ORIGINAL ARTICLE

Diffusion dynamics of influence in a social network of intellectuals

Raji Ghawi[1] · Cindarella Petz[1] · Jürgen Pfeffer[1]

Abstract

The history of intellectuals consists of a complex web of influences and interconnections of philosophers, scientists, writers, their work, and ideas. To understand how did these influences evolve over time, we mined a network of influence of over 12,500 intellectuals, enriched it with a temporal dimension dividing the history into six eras. We analyze time-sliced projections of the network into within-era, inter-era, and accumulated-era networks, and identify various patterns of intellectuals and eras and studied their development in time. We also construct influence cascades, analyze their properties: size, depth and breadth, and analyze how the cascades of influence evolve over the consecutive eras. We find out that the cascades are clustered into two categories, namely small- and large cascades. An interesting finding here is that the fraction of small cascades increases, while the fraction of larges cascades decreases over time. We also briefly analyze the community structure within the influence network of scholars.

Keywords Social networks · Intellectual influence · Diffusion dynamics

1 Introduction

"No self is of itself alone," wrote Erwin Schrödinger in 1918 (Moore 1994) and noted, "It has a long chain of intellectual ancestors." The history of intellectuals is comprised of a myriad of such long chains, embedded in a tapestry of competing influences of "ageless" ideas, which—in the words of the French scholar Bonaventura D'Argonne in 1699—"embrace [...] the whole world" (Grafton 2009).

To understand the dynamics of influence and spread of ideas through history, the embeddness and interconnections of scholarship should be taken into account. A network approach offers to identify the most influential scholars via their positions in a network of intellectual influence through the history. This allows the study of their social relations (Wasserman and Faust 1994; Hennig et al. 2012; Otte and

Rousseau 2002), and to provide deep insights into the underlying social structure.

In a previous work (Ghawi et al. 2019), we addressed the analysis of such a social network of intellectual influence, incorporating over 12,500 scholars from international origins since the beginning of historiography. In Petz et al. (2020), we also extended the analysis of that network by incorporating a temporal dimension, analyzing the network of scholars dependent to their time, and adding a longitudinal perspective on how scholars formed networks. By doing so, we opted for an inclusive, global perspective on the history of intellectuals. This perspective of a vast longitudinal global network of intellectuals is a response to recent discussions on not-global-enough research within intellectual history (Haakonssen and Whatmore 2017). We thus attempt to go beyond the traditional "master narratives" (Gänger and Lewis 2013) of a Western European centrist view on intellectual history (Subrahmanyam 2017).

The goal is not only to understand how the influence relations among scholars evolved over time, but also to get deep insights on their influence on historical periods. With this kind of longitudinal analysis, we can answer questions like: how did these influence networks evolve over time? who were the most influential scholars in a period? and which patterns of influence did emerge?

✉ Raji Ghawi
  raji.ghawi@tum.de

  Cindarella Petz
  cindarella.petz@tum.de

  Jürgen Pfeffer
  juergen.pfeffer@tum.de

[1] Bavarian School of Public Policy, Technical University of Munich, Munich, Germany

In this paper, we build upon (Petz et al. 2020), and extend the analysis of the social network of scholars by addressing the diffusion dynamics of influence among scholars over the history. As scholars get influenced by other scholars, who are influenced by others, and so on, the influence of scholars spread over time, and takes the form of *cascades*. Influence cascades can be characterized using several properties, such as size, depth, and breadth. For all scholars in the social network, we measure and analyze these properties, and categorize the scholars based on the properties of their influence cascades.

Moreover, we analyze the community structure within the network of scholars. By applying a community detection algorithm, we are able to identify the major communities of scholars who densely influence each other, forming knowledge clusters.

Our contributions are as follows:

– We incorporate a longitudinal perspective on the social network analysis of intellectuals based on a global periodization of history.
– We identify patterns of influence and their distribution in within-, inter-, and accumulated-era influence networks.
– We identify influence signatures of scholars and eras.
– We identify scholars with various knowledge broker roles.
– We construct influence cascades of scholars, and measure and their properties.
– We analyze the cascade properties over eras, and characterize them into two clusters of small- and large cascades.
– We analyze the community structure within the network of scholars, and how the identified communities influence each other.

This paper is organized as follows. Section 2 reviews related works. In Sect. 3, we briefly outline the dataset's characteristics and pre-processing. Section 4 presents the network analysis of the entire network, and its time-sliced projections into partial influence networks (within-era, inter-era, and accumulated-era), featuring their basic network metrics, degree distribution, and connectivity. In Sect. 5, we identify different influence patterns of scholars and eras. Section 6 is devoted to the longitudinal analysis of brokerage roles of scholars. In Sect. 7, we address the diffusion dynamics of influence through the analysis of influence cascades of scholars over different eras. Finally, in Sect. 8, we address the communities of scholars, and the influence between them.

2 Related work

The term of intellectual history combines a plethora of approaches on discourse analysis, evolution of ideas, intellectual genealogies, and the history of books, various scientific disciplines, political thought, and intellectual social context (Wickberg 2001; Gordon 2013). These studies are usually limited to specific regions or time spans as a trade-off for thorough comparative and textual analysis. Endeavors to write a "Global Intellectual History" (Moyn and Sartori 2013) were criticized for focusing on the more well-known intellectual thinkers despite including a transnational comparative perspective (Subrahmanyam 2015).

Network methodologies allow analyzing intellectual history and as such the history of intellectuals as big data, encompassing time and space with a focus on their inter-connections. Notably, computational methods have been used in the study of communication networks of the *respublica litteraria* of the late 17th and 18th century, in which various studies modeled the Early Modern scholarly book and letter exchanges as formal networks. Since 2008, the project "Mapping the Republic of Letters" at Stanford University spearheaded the digitization of Early Modern letters and systematically modelled the metadata on who is connected to whom in correspondence networks mapped into spatial realm, similarly to a "traffic analysis" (Edelstein et al. 2017 p. 403). More recent studies have incorporated a temporal perspective on these epistolary networks, and studied their change in time, as well as differentiated between the types of correspondence exchanged in a multi-layered perspective (Vugt 2017). While the *Republic of Letters* contains a multitude of scholarly actors in an imagined intellectual community—the so-called republic—consisting of "a palimpsest of people, books, and objects in motion" (Grafton 2009 p. 6), it is confined to the Early Modern period, and primarily studied through the in-depth analysis of selected ego networks, such as e.g. the correspondence network of Benjamin Frankling during his "London Decades" (1757–1775) Winterer (2012), or the influence of English authors on the Enlightenment philosophy of Voltaire (Edelstein and Kassabova 2020).

A recent study (Ghawi et al. 2019) proposed to research the entire history of intellectuals with the means of a network approach. This paper defined the most influential as those with the longest reaching influence (influence cascades), and identified as such Antique and Medieval Islam scholars, and Karl Marx as the one with the most out-going influences. In this paper, we extend this analysis by incorporating a temporal dimension in order to establish a deeper insight on how these influences evolved in time. Much research has been devoted to the area of longitudinal social networks (Newcomb 1961; Huisman and Snijders 2003; Snijders et al. 2010; Holme and Saramäki 2019). Longitudinal network studies aim at understanding how social structures develop or change over time, usually by employing panel data (Hennig et al. 2012). Snapshots of the social network at different points in time are analyzed in order to explain the changes in the social structure between two (or more)

points in time in terms of the characteristics of the scholars, their positions in the network, or their former interactions.

3 Data

3.1 Data acquisition and preprocessing

The source of information used in this paper originated from YAGO (Yet Another Great Ontology) (Mahdisoltani et al. 2015), which is a pioneering semantic knowledge base that links open data on people, cities, countries, and organizations from Wikipedia, WordNet, and GeoNames. At YAGO, an influence relation appears in terms of the `influences` predicate that relates a scholar to another when the latter is influenced by the ideas, thoughts, or works of the former. The accuracy of this relation was evaluated by YAGO at 95%. We extracted a dataset that encompasses all influence relationships available in YAGO, using appropriate SPARQL queries that implement mining techniques of social networks from Linked Open Data (Ghawi and Pfeffer 2019). The result consisted of 22,818 directed links among 12,705 intellectuals that made up the nodes and edges of our target social network of influence. In order to incorporate a time dimension to our analysis, we extracted birth and death dates of each scholar. Some scholars had missing birth and/ or death dates, which we deduced by subtracting 60 years from the death date, and vice versa, up to the symbolic year of 2020. When both dates were missing, we manually verified them. During this process we had to remove some entities, as they did not correspond to intellectuals. These were either (1) concepts, e.g., 'German_philosophy' and 'Megarian_school', (2) legendary characters, e.g., 'Gilgamesh' and 'Scheherazade', or (3) bands e.g., 'Rancid' and 'Tube.' To this end, we obtained a new dataset of 12,577 scholars with complete birth and death dates.

