the Roman world. In some cases, the absence of any associations is telling in how Cassiodorus presented both his competitors; note the low association between Boethius and the philosopher Aristotle, whereas Boethius

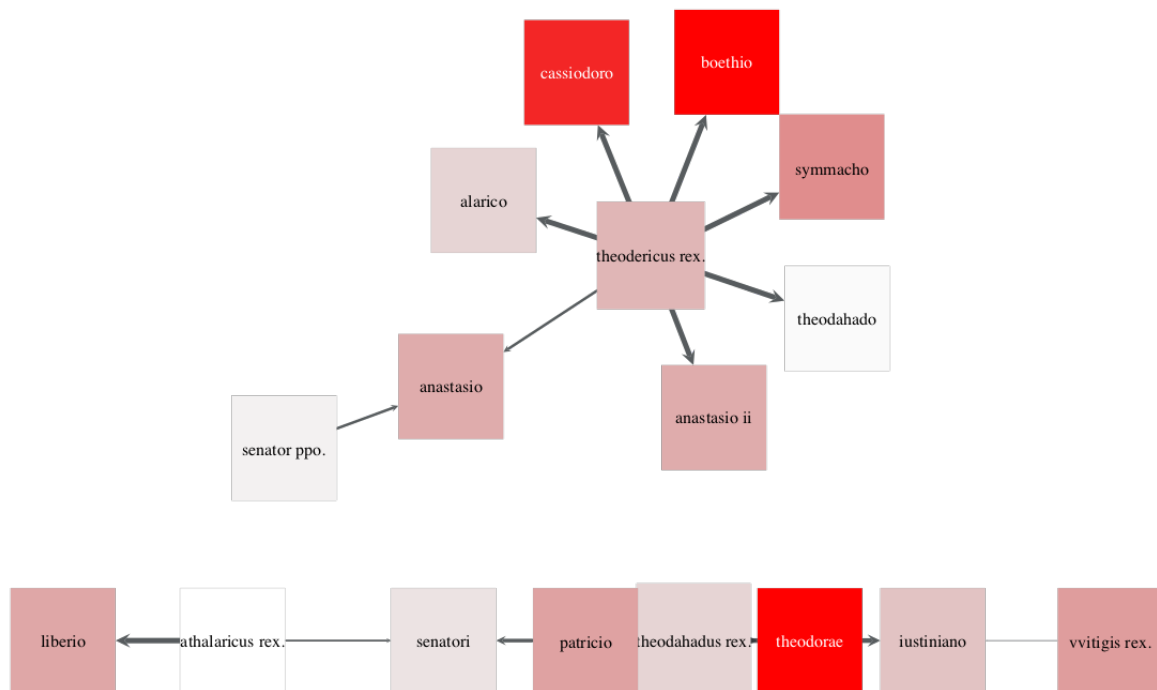


Figure 11: Network of the *Variae* – Association with Cato

translated several of Aristotle’s works!

The digital analysis of the use of literary and historical resonances has brought to light Cassiodorus’ multi-layered approach to his colleagues and competitors in the late antique political patchwork; historical and literary evidence suggests Cassiodorus is at pains to cultivate the association with aristocrats such as Symmachus and Boethius in order to improve his profile. Yet the analysis of the heat maps shows signs of the contrary; Cassiodorus carves out for himself in his *Variae* a profile which is not only significantly different from the profile of Boethius and Symmachus, but also superior.

### 4.3 Network analysis

The third section of the analysis will combine the method of ascertaining associations through heat maps with a traditional network analysis. The nodes in the network are connected to each other on the basis of who receives a letter from whom in the *Variae*, for which we used the titles preceding each letter. These titles follow a stereotypical formula: ‘To person X from king Y’. The result of the network analysis is a set of clusters with at the center the person on whose behalf the letter was sent – either an Ostrogothic king or Cassiodorus himself in the capacity of praetorian prefect. Arrows connect the sending center of the cluster with the nodes receiving the letters – the thickness of the arrow represents how many letters are addressed to the receiving person in the *Variae*. The color of the nodes represents, as in the heat maps, to what extent a node in the network is associated with a certain Greek author, Latin author or historical character. As regards the figures with whom the nodes in the network are associated, we selected the same figures which appeared in sections 4.1 and 4.2. to be the most emblematic emblems of Romanness and antiquity, namely Cato, Cicero, Trajan, Valentinian, Virgil – and to a lesser extent, Pliny and Tacitus. The results can be seen in Figures 11, 12, 13, 14, 15, 16, and