3.2 Periodization

In this paper, we do not use the classical concept of network snapshot, which is a static network depicted at a given point in time. Rather, we split the time span (i.e., the history) manually into consecutive periods (eras), and embed the network nodes (actors) into the eras in which they lived. This way, the micro-level influence among scholars can be viewed as a macro-level influence among periods of history. This enables the analysis of the influence network within each era (= within-era), between different eras (= inter-era), and in an accumulative manner (= accumulated-era). By introducing a longitudinal perspective, we split the time-span using a periodization that takes global events into account. Any periodization is a construct of analysis, as each field of research has its own timeline characterizing periods (Pot 1999) which

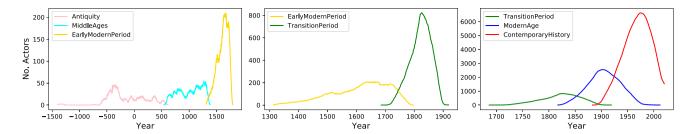**Table 1** Eras and their start- and end dates

	Era	Start–end years	Length
AN	Antiquity	(up to 600 AD)	∼ 2000
MA	Middle ages	(600, 1350)	750
EM	Early modern period	(1350, 1760)	410
TP	Transition period	(1760, 1870)	110
MR	Modern age	(1870, 1945)	75
CH	Contemporary history	(1945, 2020)	75

are dependent on different caesura for the respective object of research (Osterhammel 2002). This complicates an overarching longitudinal perspective on a global scale. In order to match the internationality of scholars, we used Osterhammel's global periodization (Osterhammel 2002) and worked with six consecutive periods (eras): Antiquity (up to 600 AD), Middle Ages (600–1350), Early Modern Period (1350–1760), Transitioning Period (1760–1870), Modern Age (1870–1945), and Contemporary History (1945–2020). We have given each of these eras an abbreviation to easily referring it throughout the paper as shown in Table 1.

One conceptual challenge was to map scholars into eras. Many scholars fit to more than one period's timeline. We opted for a single era membership approach since it is more intuitive and easier to conceptualize. A single era membership of each scholar reduces the complexity of analysis and computations, while encompassing the essential membership of each scholar to a single era. It also offers adequate results when we compare eras, since it avoids redundancy. This approach does not change the influences of the scholar to scholars of other periods.

In order to assign a single era to a scholar, we used the following method: We calculated the midpoint of the scholar's lifespan ignoring the first 20 years of their age, as we assumed that scholars in general would not be active then. Then we assigned the era in which this midpoint occurs as the scholar's membership to an era. After this initial assignment process, we verified the global validity of assignments by counting the number of influence links from one era to another. We observed that there were some reverse links of eras, i.e., an influence relation from an actor in a recent era toward an actor assigned to an older era. Those anomaly cases (about 200) were basically due to:

1. Errors in dates: Some dates were stated in the Hijri calendar, instead of the Gregorian calendar, and some dates were BC and missing the negative sign.
2. Errors in direction of the relationship: Source and target actors were wrongly switched.
3. Inappropriate era-actor assignments.

Fig. 1 Number of scholars alive in each year based on their assigned eras

Fig. 2 Number of scholars per era

Fig. 3 Number of influence links from preceding periods (rows) to target eras (columns)
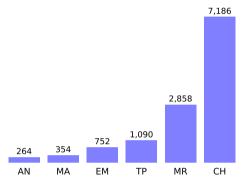
The anomalies due to errors have been manually corrected. The cases of inappropriate assignment were technically not erroneous. This usually happened when the influencer lived much longer than the influenced, elevating the influencer's period into a more recent one. We solved this by iteratively reassigning either the influencer backward to the era of the influenced, or the influenced forward to the era of the influencer. As a result, each scholar is assigned to exactly one era, such that no reverse links of eras exist. The final cleaned dataset consists of 22,485 influence links among 12,506 intellectuals. Figure 1 shows each era's continuous density of scholars based on their lifespan; whereas Fig. 2 shows the number of scholars assigned to each era.

4 Analysis

With scholars embedded in their respective eras, the entire influence network can be time-sliced: we projected it into several partial networks based on the source era (of the influencer) and target era (of the influenced scholar). When the source and target eras are the same, we call the partial network a *within-era* influence network. When the source and target eras are different, we call the partial network an *inter-era* influence network. There are no reverse links from a later era to a previous one due to preprocessing.

As a result of time-slicing the whole network, we obtain six within-era networks corresponding to all the six eras, and 15 inter-era networks, corresponding to all chronologically ordered (but not necessarily consecutive) pairs of different eras. Moreover, we constructed six *accumulated-era* influence networks of scholars living up to and including a target era.

Figure 3 shows the distribution of influence links over pairs of eras, where the rows represent source eras, and the columns represent target eras, i.e., each cell displays the number of influence links incoming from (actors in) the row era, outgoing to (actors in) the columns era. One can easily observe that the greatest deal of links occur within the Contemporary era, followed by the links from Modern Age to Contemporary, and within Modern Age. This is obviously because those recent eras comprise the largest deal of scholars in our dataset.

Figure 4 shows the proportion of influence links among all pairs of eras. There, we can already make two major observations for inter- and within-era influence relations: For one, the highest fraction of influence received by scholars of each era comes from its own era. This means that the internal impact of any era is in general higher than its external impact. In absolute numbers, the vast majority of links occur within the Contemporary era, followed by links from the Modern Age to the Contemporary period, and within the
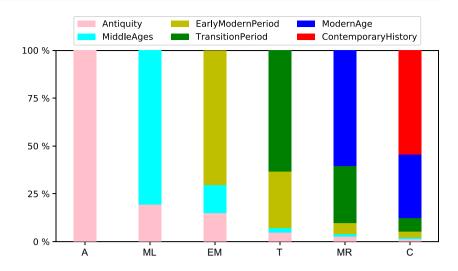
Fig. 4 Percentage of received influences in each era

Table 2 Metrics of within-era networks

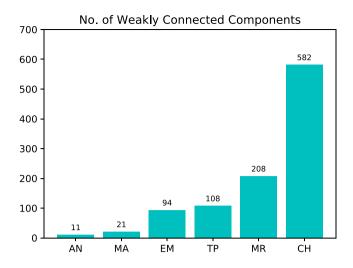| Era	AN	MA	EM	TP	MR	CH
N	219	303	610	761	2102	6081
N/A	82%	86%	81%	70%	73%	85%
E	327	387	694	927	2806	7960
Density	.0068	.0042	.0019	.0016	.0006	.0002
avg. d_{out}	1.49	1.28	1.14	1.22	1.33	1.31
max d_{in}	12	9	17	27	21	26
max d_{out}	20	16	23	32	68	58
max d	32	20	32	41	73	58
WCC	11	21	94	108	208	582
LWCC	179	233	245	436	1495	4379
	82%	77%	40%	57%	71%	72%
SCC	0	2	6	8	31	38
R	0	0.005	0.023	0.028	0.036	0.014
T	0.064	0.066	0.071	0.042	0.029	0.017

Modern Age, which is clearly owed to the increased amount of scholars in these periods.

The inter-era influences of each period is strongest on its consecutive period. As our earliest period, Antiquity receives only influence links from itself, whereas the influence received in the Middle Ages are 82% internal, and 18% from Antiquity. Subsequently, the amount of the within-era influence shrinks throughout the consecutive periods, but still remains the biggest influence. Noteworthy here is the high proportion of influences of Antiquity on the Early Modern period, which represents their increased reception during the Renaissance. However, the proportionately many links of Antiquity to the Middle Ages reassert the shift in historical research that the Renaissance did not "rediscover" Antiquity, but was received before in the Middle Ages as well (Fejfer et al. 2003 p. 3–4).

4.1 Within-eras influence networks

In the following, we analyzed the six *within-era* influence networks, which represent the internal impact of an era. We extracted the following metrics, as shown in Table 2:

– Number of nodes N, and edges E, and density D.
– Average out-degree (= avg. in-degree due to the properties of a directed graph).
– Max. in-degree, max. out-degree, and max. degree.
– WCC: number of weakly connected components.
– LWCC: size of the largest weakly connected component.
– SCC: number of strongly connected components, when the number of nodes is > 1).
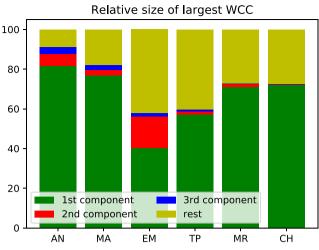– Reciprocity (R) and transitivity (T).

Fig. 5 Weakly connected components in within-era influence networks

We included $\frac{N}{A}$ in Table 2 in order to contain that the number of nodes N in a within-era network could be less than the number of actors of that of era A. This is owing to the fact that not all scholars of an era necessarily participated in its within-era influence network. Some scholars influenced or were influenced by actors of different eras only. However, around 80% of scholars in each era were active in these within-era networks. The highest value of 86% of the Middle Ages refers to their relative self-containment as an era, as well as the lowest value in the Transitioning period of 70% refers to its high out-going influences.

Over all eras, the amount of nodes and edges steadily increased, while the density of networks decreased. On average, the out-degree revolves around 1.25, where the highest value of 1.5 occurs in Antiquity, and the lowest of 1.14 in the Early Modern period. When we compare the evolution of the max. out-degree in time, we find that the expected continuous increase did not always hold due to two exceptionally high observations at Antiquity and the Modern Age. Mutual ties among contemporaries were in general very low. We can report none in Antiquity, and only one in the Middle Ages between Avicenna and Al-Bīūī. In the Early Modern period, eight mutual relations were observed, including, e.g., Gottfried Leibniz (1646–1716) and David Bernoulli (1700–1782), whereas 13 mutual relations in the Transitioning period, such as Friedrich Engels (1820–1895) and Karl Marx (1818–1883), or Johann Goethe (1749–1832) and Friedrich Schelling (1775–1854). In the Modern Age, the number of mutual ties increased to 51 (e.g., Jean-Paul Sartre (1905–1980) and Simone de Beauvoir (1908–1986)); and to 54 in the Contemporary period.

Figure 5 shows the number of weakly connected components (WCCs) in the within-era networks of each era, and the relative size of the largest ones w.r.t the whole corresponding network.

The number of WCCs increased gradually over the consecutive eras. In general, the networks consisted of one giant component, which encompassed the majority of nodes, while the rest of components were relatively smaller. This was particularly developed in Antiquity and the Middle Ages, where the giant components constitute of 82% and 77% of the nodes, while the second largest were at 6% and 3%, respectively. The Early Modern period constitutes an exception to this giant component rule: the largest one was only at 40%, and the second largest at 16%. Looking at their composition, the first consisted of natural scientists, mathematicians, and philosophers, such as Descartes, Newton, and Leibniz, while the smaller one was compromise of artists and painters, such as Rembrandt and Raphael. The single giant component phenomenon appeared again in subsequent eras. For instance, in the Transitioning period, there were 108 WCCs, where the largest two incorporated 57% and 1.3% of the nodes. In the Modern and Contemporary Age, the largest components comprised about 70% of nodes.

Who was most influential on their contemporaries? Table 3 lists the top five scholars per era based on their out-degree in the within-era influence networks. The highest within-era out-degree over all times was achieved by Friedrich Nietzsche (1844–1900) of the Modern Age with 68 outgoing influence links to other scholars of his era.

4.2 Inter-era influence networks

Inter-era influence networks are partial networks where the source era precedes the target era. We interpreted these networks as bipartite, as the actors belong to different groups; the source era and the target era. Therefore, only edges between nodes sets are possible.

Table 3 Top 5 actors, per era, based on out-degree in within-era influence networks

Antiquity		MiddleAges	
Plato	20	Avicenna	16
Aesop	13	Muhammad	11
Pythagoras	10	Al-Ghazali	11
Plotinus	10	BanūMūsā	8
Euhemerus	10	J. S. Eriugena	8
EarlyModern		Transition	
John Locke	23	Goethe	32
René Descartes	22	Hegel	29
Isaac Newton	15	Lord Byron	24
Hugo Grotius	13	Immanuel Kant	22
Leibniz	11	von Schelling	17
Modern		Contemporary	
Nietzsche	68	Vladimir Nabokov	58
Jules Verne	35	Friedrich Hayek	50
Henri Bergson	35	Richard Pryor	50
Leo Tolstoy	24	Jacques Derrida	48
Edmund Husserl	22	Michel Foucault	47

Table 4 shows the metrics for those inter-era influence networks. In general, each era had the most links with its consecutive era, and additionally with the Contemporary period's scholars. Exception to this was Antiquity, which saw its first peak with the Early Modern period relating to Renaissance interests.

Their densities were again decreasing through the combinations, except for those periods that had less links to other periods, such as the Middle Ages to the Transitioning period.

Which scholar influenced a successive era the most? Table 5 shows the scholars with the highest degrees in the inter-era networks. Noteworthy here is Karl Marx, who had the highest out-degree over all times from the Transitioning period to the Contemporary age, followed by modern philosopher Friedrich Nietzsche and Martin Heidegger on Contemporary scholars.

4.3 Accumulative influence networks

For each era, we constructed an accumulative network of all influence links among scholars who lived up to and including that era. We performed essential social network analysis on these six *accumulated-eras* networks, which combine the internal and external impact of eras. The final network of the Contemporary Age is the same as the complete network over all periods (Ghawi et al. 2019).

Figure 6 shows the best connected scholars for each era those that influenced at least 10 others - in the final accumulated network. We clearly see two joined networks of hubs. The right section is very diverse in terms of including different eras and different fields such as philosophy, theology, and science scholars. The left section consists mainly of writers since the Long 19th Century (1789–1914); Alexander Pushkin (1799–1837) is one of the eldest nodes there. This writers' network shows little diversity in comparison to other historical periods and consists mostly of Modern and Contemporary age writers. That writers are less connected to the philosophy, theology, and science scholars show that these groups referenced themselves more consistently.

Table 6 shows the metrics of accumulated-era networks. Regarding node degrees change over consecutively accumulated eras, we observe that at all eras the maximum

Table 4 Metrics of inter-eras influence networks

Source → Target	N	E	N_s	N_t	D	In-degree avg	In-degree max	Out degree avg	Out degree max
AN → MA	82	87	38	44	.052	1.98	7	2.29	12
AN → EM	117	145	46	71	.044	2.04	7	3.15	19
AN → TP	66	66	29	37	.062	1.78	5	2.28	11
AN → MR	101	114	42	59	.046	1.93	11	2.71	23
AN → CH	169	177	49	120	.030	1.47	6	3.61	46
MA → EM	149	144	66	83	.026	1.73	9	2.18	21
MA → TP	52	36	22	30	.055	1.20	5	1.64	6
MA → MR	77	62	27	50	.046	1.24	4	2.30	12
MA → CH	146	121	50	96	.025	1.26	6	2.42	34
EM → TP	392	432	159	233	.012	1.85	16	2.72	24
EM → MR	262	269	101	161	.016	1.67	13	2.66	15
EM → CH	437	432	125	312	.011	1.38	7	3.46	35
TP → MR	1111	1373	436	675	.005	2.03	19	3.15	53
TP → CH	888	1041	212	676	.007	1.54	9	4.91	112
MR → CH	3817	4885	1271	2546	.002	1.92	17	3.84	78

Table 5 Top scholars w.r.t out-degree in inter-era networks

source → target	First rank		Second Rank	
AN → MA	Aristotle	12	Augustine of Hippo	6
AN → EM	Aristotle	19	Plato	14
AN → TP	Aristotle	11	Plato	9
AN → MR	Plato	23	Aristotle	16
AN → CH	Aristotle	46	Plato	32
MA → EM	Ibn Tufail	21	Thomas Aquinas	9
MA → TP	Petrarch	6	Dante Alighieri	5
MA → MR	Dante Alighieri	12	Thomas Aquinas	11
MA → CH	Thomas Aquinas	34	Dante Alighieri	10
EM → TP	J. J. Rousseau	24	Shakespeare	21
EM → MR	Baruch Spinoza	15	Shakespeare	15
EM → CH	Shakespeare	35	David Hume	25
TP → MR	Immanuel Kant	53	Karl Marx	43
TP → CH	Karl Marx	112	Hegel	67
MR → CH	Nietzsche	78	Martin Heidegger	73

out-degree is greater than the maximum in-degree. Moreover, those maximum degrees continuously increase over eras, in contrast to within-era networks. The average out-degree

changes slightly over time, taking its lowest value of 1.45 at Middle Ages, and highest value of 1.8 at Contemporary age. Noteworthy is the drastic collapse of the largest Weak Component in the Early Modern period, which has steadily risen since.

Who was the most influential intellectual in an era? Figure 7 shows the evolution of the 10 most influential scholars in the complete network based on their out-degree progression in the accumulative networks.

The top two ranks of the most prolific scholars were consistently taken over by Antique philosophers Plato, and Aristotle (who among contemporaries was only in rank 6). Contemporary scholars came on third rank in the Middle Ages (Avicenna), in the Early Modern period (Ibn Tufail, John Locke, René Descartes), and in the Transitioning period (John Locke, Johann Goethe). This changed in the Modern Age, when Transitioning period scholars Immanuel Kant and Hegel took the first ranks. Aristotle still remained in the top five. The highest out-degree over all times is observed at the Contemporary Age, where Karl Marx had 158 outgoing influence links to other scholars of all eras, followed by Nietzsche, Hegel, and Kant.
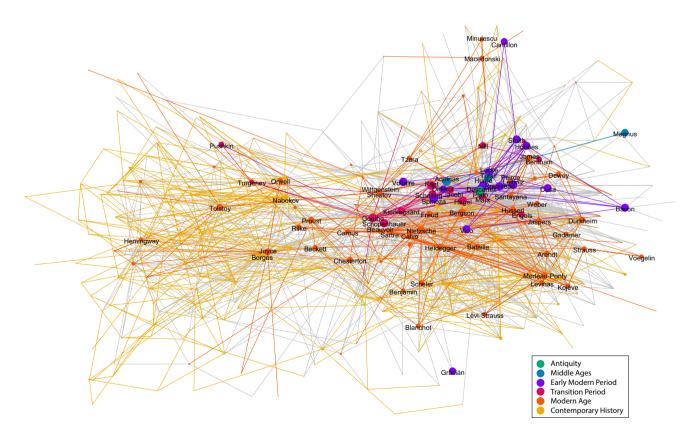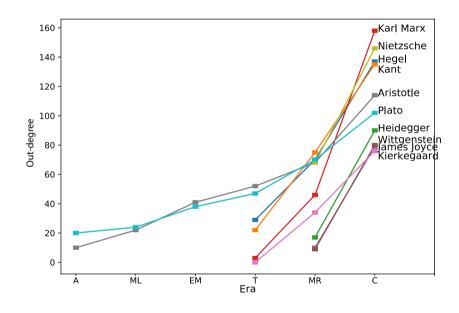

Fig. 6 Network of the most influential actors with at least 10 out-going influences. Node size = proximity prestige, node color = era, links within an era are colored with the color of the era, the other links are gray

Table 6 Metrics of accumulative-era networks

Era	AN	MA	EM	TP	MR	CH
N	219	552	1227	2141	4697	12,506
E	327	801	1784	3245	7869	22,485
N_{source}	54	155	388	677	1501	3890
N_{inner}	71	178	353	597	1331	3080
N_{sink}	94	219	486	867	1865	5536
Density	.0068	.0026	.0012	.0007	.0004	.0001
avg. d_{out}	1.49	1.45	1.45	1.5	1.68	1.80
max d_{in}	12	16	26	38	48	48
max d_{out}	20	24	41	52	75	158
max d	32	36	50	60	116	196
WCC	11	30	110	211	390	817
LWCC	179	441	797	1513	3550	10192
	82%	80%	65%	71%	76%	81%
SCC	0	2	8	16	47	85
R	0	0.002	0.010	0.014	0.019	0.011
T	0.064	0.067	0.064	0.056	0.039	0.021

Fig. 7 Top 10 of the most influential intellectuals of the complete network based on their out-degree, and their progression in the accumulated-era networks

5 Patterns of influence over eras

In this section, we study the influence patterns of scholars over eras. We construct influence signatures based on how much on average a scholar influenced an era, and which patterns of directed influences characterize an era.