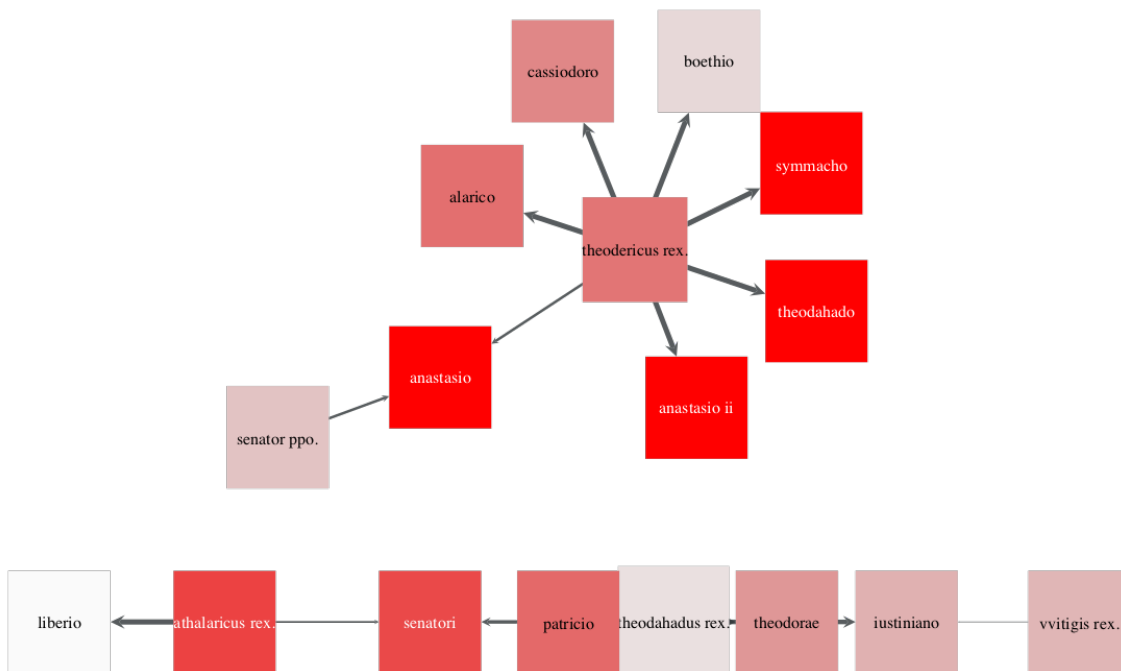


Figure 12: Network of the Variae – Association with Traianus

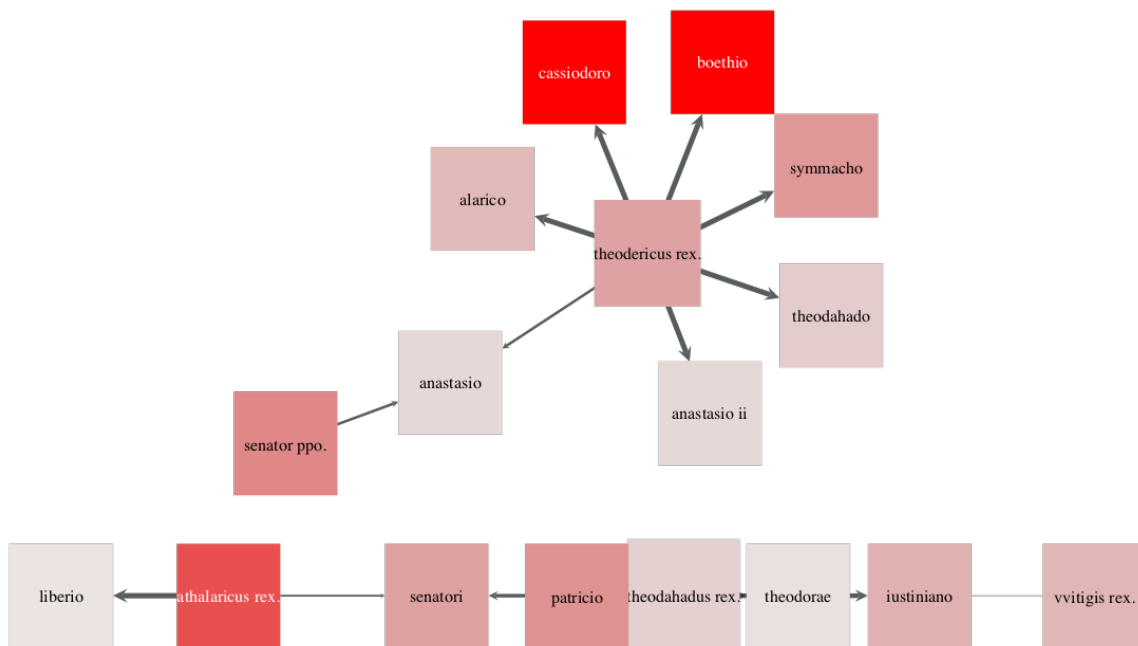


Figure 13: Network of the Variae – Association with Valentinianus



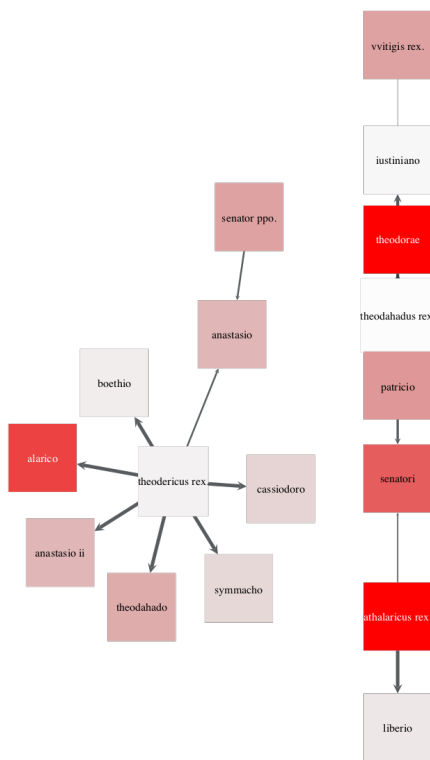


Figure 14: Network of the Variae – Association with Cicero

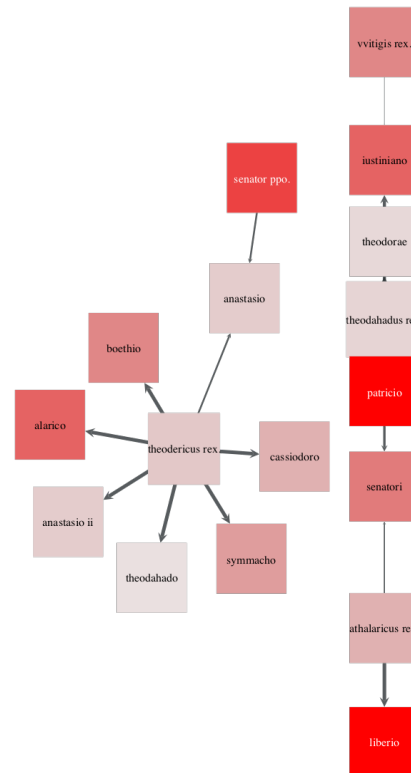


Figure 15: Network of the Variae – Association with Vergil

17. We will focus, in this analysis, on the profiling of the Ostrogothic kings on the one hand and the profiling of Cassiodorus in comparison with his colleagues and competitors.

An analysis of the network and associations of the three kings Theodoric, Athalaric and Theodahad shows the following results. The associations between the kings and political emblems of the Roman legacy (namely Cato, Trajan and Valentinian), helps us to nuance the tale of decline of the Ostrogothic dynasty after Theodoric. Indeed, it is true that Theodoric and Alaric are more associated with these political emblems than Theodahad is, which would indicate a decline of the Roman image of the dynasty at its end. Yet on the other hand we can see that Athalaric is more associated with Cato, Trajan and Valentinian than Theodoric is. Apparently the Ostrogothic dynasty did not dissipate at once after the decease of its presumed champion Theodoric, but the dynasty retained and enhanced even its reputation as being Roman under the rule of Athalaric – at least in the presentation of reality by Cassiodorus in his letter collection.