5.1 Influence power of scholars

For each scholar, we construct their influence signature as a sequence of their influence links toward each era, starting from their own. For example, the influence signature of Aristotle was [10, 12, 19, 11, 16, 46], which meant he had 10 influence links within Antiquity, 12 links toward the Middle Ages, etc. Using those signatures, we define the *longitudinal influence power* of a scholar as the average of their influence signature. A scholar would have a high influence power when he has (1) a high number of influence links, (2) over all or many eras. In contrast, having few influence links over several eras, or many links over few eras would give a low value of this influence power measure. For example, with an average around 19 both Aristotle and Shakespeare had similar influence powers. In absolute numbers, Aristotle had almost twice the number of Shakespeare's influence links (114 to 73, respectively). While Aristotle influenced all 6 eras, and Shakespeare only 4, the ratio of the links per era decreased for Aristotle, resulting in their similar influence

Table 7 Top scholars, in each era, with respect to *longitudinal influence power*

Antiquity		MiddleAges	
Aristotle	19.0	Thomas Aquinas	12.6
Plato	17.0	Dante Alighieri	6.0
Augustine of Hippo	6.0	Ibn Tufail	5.8
Plotinus	4.7	Avicenna	4.6
Heraclitus	4.2	Al-Ghazali	3.6
EarlyModernPeriod		**TransitionPeriod**	
William Shakespeare	18.2	Karl Marx	52.6
Baruch Spinoza	14.8	Hegel	45.7
René Descartes	14.0	Immanuel Kant	45.0
John Locke	13.0	Søren Kierkegaard	25.3
David Hume	12.5	Fyodor Dostoyevsky	23.0
ModernAge		**ContemporaryHistory**	
Friedrich Nietzsche	73.0	Vladimir Nabokov	58.0
Martin Heidegger	45.0	Friedrich Hayek	50.0
Ludwig Wittgenstein	40.0	Richard Pryor	50.0
James Joyce	39.5	Jacques Derrida	48.0
Sigmund Freud	32.0	Michel Foucault	47.0

powers. This measure provides an indicator of the influence power of an intellectual throughout history, and combines both the intensity and the diversity of influence.

Influence power also allows us to compare scholars from different eras. Table 7 shows the top 5 scholars based on the longitudinal influence power. Here, Aristotle, Thomas Aquinas, William Shakespeare, Karl Marx, Friedrich Nietzsche, and the writer Vladimir Nabokov (1899–1977) are identified by their influence power as the most influential intellectuals of their respective periods. The highest longitudinal influence powers over all times had Nietzsche (73), followed by Nabokov (58) and Marx (52).

5.2 Influence patterns

Which directed influences were most common in an era? We derive these influence patterns of eras by replacing any nonzero entries by X of the scholar's influence signatures, and aggregate all occurrences of each pattern for each era. We thus ignore the actual values of influence (intensity), but keep the temporal effect (diversity). For example, the influence pattern $[X, 0, \cdots, 0]$ means that the scholarly influences goes to the first (own) era only, with no influence on other eras. The pattern $[X, X, \cdots, X]$ signifies that the influence is distributed over all applicable eras, regardless of the actual values. Table 8 gives the top patterns of each era with the pattern's frequency of occurrence with regard to the respective era.

For example, for the Middle Ages the most frequent pattern is $[-, X, 0, 0, 0, 0]$, which represents that 56% of scholars only influenced contemporaries with no influences on other eras. Over all eras, the most common pattern was within-era influence, followed by the influence on the consecutive period. Exception to this rule is the Modern period, which experienced the reverse, and had a higher influence on the Contemporary period than on its own. Since the Early Modern period, the pattern of influencing all successive eras including its own becomes more frequent (from 7% on), and rises with each successive period.

6 Brokerage roles

Which roles had scholars in regard to their influence on others? By following the brokerage approach by Gould and Fernandez (Gould and Fernandez 1989), we infer on the roles of scholars by analyzing the non-transitive triads, in which node A has a tie to node B, and B has a tie to node C, but there is no tie between A and C. In these triads, B is thought to play a structural role called a *broker*.

The possible brokerage roles are shown in Fig. 8, which are adapted from the work of Gould and Fernandez in Gould and Fernandez (1989), and Everett and Borgatti (2012). These brokerage roles are:[1]

1. *Coordinator*, where A, B and C all belong to the same group;
2. *Representative*, where A and B belong to one group, and C belongs to another;
3. *Gatekeeper*, where A belongs to one group, and B and C belong to another;
4. *Liaison*, where A, B and C each belong to a different group.

In this paper, we interpret nodal membership in groups as eras. This allows us to consider to what extent a node's importance is based on joining two nodes that are members of the node's own era, or on joining others outside their era.

Table 9 shows, for every type of brokerage roles, and for each era, the top three scholars with that role in that era. The number besides each scholar is the number of non-transitive triads of that scholar w.r.t the specified brokerage role and the specified era. Since reverse links, from an era to an older one, are not allowed (as per preprocessing), some brokerage roles are not possible in some eras. Namely, Representative

---

[1] There is a fifth brokerage role, the *Consultant*, where A and C belong to one period, and B belongs to another. This role is not possible in our network, as we did not allow reverse influences of a more recent period onto a previous one.

Table 8 Top frequent influence patterns of eras

	AN	MA	EM	TP	MR	CH	
Antiquity	×	0	0	0	0	0	43%
	0	0	0	0	0	×	8%
	0	×	0	0	0	0	7%
	0	0	×	0	0	0	7%
MiddleAges	×	0	0	0	0	0	56%
	0	×	0	0	0	0	9%
	×	×	0	0	0	0	7%
	0	0	0	0	0	×	6%
EarlyModern			×	0	0	0	51%
			0	×	0	0	13%
			0	0	0	×	7%
			×	×	×	×	7%
Transition				×	0	0	35%
				0	×	0	29%
				×	×	×	11%
				×	×	0	9%
				0	0	×	8%
				0	×	×	7%
ModernAge					0	×	38.8%
					×	0	36.7%
					×	×	24.5%
Contemporary						×	100%

Fig. 8 Brokerage Roles of the top right node of each triad, adapted from Gould and Fernandez (1989) (Gould and Fernandez 1989)

and Liaison brokerage roles are impossible for Contemporary, as well as Liaison and Gatekeeper brokerage roles for Antiquity.

For Coordinator role, A, B and C belong to the same era. Hence, a scholar with this role gets influence from- and influences other scholars from the same era. The scholars with the highest scores for Coordinator in their respective periods are: the ancient Greek philosopher Plato, the medieval polymath Avicenna (980–1037), the Early Modern philosopher John Locke (1632–1704), Johann Goethe (1749–1832), Friedrich Nietzsche (1844–1900), and the contemporary horror writer Stephen King (born 1947).

For Representative role, A and B belong to one era, and C belongs to another (more recent) era. Hence, a scholar with this role gets influence from other scholars from his own era, and influences other scholars from another era. The top scholars with this role are: Plato and Aristotle in Antiquity, Ibn Tufail (1105–1185) and Tomas Aquinas (1225–1274) in Middle Ages, David Hume (1711–1776) and

Leibniz (1646–1716) in Early Modern period, Karl Marx (1818–1883) and Hegel (1770–1831) in Transition period, and the modern philosophers Martin Heidegger (1889–1976) and Ludwig Wittgenstein (1889–1951).

For Gatekeeper role, A belongs to one era, and B and C belong to another more recent era. Hence, a scholar with this role gets influence from other scholars from an older era, and influences other scholars from his own era. The top scholars with this role are: Avicenna and Tomas Aquinas in Middle Ages, René Descartes (1596–1650) and John Locke in Early Modern period, Immanuel Kant (1724–1804), Hegel, and Goethe in Transition period, Nietzsche in Modern Age, and the contemporary French philosopher Michel Foucault (1926–1984).

For Liaison role, A, B and C each belong to a different group. Hence, a scholar with this role gets influence from other scholars from an older era, and influences other scholars from another more recent era. The top scholars with this role are: Tomas Aquinas in Middle Ages, the Early Modern

Era	Coordinator		Representative		Gatekeeper		Liaison	
AN	Plato	207	Plato	970				
	Aristotle	71	Aristotle	770				
	Zeno of Citium	47	Augustine of Hippo	137				
MA	Avicenna	61	Thomas Aquinas	503	Avicenna	40	Thomas Aquinas	351
	Thomas Aquinas	50	Ibn Tufail	203	Pseudo-Denys	37	Dante Alighieri	88
	Al-Ghazali	33	Rumi	85	Thomas Aquinas	37	Meister Eckhart	43
EM	John Locke	194	David Hume	686	René Descartes	221	Baruch Spinoza	559
	Leibniz	171	Leibniz	603	John Locke	125	Shakespeare	353
	René Descartes	146	J.J. Rousseau	461	Leibniz	91	René Descartes	351
TP	Goethe	262	Karl Marx	4,094	Immanuel Kant	467	Immanuel Kant	2,496
	Hegel	247	Hegel	901	Hegel	416	Karl Marx	1,666
	Alexander Pushkin	124	Søren Kierkegaard	715	Goethe	361	Hegel	1,532
MR	Nietzsche	339	M. Heidegger	971	Nietzsche	2,808	Nietzsche	3,242
	A. Macedonski	255	L. Wittgenstein	682	A. Macedonski	386	Martin Heidegger	1,243
	Bertrand Russell	223	J.P. Sartre	597	Henri Bergson	365	Samuel Beckett	770
CH	Stephen King	480			Michel Foucault	1,085		
	Gilles Deleuze	400			Jacques Derrida	1,084		
	Michel Foucault	395			Friedrich Hayek	1,007		

Fig. 9 Brokerage Roles

philosopher Baruch Spinoza (1632–1677), Immanuel Kant in Transition period, and Nietzsche in Modern Age.

7 Diffusion dynamics of influence

In order to get insight on how the influence spread throughout the network, and how this spread change over time, we study the diffusion of influence throughout the network, similarly to Ghawi et al. (2019).

We refer to the influence path formed as scholars, influenced by an original scholar, influence other scholars, as a *cascade*; and we refer to the original scholar as the *root* (see Fig. 10). For each scholar, we construct his influence cascade, by considering out-going edges starting from that scholar. However, in order to avoid exhaustive search (due to cyclicity), we construct scholar's cascade as a directed acyclic graph (DAG), i.e., in case of reciprocal edges between a pair of nodes, we arbitrarily choose one of the reciprocal edges. Thus, the result we obtain is a directed acyclic graph (DAG), which we call henceforth a *cascade*.

In order to characterize the influence cascades, we employ the following features as used in Mathew et al. (2018); Vosoughi et al. (2018):

– **Size**: the number of nodes in the DAG which are reachable from the root node, i.e., the total number of unique nodes involved in the cascade.
– **Depth**: the length of the longest path from the root node of the cascade. The depth of a cascade, *D*, with *n* nodes is defined as

$$D = \max(d_i), 0 \le i \le n$$

where d_i is the distance (length of the shortest path) from the root to node *i*.

– **Average depth**: the average path length of all nodes reachable from the root node.
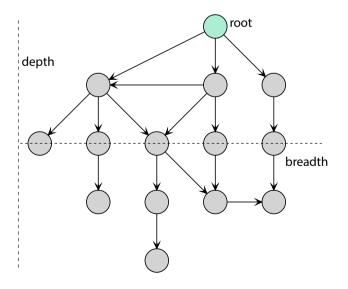
$$AD = \frac{1}{n-1} \sum_{i=1}^{n} d_i$$

Fig. 10 Influence cascade

Era	Size		Depth		Avg. Depth		Breadth	
Antiquity	Thales	5,725	Moses	17	Moses	9.4	Aesop	2,167
	Anaximander	5,724	Aaron	17	Aaron	9.4	Philo the Dialectician	2,163
	Hesiod	5,723	John the Evangelist	15	John the Evangelist	8.0	Pasicles of Thebes	2,163
	Pherecydes of Syros	5,722	Clement of Alexandria	15	Speusippus	7.8	Diodorus Cronus	2,163
	Pythagoras	5,721	Athanasius of Alexandria	15	Pasicles of Thebes	7.4	Thrasymachus of Corinth	2,163
Middle Ages	Ja'far al-Sadiq	5,478	Abū Ḥanīfa	20	Abū Ḥanīfa	12.1	Al-Ghazali	2,125
	Harbi al-Himyari	5,434	Abu Suhail an-Nafi	20	Abu Suhail an-Nafi	12.1	Al-Juwayni	2,125
	Jābir ibn Hayyān	5,433	Hisham ibn Urwah	20	Hisham ibn Urwah	12.1	Abu Talib al-Makki	2,125
	Muhammad	5,402	Malik ibn Anas	19	Al-Juzajani	11.2	Harith al-Muhasibi	2,125
	Al-Kindi	5,401	Al-Shaybani	19	Ibn Jurayj	11.2	Al-Balkhi	2,104
Early Modern	Nicholas of Cusa	5,203	William Derham	17	Jan van Scorel	9.2	Samuel Richardson	2,126
	Michel de Montaigne	5,180	John Ray	17	William Derham	8.5	Denis Diderot	2,118
	Francisco Suárez	5,175	Giulio Cesare Croce	17	John Ray	8.5	G-F. Rouelle	2,118
	Gemistus Pletho	5,172	N. Boileau-Despréaux	16	Giulio Cesare Croce	8.5	Pierre Bayle	2,114
	Basilios Bessarion	5,170	Thomas Gray	15	M. van Heemskerck	8.2	John Selden	1,965
Transition	J. N. Tetens	4,699	H. T. Buckle	18	F. K. Forberg	9.9	Oliver Goldsmith	1,960
	J. H. Lambert	4,699	J-F. Millet	18	J-F. Millet	9.3	Goethe	1,895
	Immanuel Kant	4,698	Monticelli	17	H. T. Buckle	9.1	James Macpherson	1,866
	Thomas Reid	4,644	F. K. Forberg	17	Bettina von Arnim	8.9	F. G. Klopstock	1,844
	J. G. Hamann	4,641	Costache Caragiale	16	A. P. de Candolle	8.6	Friedrich Schiller	1,838
Modern Age	Hippolyte Taine	3,675	Anton Mauve	17	Alfred Binet	9.1	Hippolyte Taine	1,357
	African Spir	3,536	Edmund John	17	James Mark Baldwin	9.1	African Spir	1,283
	Jean-Marie Guyau	3,256	Thomas Mann	17	Jacob L. Moreno	9.1	Herbert Spencer	1,071
	Paul Bourget	3,253	Georges Seurat	17	E. E. Cummings	9.0	Nietzsche	1,008
	Paul Rée	3,253	Vasile Pogor	17	Hans Vaihinger	8.9	Paul Bourget	1,008
Contemporary	R. E. Schultes	884	Witold Gombrowicz	15	Kenneth Patchen	8.0	Jean Genet	214
	Witold Gombrowicz	857	Kenneth Patchen	15	Witold Gombrowicz	7.9	Henry Miller	193
	Kenneth Patchen	857	Thomas Bernhard	14	Robert Duncan	7.2	R. E. Schultes	155
	Robert Duncan	856	Allen Ginsberg	14	Allen Ginsberg	7.1	J-T. Desanti	154
	Thomas Bernhard	853	Robert Duncan	14	Thomas Bernhard	6.9	J. L. Austin	148

Fig. 11 Analysis of influence diffusion cascades of scholars. Table shows the top 5 scholars of each era

– **Breadth**: the maximum number of nodes present at any particular depth in the cascade.

$$B = \max(b_j), 0 \le j \le D$$

where b_j denotes the breadth of the cascade at depth j and D denotes the maximum depth of the cascade. For all scholars, we extracted their cascades, and computed the properties of cascades: size, depth, average depth and breadth. Clearly, there are some nodes that do not have cascades, since they do not have successors (they are not influencers), thus, we had to exclude those scholars. We also excluded cascades of size 1 (scholars who influence one another only). Therefore, we end up with 4,537 cascades (36% of all scholars).

Table 11 shows for each era the top 5 scholars based on the four features of cascades. We observe that top cascades by size and by breadth correspond to Antiquity intellectuals, whereas, top cascades by depth (and avg. depth) correspond to intellectuals of the Middle Ages (Islam theologians).

However, in order to get insight on how the features of those cascades evolve over time, we compare them over the consecutive eras. Figure 12 shows the distribution of the four features (size, breadth, depth, and avg. depth) over the eras; whereas Table 9 provides, for those features, a statistical summary including the mean, median (50% quantile) and maximum, over the different eras.