When we consider how the three monarchs were presented *as intellectuals* through their associations with cultural emblems such as Cicero and Virgil, the picture of Ostrogothic representation becomes even more nuanced. For we can see how Athalaric is associated more intensely with both Cicero and Virgil than Theodoric and Theodahad. Not only is Athalaric more intensely associated with political emblems of the Roman legacy. His profile is also more differentiated in comparison with the other Ostrogothic kings as he combines both political and cultural links to the Roman legacy. This markedly cultural nature of Athalaric's Roman profile could be coupled to his education. Athalaric's mother Amalasuenta had Athalaric educated as a Roman

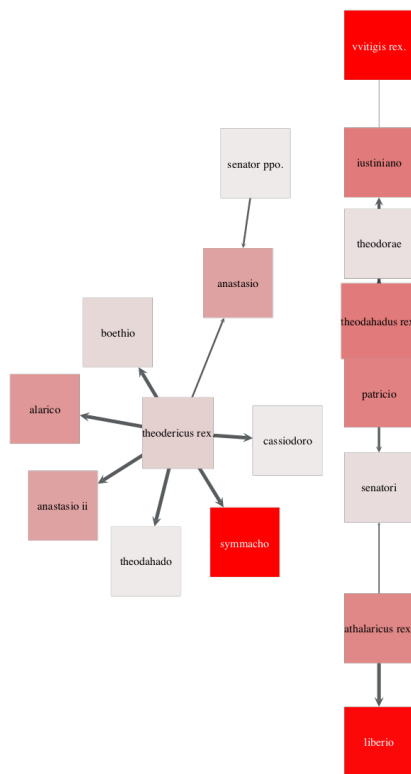


Figure 16: Network of the *Variae* – Association with Pliny

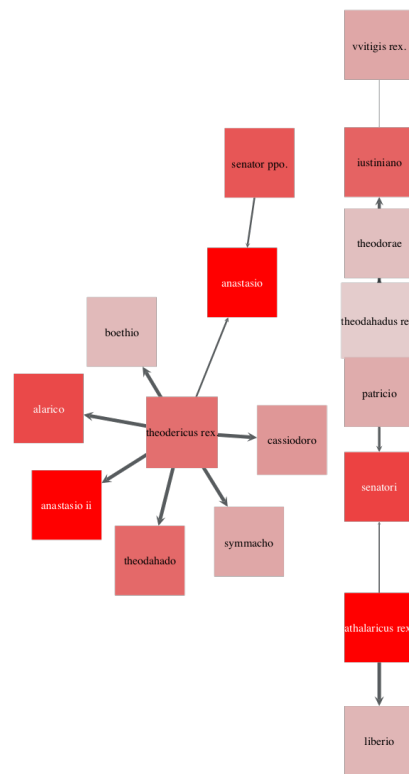


Figure 17: Network of the *Variae* – Association with Tacitus

aristocrat, a decision which caused upheaval under the Gothic military aristocracy, but which can account for his Roman cultural and political profile.<sup>17</sup>

We began our paper with the adagio of Firth [1957]: ‘You shall know a word by the company it keeps’. This principle we can actually transpose onto the presentation of the Ostrogothic kings in the *Variae*. Not only does one recognise a word by the company it keeps, but also a king by the company he keeps. When we look at the nodes surrounding the three monarchs, we can see how they corroborate the tendencies in the presentation of the monarchs sketched above. In general, the persons receiving letters from Theodoric and Theodahad are more associated with political emblems of the Roman legacy (Cato, Trajan and Valentinian) than persons surrounding Athalaric are. Reversely, the persons surrounding Athalaric in the network are in general more associated with Cicero and Virgil as literary emblems of the Roman heritage than the persons surrounding Theodoric and Theodahad are. The differences in political profiling of the monarch are therefore also reflected in the separate profiles of the persons present in the network of the monarch.