At a glance, one can see that the size and breadth features exhibit similar behaviors; while the depth and avg. depth features exhibit similar behaviors as well. We observe that the size of cascades decreases over time until it almost vanishes at Contemporary period. At the first four eras (up to Transition Period) this decrease in size is smooth (average size is above 1500 on average), but it becomes more sharp in the last two eras (average size of less than 500 in Modern Age, and only 40 in contemporary). Moreover, we can see from Table 9, that in the first two eras (AN and MA) the mean size is less than the median, which means that the distribution is negatively skewed; but starting from Early Modern Age, the mean becomes greater than the median, hence, the
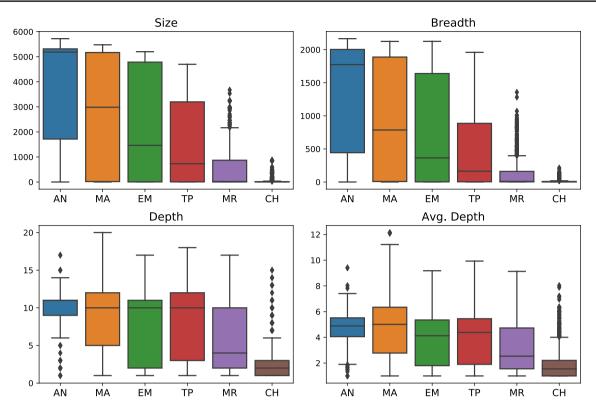
Fig. 12 Distribution of cascades features over eras

Table 9 Statistical summary (average, median and max) of cascade features over the different eras

Era	#	Size			Breadth		
		avg	50%	max	avg	50%	max
AN	177	3857	5189	5723	1371	1775	2167
MA	199	2493	2985	5478	882	787	2125
EM	400	2237	1463	5203	752	365	2126
TP	594	1568	731	4699	455	165	1960
MR	1437	477	20	3675	122	10	1357
CH	1730	40	5	884	13	4	214
Era	#	Depth			Avg. depth		
		avg	50%	max	avg	50%	max
AN	177	9.3	11	17	4.6	4.9	9.4
MA	199	9.1	10	20	5.0	5.0	12.1
EM	400	7.6	10	17	3.9	4.1	9.2
TP	594	7.9	10	18	3.9	4.4	9.9
MR	1437	6.0	4	17	3.2	2.5	9.1
CH	1730	2.9	2	15	1.9	1.5	8.0

distribution is positively skewed. A similar behavior can be observed for the breadth of cascades.

The depth, and avg. depth features also decrease smoothly over time. The median depth is constant (at 10) over three eras, from MA until TP, and drops afterward. We can observe a slight increase in depth and avg. depth in Middle Ages comparing to Antiquity: maximum depth goes from 17 to 20, mean avg.depth goes from 4.6 up to 5.0, and maximum avg-depth goes from 9.4 up to 12.1. Another slight increase in depth and avg. depth is also observed in Transition period

Table 10 Correlation of cascade features

	Breadth	Depth	Avg. depth
Size	0.98	0.71	0.64
Breadth		0.63	0.57
Depth			0.97

comparing to Early-Modern period: mean depth goes from 7.6 to 7.9, maximum depth goes up to 18, median avg-depth goes from 4.1 to 4.4, and maximum avg-depth goes from 9.2 up to 9.9. Moreover, in the first four eras (until TP) the mean depth is less than the median, which means that the distribution is negatively skewed; but starting from Modern Age, the mean depth becomes greater than the median, hence, the distribution becomes positively skewed, with many outliers to the right (high values). The same applies for avg. depth.

Table 10 shows the correlation values between the different features. We can see that there is a very strong correlation between size and breadth (0.98), and between depth and avg. depth (0.97). On the other hand, the correlation between Those values are almost the same over all eras.

In order to get insight on how the cascades evolve over the different eras, Fig. 13 shows scatter plots of several pairs of cascade features.

Figure 13-a shows the relation between the size and the breadth of cascades over time. Besides the linear relation that we can clearly see between these features, we can also observe that in early eras, starting from AN, these features tend to have large values, and over time the values decreases gradually until they become relatively very small at CH. For instance, if we consider the cascades with size ≥ 4500 and breadth ≥ 1500, the fraction of such large cascades is 66% in Antiquity. This fraction drops to 39% in Middle Ages, 31% in Early Modern, and only 3% in Transition period; then it becomes 0% in Modern and Contemporary periods.

Figure 13-b shows the relation between the size and the depth of cascades over time. We can see that at Antiquity most of the cascades have large values of size and depth, while some cascades have small size and small depth. On the one hand the fraction of large cascades with size ≥ 4500 and depth ≥ 9 is 69% in Antiquity, and it drops to 39% in MiddleAges, 33% in EarlyModern, and 3% in Transition period, and 0% afterwards. On the other hand, we observe that cascades that have a small depth (≤ 7) have always a very small size (≤ 400). In other words, we can say that the necessary condition to have a non-small size cascade, is to have a depth of at least 8. The fraction of such small cascades increases over time from 23% in Antiquity, to 93% in Contemporary period.

Figure 13-c shows the relation between the depth and the avg. depth of cascades over time. We can clearly see the linear relation between these features (correlation 0.97) over all eras.

Now, in order to characterize this linear relation we apply a linear regression model, using depth as a dependent-, and avg. depth as independent variable. The result is:

$$depth \sim -1.17 + 2.22 \times avgdepth$$

Similarly, if we use avgdepth as a dependent-, and depth as independent variable, the result is:

$$avgdepth \sim 0.67 + 0.42 \times depth$$

In both cases, the prediction accuracy is pretty high, with $R^2 = 0.94$.

Moreover, in order to characterize the relation between the size of cascades and the other features, we apply a multiple regression model using size as a dependent variable, and depth and breadth as independent variables (we exclude avg. depth to avoid multicollinearity, due to its linearity with depth).

The result of this model is:

$$size \sim -139 + 52.1 \times depth + 2.61 \times breadth$$

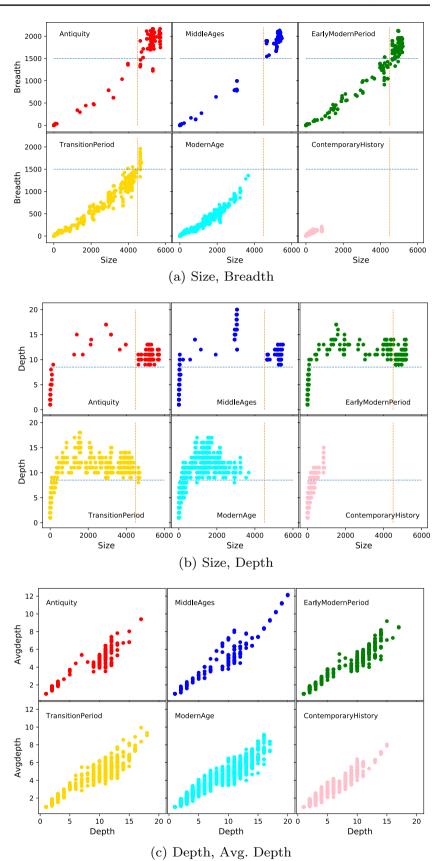with a very high accuracy of $R^2 = 0.98$.

7.1 Clustering of cascades

Based on previous discussion, we notice that most of the cascades tend to have either pretty small values of features, or pretty large values; while intermediate values are little frequent.
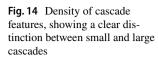
This observation can be verified be looking at Fig. 14 which depicts a kernel density estimate (KDE) plot of each feature, showing the data using a continuous probability density curve. All the four features exhibit two dense regions, that are clearly distinguishable (we approximately separate them using a vertical dashed line), that correspond to small and large cascades.

Moreover, Fig. 15 depicts a violin plot of each feature over all eras, showing the full distribution of features. Here also we can clearly see the two regions, that distinguish small- and large cascades, over the different consecutive eras.

In order to categorize influence cascades based on their aforementioned features, we apply a clustering algorithm, namely K-Means (Lloyd 1982; MacQueen 1967), using the four features of cascades: size, breadth, depth, and avg. depth. The goal is to obtain two clusters of cascades, namely small- and large cascades. Hence, we use $k = 2$ as the number of desired clusters. However, as we have seen, the features are on different scales, for instance, the depth and avg. depth are below 20, while the size and breadth can be above 5000. Therefore, we need to normalize the features
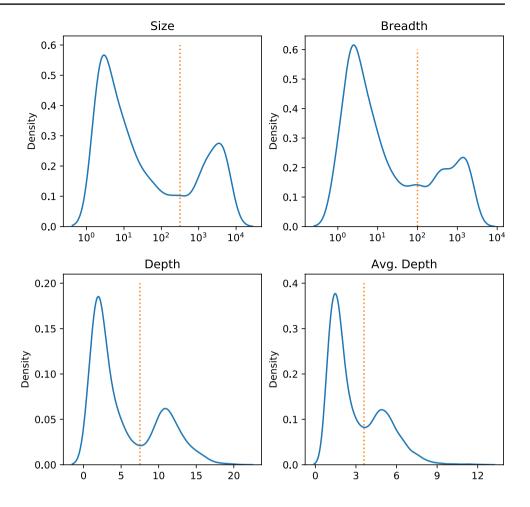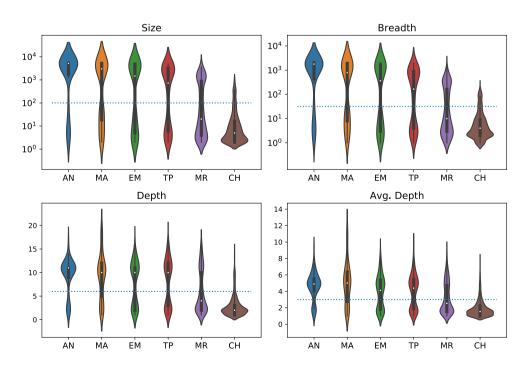
Fig. 13 Relation between features of cascades over eras

(a) Size, Breadth

(b) Size, Depth

(c) Depth, Avg. Depth

Fig. 14 Density of cascade features, showing a clear distinction between small and large cascades

Fig. 15 Distribution of cascade features over eras

Table 11 Statistical summary of the features of small- and large cascades

	Small cascades			Large cascades		
count	3425 (75.5%)			1,112 (24.5%)		
	avg	min	max	avg	min	max
Size	54.4	2	1425	3214.4	653	5723
Depth	3.3	1	13	12.1	8	20
Avg. depth	2.1	1	8.1	5.6	3.4	12.1
Breadth	17.3	1	389	1013.8	102	2167

Fig. 16 Feature distribution of the 2 clusters of cascades

to put them on the same scale. This is done by dividing each feature by its maximum, thus each feature becomes in the range [0,1].

As a result of the clustering, we obtain two clusters of cascades, that can be indeed categorized as small cascades (CS), and large cascades (LC). As shown in Table 11, SC cluster comprises 3425 cascades (75.5%), while the remaining 1112 cascades (24.5%) belong to LC cluster.

The differences between the two clusters are clear. For instance, the size of small cascades is 54.5 on average (median = 6), and 3,214.4 for large cascades (median = 3,056). On the other hand, the depth of small cascades is 3.3 on average (median = 2), and 12.1 for large cascades (median 12).

Figure 16 depicts a violin plot for each feature showing its distribution. One can easily see how distinct the two clusters of small- and large-cascades are.

Moreover, we can also look more closely at these two clusters by looking at the scatter plots of Fig. 17, that show the relation between different pairs of cascade features. For instance, when we look at the relation between size and breadth, we see the cluster of small-cascades (blue) located in a small area at bottom-left (size ≤ 1000, and breadth ≤ 400); however, this small area comprises all small cascades that are more than 75% of all cascades!

Finally, it is of great importance to look at how these two clusters of cascades evolve over time. Table 12 shows the number and fraction of cascades in the small- and large-cascade clusters, over the different consecutive eras. We see that, although the raw number of cascades in both clusters gradually increases over time (except for LC in MA and CH), the fraction of small cascades increases, while the fraction of large cascades decreases over time.

This change in the fractions of SC and LC clusters over time is also reflected in Fig. 18. We observe that in Antiquity, about 75% of cascades are classified as large, and 25% as small. In the next three eras, Middle Ages, Early Modern period, and Transition period, small- and large cascades are almost equally distributed (about 50% each). Then, in Modern Age, large cascades make only 25% of cascades, and the remaining 25% are small. During Contemporary History, almost all cascades (99.4%) are small.

This result makes sense and is pretty reasonable. The longer history a scholar has, the more influence he can give, the bigger his legacy is; and the larger his chains of influence become. In other words, the influence cascade of a scholar is somewhat proportional to how long his history is. On the other hand, recent scholars have not yet enough time to develop large cascades of influence.
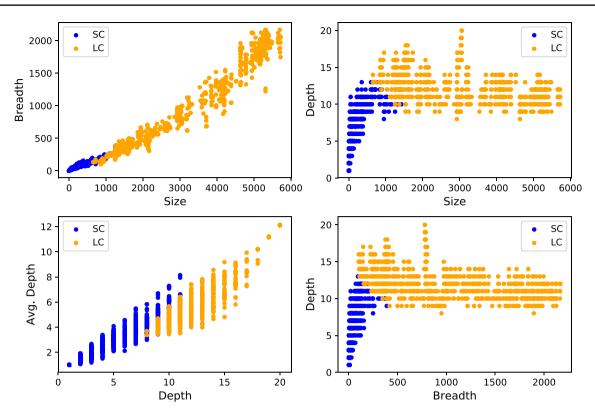
Fig. 17 Distribution of the two clusters of small- and large cascade w.r.t different pairs of features

Table 12 Number and fraction of cascades in small- and large clusters, over eras

	AN	MA	EM	TP	MR	CH
SC	42	87	190	296	1090	1720
	23.7%	43.7%	47.5%	49.8%	75.9%	99.4%
LC	135	112	210	298	347	10
	76.3%	56.3%	52.5%	50.2%	24.1%	0.6%

Fig. 18 Fraction of SC and LC clusters over eras

8 Communities of scholars

In order to get deep insights on how the scholars influence each other, we analyze the community structure in the social network of scholars. A community in a social network is a group of nodes that are relatively densely connected to each other but sparsely connected to other dense groups in the network (Porter et al. 2009).

For this purpose, we applied a community detection algorithm, namely InfoMap[2] algorithm (Bohlin et al. 2014), on our complete influence-based social network of scholars (over all eras). As a result, we obtained 1,772 communities. However, since many of those communities are of small size,

---

[2] InfoMap algorithm is a commonly used community detection algorithm that suits directed networks; it is implemented in the map equation framework.

Table 13 Communities of scholars, top 10 by size. Distribution of member scholars over eras, and notable scholars

#	size	AN	MA	EM	TP	MR	CH	Notable scholars
1	180	0	0	0	0	6	174	R. Pryor, L. Bruce, G. Carlin
2	91	2	2	13	47	12	15	Hegel, Immanuel Kant, Kierkegaard
3	88	2	0	12	14	21	39	F. Hayek, Ludwig von Mises, Adam Smith
4	87	0	0	0	5	28	54	E. Kocbek, F. Prešeren, S. Kosovel
5	65	18	27	13	0	3	4	Thomas Aquinas, René Descartes, Augustine of Hippo
6	60	0	1	0	11	43	5	A. Macedonski, E. Lovinescu, Ion Minulescu
7	60	2	0	0	15	14	29	Karl Marx, L. Althusser, Friedrich Engels
8	60	5	0	19	16	9	11	John Locke, David Hume, John Stuart Mill
9	49	0	0	0	0	21	28	Ludwig Wittgenstein, Bertrand Russell, Rudolf Carnap
10	48	0	0	1	7	25	15	Pablo Picasso, Paul Cézanne, Henri Matisse

we opted to exclude communities that have 5 or less scholars; hence, we have 716 remaining communities.

In each of such detected communities, most of the influence of member scholars goes toward other members of the same community. This means that those scholars belonging to the same community, while influencing each other, are forming a cluster of knowledge, that simulates a school of thought.

It is noteworthy that each of those communities comprises scholars who belong to different eras. This means that the communities are mostly diverse, and open (rather than closed), and evolved over time.

Table 13 provides an overview of the largest 10 communities, sorted by community size (number of member scholars). This table also shows the distribution of member scholars over the different eras, and lists few of notable scholars belonging to that community (top 3 scholars based on out-degree).

The largest community consists of 180 scholars, who are mostly contemporary American actors (mainly comedians). The second largest community comprises 91 scholars, who are mostly philosophers of Transition period, including Hegel, Kant and Kierkegaard. Third and fourth communities, respectively, comprise economists and poets from modern and contemporary periods.

The fifth community mainly comprises scholars from Antiquity and Middle Ages, including Descartes and Tomas Aquinas. Among other noteworthy communities, the community no. 7 which mainly represents the communism school of thought comprising modern and contemporary philosophers such as Marx and Engels; and the community no. 10 which comprises a group of modern famous painters, including Picasso, Cézanne and Matisse.

However, although most of the influence of communities is internal, there is still some observable external influence. That is, in some communities the scholars have influence on other scholars of other communities. Thus, we can measure how a community influences another one by aggregating the influence of individual scholars of the first community on the scholars of the second. In other words, we define the influence of a community A on another community B, denoted $f(A, B)$ as the sum of individual scholar influence over all scholars of A and all scholars of B:

$$f(A, B) = \sum_{a \in A} \sum_{b \in B} f(a, b)$$

where $f(a, b)$ is a function defining the influence of a scholar a on another b, and is given by:

$$f(a, b) = \begin{cases} 1 & \text{if scholar } a \text{ influences scholar } b \\ 0 & \text{otherwise} \end{cases}$$

Using this formula, we calculated the community influence over all possible pairs of the 716 detected communities in our social network of scholars. In fact, there are more than a half million of such pairs of communities; however, only 1% of those pairs exhibit a nonzero influence (about 5 thousand pairs). Even in this tiny portion, for many of these community pairs, the influence was negligible, with value 1 for 73% of cases (i.e., only 1 scholar from one community influences 1 scholar from the other), and value 2 for 16% of cases.

Thus, we opted to retain only the pairs of communities where the value of community influence is greater than or equal 15. The result can be expressed as a directed and weighted network of communities, where the nodes represent the communities, and the directed edges represent the community influence, and the weights represent the aggregated value of individual scholars (of one community towards another). This network of communities is shown in Fig. 19, where each node is labeled by the id of the community, the node size is proportional to the size of the community, the color represents the dominant era of the
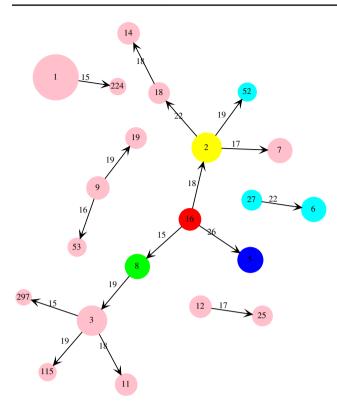
Fig. 19 Network of communities influencing each other

community scholars, and the edges are labeled by the aggregated influence.