The second part of this section will return to Cassiodorus and his profiling in comparison with his colleagues and competitors. The network allows us to trace the evolution of Cassiodorus’ profile throughout his long political career. The node with the designation ‘Cassiodoro’ in the networks refers to the earlier part of his career, before his accession to the office of praetorian prefect (ca. 507–532 AD),<sup>18</sup> whereas the nodes with the designation ‘Senatori’ or ‘Senator

<sup>17</sup>On his education, see the account of Procopius, *Wars* 5.2.6-17.

<sup>18</sup>In one exceptional case, the designation ‘Cassiodoro’ refers to another Cassiodorus, the father of Cassiodorus in letter I.3. Letters I.3 and I.4 namely treat the promotion of Cassiodorus senior to the rank of patrician, with the

PPO' refer to Cassiodorus in the capacity of praetorian prefect (533 – 540).<sup>19</sup> We can see that the node representing the earlier part of Cassiodorus' career is strongly associated with the political emblems of the Roman legacy and weakly with its cultural emblems. The reverse is true for the nodes designating Cassiodorus as praetorian prefect. Apparently, Cassiodorus first profiled himself mainly through the political emblems of Roman rule whereas he later on focused more on cultural emblems such as Cicero and Virgil to give expression to his image. Perhaps this shift from the political to the cultural already foreshadowed Cassiodorus' cultural preoccupations during his redacting his letters in Constantinople.

As regards Cassiodorus' profiling in comparison with his competitors, the results of the analysis of the triad Cassiodorus-Boethius-Symmachus (section 4.2) are corroborated by the network graphs. Literary authors which are highly associated with Cassiodorus (Cicero and Tacitus) are not associated with Boethius and Symmachus and vice versa for the authors Virgil and Pliny. Furthermore, the same tendency can be seen when comparing Cassiodorus with two other competitors, namely Liberius [O'Donnell, 1981] and Peter the Patrician (ca. 500–565) [Antónopoulos, 1990]. These are more associated with Virgil and Pliny and less with Tacitus and Cicero, whereas the reverse is the case for Cassiodorus. Not only is Cassiodorus at pains to distinguish himself on a cultural plane from Boethius and Symmachus, but also from other intellectuals and politicians as Peter the Patrician and Liberius.

As regards the associations between Cassiodorus and his colleagues on the one hand and political emblems of the Roman legacy such as Cato, Trajan and Valentinian on the other hand, the picture is less straightforward – which is perhaps an indication that Cassiodorus aimed predominantly at distinguishing himself from the others in cultural terms. For instance, although he ousts Symmachus, Peter the Patrician and Liberius in his association with Cato, Cassiodorus shares this strong association with Boethius. Symmachus is more associated with Trajan than Cassiodorus but Cassiodorus is more associated with Trajan than Boethius, Peter the Patrician and Liberius. As regards the associations with Valentinian, Cassiodorus has to share his place with Boethius and Peter the Patrician, whereas the leaves Symmachus and Liberius behind.

## V CONCLUSIONS AND FUTURE WORK

The combination of different digital methods of analysis aids in efforts to nuance the image we have of Cassiodorus' ways of profiling himself and others in his *Variae*. Not only was the combination of methods beneficial to expand on preliminary work done on the digital analysis of late antique ideology (section 4.1) [Bjerva and Praet, 2015]. This integrated digital approach also nuanced some of the more debated aspects of the image of Cassiodorus and the Ostrogothic kings in 6th-century Italy. Contrary to the claims made by [Bjornlie, 2009, 2012], Cassiodorus did not only seek to be associated with Boethius and Symmachus, but also used literary and historical references in order to profile himself as an intellectual notably distinct from Symmachus, Boethius and other competitors in the political patchwork (sections 4.2 and 4.3). Furthermore, the analysis of the imago of the Ostrogothic kings served to nuance the narrative of an Ostrogothic dynasty depending solely on its champion Theodoric and dissipating quickly after his decease. We have shown how the Roman image of the dynasty was even enhanced under Athalaric before reaching its nadir under Theodahad (section 4.3).

---

first letter intended for the newly promoted Cassiodorus and the second letter addressed to the senate to notify this body of the promotion.

<sup>19</sup>Cassiodorus' full name was Flavius Magnus Aurelius Cassiodorus Senator. He was mainly known under the name Senator during late antiquity and the middle ages. PPO is an abbreviation for Cassiodorus' function at that time, namely Praefectus Praetorio, or praetorian prefect.

In short, an application of the digital method to late antique corpora of texts will provide the necessary grounds for further interpreting the political and cultural position of late antique intellectuals, scholars and politicians in the cases where the massive presence of literary attestations is accompanied by a dearth of biographical and factual data. The case of Cassiodorus is notorious; whereas the *Variae* are formidable reading, traditional historians still have to speculate on Cassiodorus' activities and political status during his stay in Constantinople. Future research could cast an equally clarifying light on other such notoriously visible and at the same time invisible historical characters, such as John of Lydia (ca. 490–ca. 565 AD), Liberius and Peter the Patrician.

Another possibility lies in the diachronic application of the methods used in this paper in order to trace the ideological redefinitions of such key values as Romanness during the late antique and early medieval periods, periods which themselves witnessed crucial transformations of the Roman empire itself [Dmitriev, 2010].

We further provided two aims of investigation for this paper, dealing with the applicability of word embeddings to otherwise qualitative research. The thorough and groundbreaking analysis that was carried out using the output of our system indicates that they are indeed an appropriate tool, and highlight the possibilities of future research in this direction.

On a theoretical level, this paper contributed through the case study of Cassiodorus' *Variae* to the redefinition of the concept of intertextuality on a more conceptual level, namely, as 'presence of an author in a literary work'. This conceptual definition is more suited to texts which are intended to different groups of readers with different grasps of the presence of literary and historical references in a literary text. Furthermore, the redefined concept of intertextuality is useful in tracing the presence of authors which are otherwise untracable through traditional intertextual research, as their texts are partially or not transmitted. In the future, different case studies with different digital methods could contribute to our understanding of the multi-layered phenomenon of intertextuality.

## ACKNOWLEDGEMENTS

We would like to thank Prof. dr. Peter Van Nuffelen, dr. Lieve Van Hoof, drs. Renaat Meesters (Ghent University) for reading drafts of this paper and providing us with useful suggestions.

## References

- P. Antónopoulos. *Petros Patrikios: ho Vyzantinos diplōmatēs, axiōmatouchos kai syngrapheas*, volume 7. *Ιστορικές εκδόσεις Στ. Δ. Βασιλόπουλος*, 1990.
- D. Bamman and D. Smith. Extracting two thousand years of latin from a million book library. *Journal on Computing and Cultural Heritage (JOCCH)*, 5(1):2, 2012.
- S. J. B. Barnish. Sacred texts of the secular: Writing, hearing, and reading cassiodorus variae. *Studia Patristica*, 38:362–70, 2001.
- S. J. B. Barnish et al. *The Variae of Magnus Aurelius Cassiodorus Senator...*, volume 12. Liverpool University Press, 1992.
- M. Baroni, G. Dinu, and G. Kruszewski. Dont count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, volume 1, 2014.
- J. Bjerva and R. Praet. Word embeddings pointing the way for late antiquity. In *9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences and Humanities (LaTeCH 2015)*, pages 53–57. Association for Computational Linguistics, 2015.
- S.M. Bjornlie. What have elephants to do with sixth-century politics?: A reappraisal of the" official" governmental dossier of cassiodorus. *Journal of late Antiquity*, 2(1):143–171, 2009.