For instance, we realize that one of the central communities in this network is the community no. 16, which comprises a group of famous Antiquity scholars, including Aristotle and Plato. This community has a great influence on several other communities, including: community no. 5 (Middle Ages, incl. Descartes and Aquinas), community no. 8 (Early Modern age, incl. John Locke and David Hume), and community no. 2 (Transition period, incl. Hegel and Kant).

9 Conclusion

In this paper, we incorporated a longitudinal aspect in the study of the influence networks of scholars. First, we extracted their social network of influence from YAGO, a pioneering data source of Linked Open Data, which records the main influences of and by intellectuals We opted for a global approach for the periodization of history to match the internationality of scholars, resulting in six consecutive eras to study.

Our main question was whether we could identify patterns of influence, and their change over time. Therefore, we performed essential network analysis on every time-sliced projection of the entire network in within-era, inter-era,

and accumulated-era influence networks. We investigated their social network metrics, degree distribution, and connectivity. An influence pattern throughout all eras was that the internal impact of any era was higher than its external impact. The vast majority of scholars influenced scholars of their own period (= within-era influence) with a relatively stable average out-degree. There were only a few instances of reciprocity. When accumulating eras, the max. degrees drastically increased. However, over all eras the maximum out-degree stayed greater than the maximum in-degree. In inter-era influence networks, each era has the most influence on the consecutive one, and the Contemporary period. The exception to this rule was a spike in the absolute links of antique influences on the Early Modern period, representing the increased reception of antique scholars during the Renaissance. However, proportionally Antiquity's influence on Early Modernity was as high as on the Middle Ages, which reasserts the shift in historical research that the Renaissance thinkers did not "rediscover" Antiquity, but that medieval scholars also received it (Fejfer et al. 2003 p. 3–4).

With a longitudinal perspective, we can add a more pronounced view on who the most influential intellectuals are. The scholar with the highest out-degree over all periods on contemporaries (= within-era) was Modern age scholar Friedrich Nietzsche. Plato in Antiquity, Avicenna in the Middle Ages, John Locke in the Early Modern period, Johann Goethe in the Transition period and Vladimir Nabokov in the Contemporary period were the most influential on the contemporaries of their respective periods.

When accumulating eras, the most influential intellectuals of an era change: here, Plato was the most influential for Antiquity and the Middle Ages, Aristotle for the Early Modern and Transitioning period, Immanuel Kant for the Modern Age. In the Contemporary period, and therefore for the complete network of intellectuals, Karl Marx.

In the inter-era network analysis, Transitioning period scholar Karl Marx had the highest out-degree over all times to the Contemporary age. Modern intellectuals Friedrich Nietzsche and Martin Heidegger took second place over all time for the Contemporary period.

To understand the diffusion dynamics of influence, we constructed influence cascades of scholars, and measured their properties, such as size, depth and breadth. First, we found that those properties decrease over time, which means that the influence cascades are larger for older scholars than for more recent ones. We also analyzed the inter-relations between the properties of cascades, and found that they are positively correlated, in particular, size with breadth, and depth with average depth. We also characterized such relations in form of different linear models, with high accuracy.

Moreover, we found out that the cascades are clustered into two categories, namely small- and large cascades. An interesting finding here is that the fraction of small cascades

increases over time, while the fraction of larges cascades decreases. In particular, the majority of the cascades in Antiquity belong to the large category, whereas in Middle Ages, Early Modern, and Transition periods the cascades were evenly distributed into the small and the large categories. The large cascades became the minority in Modern Age, and almost disappeared in Contemporary History. Hence, we could conclude that the influence cascade of a scholar is somewhat proportional to how long his history is. The longer history a scholar has, the more influence he can give, the bigger his legacy is; and the larger his chains of influence become.

This study of the longitudinal patterns of influence is such suited to further the insights on the interconnections of influence of thinkers and the dynamics of eras alike. Therefore, we plan to study the evolution of communities in these accumulated networks in future work. Another direction of research would be to study the effects of different periodizations on the importance of scholars, as well as deriving an automated periodization based on the dataset. In addition, we would like to compare this YAGO network of intellectual influence with a more detailed network of scholars based on the main books on intellectual history, in order to establish their differences and insights in this field of study.

References

Bohlin L, Edler D, Lancichinetti A, Rosval M (2014) Community detection and visualization of networks with the map equation framework. In: Measuring scholarly impact: methods and practice, pp 3–34

Edelstein D, Findlen P, Ceserani G, Winterer C, Coleman N (2017) Historical research in a digital age: reflections from the mapping the republic of letters project. Am Histor Rev 122(2):400–424

Edelstein D, Kassabova B (2020) How England fell off the map of Voltaire's Englightenment. Modern Intellect History 17(1):29–53. https://doi.org/10.1017/S147924431800015X

Everett M, Borgatti S (2012) Categorical attribute based centrality: E–I and G–F centrality. Soc Netw 34:562–569. https://doi.org/10.1016/j.socnet.2012.06.002

Fejfer J, Fischer-Hansen T, Rathje A (2003) Introduction. In: Feijfer J, Fischer-Hansen T, Rathje A (eds) The rediscovery of antiquity: the role of the artist. Museum Tusculanum Press, Copenhagen, pp 11–20

Gänger S, Lewis SL (2013) Forum: a world of ideas: new pathways in global intellectual history, c.1880–1930. Modern Intellect History 10(2)

Ghawi R, Petz C, Pfeffer J (2019) 'On the shoulders of giants', Analysis of a social network of intellectual influence. In: 6th International conference on social networks analysis, management and security (SNAMS). Granada, Spain, pp 248–255

Ghawi R, Pfeffer J (2019) Mining social networks from linked open data. In: Endres D, Alam M, Şotropa D (eds) Graph-Based representation and reasoning. Springer, Cham, pp 221–229

Gordon PE (2013) What is intellectual history? A frankly partisan introduction to a frequently misunderstood field

Gould RV, Fernandez RM (1989) Structures of mediation: a formal approach to brokerage in transaction networks. Sociol Methodol 19:89–126

Grafton A (2009) A sketch map of a lost continent: the republic of letters. Republic Lett J Study Knowl Politics Arts 1(1)

Grafton A (2009) Worlds made by words: scholarship and community in the modern west. Harvard University Press, Cambridge

Haakonssen K, Whatmore R (2017) Global possibilities in intellectual history: a note on practice. Global Intellect History 2(1):18–29

Hennig M, Brandes U, Pfeffer J, Mergel I (2012) Studying social networks: a guide to empirical research. Campus Verlag, Frankfurt

Holme P, Saramäki J (2019) A map of approaches to temporal networks. In: Holme P, Saramäki J (eds) Temporal network theory, computational social sciences. Springer, Cham, pp 1–24

Huisman M, Snijders TAB (2003) Statistical analysis of longitudinal network data with changing composition. Sociol Methods Res 32(2):253–287

Lloyd SP (1982) Least squares quantization in PCM. IEEE Trans Inf Theory 28:129–137

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, Volume 1: Statistics. University of California Press, Berkeley, CA, pp 281–297

Mahdisoltani F, Biega J, Suchanek FM (2015) YAGO3: a knowledge base from multilingual Wikipedias. In: CIDR 2015, 7th biennial conference on innovative data systems research, Asilomar, CA, USA, Online Proceedings

Mathew B, Dutt R, Goyal P, Mukherjee A (2018) Spread of hate speech in online social media. CoRR abs/1812.01693

Moore W (1994) A life of Erwin Schrödinger. Cambridge University Press, Cambridge

Moyn S, Sartori A (eds) (2013) Global intellectual history. Columbia University Press, New York

Newcomb TM (1961) The acquaintance process as a prototype of human interaction

Osterhammel J (2006) über die Periodisierung der neueren Geschichte (Vortrag in der Geisteswissenschaftlichen Klasse am 29. November 2002). Akademie Verlag, Berlin, pp 45–64

Otte E, Rousseau R (2002) Social network analysis: a powerful strategy, also for the information sciences. Inf Sci 28(6):441–453

Petz C, Ghawi R, Pfeffer J (2020) A longitudinal analysis of a social network of intellectual history. In: IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM). The Hague, Netherlands (virtual), Dec 7–10 (2020)

Porter MA, Onnela JP, Mucha PJ (2009) Communities in networks. Notices AMS 56(9):1082–1097

Pot JHJvd, (1999) Sinndeutung und Periodisierung der Geschichte. Eine systematische Übersicht der Theorien und Auffassungen, Brill, Leiden Bosten Köln

Snijders T, Bunt GG, Steglich C (2010) Introduction to actor-based models for network dynamics. Soc Netw 32:44–60

Subrahmanyam S (2015) Global intellectual history beyond hegel and marx. History Theory 54(1):126–137

Subrahmanyam S (2017) Beyond the usual suspects: on intellectual networks in the early modern world. Global Intellect History 2(1):30–48

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359(6380):1146–1151. https://doi.org/10.1126/science.aap9559

Vugt I (2017) Using multi-layered networks to disclose books in the republic of letters. J Histor Netw Res 1(1):25–51

Wasserman S, Faust K (1994) Social network analysis: methods and applications, 1st edn. Cambridge University Press, New York

Wickberg D (2001) Intellectual history vs. the social history of intellectuals. Rethinking History 5(3):383–395

Winterer C (2012) Where is America in the republic of letters? Modern Intellect History 9(3):597–623. https://doi.org/10.1017/S1479244312000212

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.