- S.M. Bjornlie. *Politics and Tradition between Rome, Ravenna and Constantinople: A Study of Cassiodorus and the Variae, 527–554*, volume 89. Cambridge University Press, 2012.
- N. Coffee, J-P. Koenig, S. Poornima, C. W. Forstall, R. Ossewaarde, and S.L. Jacobson. The tesserae project: intertextual analysis of latin poetry. *Literary and Linguistic Computing*, 28(2):221–228, 2013.
- S. Dmitriev. John lydus and his contemporaries on identities and cultures of sixth-century byzantium. *Dumbarton Oaks Papers*, 64:27–42, 2010.
- L. Edmunds. *Intertextuality and the reading of Roman poetry*. JHU Press, 2001.
- J. R. Firth. A synopsis of linguistic theory. pages 1930–1955. 1952–1959:1–32, 1957.
- A.J. Fridh and J.W. Halporn. *Cassiodorus' Variarum libri XII*. Typographi Brepols Editores Pontificii, 1973.
- A.J. Fridh and J.W. Halporn. Magni aurelii cassiodori senatoris opera pars i: Variarum libri xii. *Corpus Christianorum, Series Latina*, 96., 1973.
- T.M.J. Fruchterman and E.M. Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.
- I.C. Ghiban and Ş. Truşan-Matu. Network based analysis of intertextual relations. In *Advances in Information Systems and Technologies*, pages 753–762. Springer, 2013.
- A. Giardina. *Cassiodoro politico e il progetto delle "Variae"*. Centro Italiano di Studi sull'Alto Medioevo, 1993.
- A. Giardina, G.A. Cecconi, I. Tantillo, et al. Cassiodorus' varie. 2015.
- A. Gillett. The purposes of cassiodorus variae. *After Romes Fall: Narrators and Sources of Early Medieval History*, pages 36–50, 1998.
- J.W. Halporn, M. Vessey, et al. *Cassiodorus: Institutions of divine and secular learning and On the soul*, volume 42. Liverpool University Press, 2004.
- P.J. Heather. *The historical culture of Ostrogothic Italy*. na, 1993.
- E. Jouanaud. Pour qui cassiodore at-il publié les variae?. *Teoderico il Grande ei Goti d'Italia*, pages 721–741, 1993.
- R. Koopman, S. Wang, A. Scharnhorst, and G. Englebienne. Ariadne's thread: Interactive navigation in a world of networked information. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1833–1838. ACM, 2015.
- L. Kotlerman, I. Dagan, I. Szpektor, and M. Zhitomirsky-Geffet. Directional distributional similarity for lexical inference. *Natural Language Engineering*, 16(04):359–389, 2010.
- A. Lenci and G. Benotto. Identifying hypernyms in distributional semantic spaces. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 75–79. Association for Computational Linguistics, 2012.
- O. Levy, Y. Goldberg, and I. Dagan. Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association of Computational Linguistics*, pages 211–225, 2015.
- M.D. MacLeod. Lucian vii. *Loeb Classical Library. London and Cambridge, Mass*, 1961.
- K. Małajowicz. A stylometric approach to intertextuality: An analysis of english and non-english authors, 2013. URL <https://sites.google.com/site/computationalstylistics/projects/computer-methods-in-textual-studies-2014/malajowicz>.
- T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013a.
- T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119, 2013b.
- T. Mikolov, W-t. Yih, and G. Zweig. Linguistic regularities in continuous space word representations. In *Proceedings of NAACL-HLT*, pages 746–751, 2013c.
- W.G. Müller. Interfigurality. a study on the interdependence of literary figures. *Intertextuality*, pages 101–121, 1991.
- J.J. O'Donnell. *Cassiodorus*. Los Angeles: University of California Press, 1979.
- J.J. O'Donnell. Liberius the Patrician. *Traditio*, pages 31–72, 1981.
- T.P. Peixoto. The graph-tool python library. *figshare*, 2014. doi: 10.6084/m9.figshare.1164194. URL [http://figshare.com/articles/graph\\_tool/1164194](http://figshare.com/articles/graph_tool/1164194).
- C. Rapp. Literary culture under justinian. *The Cambridge Companion to the Age of Justinian*, Cambridge, pages 376–97, 2005.
- N. Tahmasebi, L. Borin, G. Capannini, D. Dubhashi, P. Exner, M. Forsberg, G. Gossen, F.D. Johansson, R. Johansson, M. Kågeback, et al. Visions and open challenges for a knowledge-based culturomics. *International Journal on Digital Libraries*, 15(2-4):169–187, 2015.
- L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(2579-2605):85, 2008.

- G. Zecchini, A. Giardina, G. Cecconi, I. Tantillo, F. Oppedisano, A. Marcone, E. Lo Cascio, A. LA Rocca, C. La Rocca, V. Neri, et al. *Cassiodoro Varie. Volume 2: Libri III, IV, V*. Erma di Bretschneider, 2014.
- Y. Zhang, A. Jatowt, S. Bhowmick, and K. Tanaka. Omnia mutantur, nihil interit: Connecting past with present by finding corresponding terms across time. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 645–655, Beijing, China, July 2015. Association for Computational Linguistics.