Identifying potentially disruptive trends by means of keyword network analysis

Dotsika, F. and Watkins, A.

Identifying potentially disruptive trends by means of keyword network analysis

Dr Fefie Dotsika
Westminster Business School, University of Westminster, London, UK
F.E.Dotsika@westminster.ac.uk

Andrew Watkins
Birkbeck, University of London, London, UK
andrew@dcs.bbk.ac.uk

Abstract

Identifying potentially disruptive technologies is crucial to safeguarding competitive advantage by enabling stakeholders to assign resources in a manner that increases the chances of exploiting the disruption and/or mitigating the ensuing risks. However, disruptive technologies and emergent trends within known disruptive domains are mostly identified ex-post. This paper contributes to the ex-ante prediction of emergent technologies within disruptive domains by proposing a literature-driven method for the forecasting of potentially disruptive technological trends. It adopts a keyword network analysis and visualisation approach for uncovering emergent thematic, structural and temporal developments within publications and applies it as a forecasting tool to an empirical study of seven disruptive domains: 3D Printing, Big Data, Bitcoin, Cloud Technologies, Internet of Things, MOOCs and Social Media. Maturing trends were found to share influential common topics identified by high degree, betweenness and closeness centrality scores. Niche and potentially emerging trends within groups were detected by means of eccentricity and farness metrics. Visualisation techniques were found effective for further clarification and trend identification. Finally, potentially disruptive trends within domains were found to be associated with high closeness paired with low degree centrality. The findings were distilled into a framework for assisting the forecasting of potentially disruptive trends.

Keywords: disruptive technologies, emerging technologies forecasting, keyword network analysis, trend forecasting.

1. Introduction

In an ever-changing technological landscape where innovation is a crucial driver for economic growth and survival, it is desirable to be able to predict which technologies, when established, have the potential to revolutionise an industry, create new markets, increase accessibility and affordability. Studies describing innovation trends, trajectories and future patterns identify drivers such as geographical factors, firm clusters, knowledge flows and spill-over effects (Hausman & Johnston, 2014; Doloreux & Shearmur 2012; Huber 2012; Tappeiner et al., 2008; Gertler & Levitte, 2005). With information technology an integral part of all aspects of organisational life, research on IT innovation constitutes an important driver of organisational competitiveness (Hamel, 1998; Finchman 2004) promoting scalability, sustainability and affordability (Helal, 2015).

Unlike sustaining innovation which supports established improvements to existing products and services, disruptive innovation defines the process transforming a product or service that historically has been accessible at the top of a market access (i.e. for a high price or specialised skill-set) to become accessible to a new and larger population of consumers at the bottom of that market (Christensen, 1997). Disruptive innovation creates a new market and value network which eventually disrupt and displace their predecessors (Christensen & Raynor 2003; Christensen et al., 2004). Depending on the application aspect, disruptive innovations can be categorised as product, business-model or technological innovations (Markides, 2006). A disruptive technology can be thought of as a technology that changes the essence of competition among firms by transforming the performance metrics (Danneels, 2004).

There is evidence suggesting that smaller firms have a potential advantage over larger organisations in that they can leverage their capabilities for innovative solutions and are more agile in dealing with organisational change and with managing disruptive innovation (Hyvonen & Touminen, 2006; Moore & Manring, 2009). In contrast, organisational barriers in large corporations may hinder innovation. These barriers include the existence of a successful dominant design or profitable business concept, possible inability to learn and adopt change, a risk-averse management, the mishandling of the innovation process and an absent or underdeveloped infrastructure (Assink, 2006). Learning competencies in smaller but established companies have been found to have high impact on the degree of novelty of innovation (Amara et al., 2008), which influences marketing positioning and boosts growth (Dotsika & Patrick, 2013). European funding, such as the Open and Disruptive Innovation (ODI, 2014) scheme, aims to promote ideas of high disruptive potential through business innovation grants and the facilitation of consequent commercial exploitation.

Despite these initiatives disruptive innovation trajectory and forecasting are inadequately covered and poorly understood. Disruptive technologies are difficult to predict and are mostly identified ex post (Christensen & Raynor, 2003). Studies on disruptive innovation forecasting recognise the potential of literature based methods (Kostoff et al., 2004). However, no actual method has been proposed. Can the existing literature on current disruptive technologies provide clues on determining future potential trends? What can we learn from the bibliographic differences between business and academic publications on disruptive technologies? And, can keyword network analysis help identify disruptive trends and influencing themes by interpreting the thematic relationships within subject groups? These are the questions addressed in this paper. In it we present a literature-founded approach to uncovering emergent, potentially disruptive trends by analysing the sub-theme associations and timeline of disruptive technologies identified through their presence in business and scholarly articles. In order to do this, we first:

- Identify the major current trends in the field of disruptive technologies.
- Determine and compare the distribution of each of these trends from onset to present in leading business reports and academic publications.

Then, adopting a network approach, we perform a statistical and visual analysis of the data concentrating on its thematic, structural and temporal characteristics with the intention to:

- Investigate and demonstrate the thematic and temporal relationships of relevant academic publications in terms of domain, influence and popularity
- Propose a literature-based framework for assisting the forecasting of emergent trends within disruptive domains.

The rest of the paper is organised as follows: section 2 reviews disruptive technologies' forecasting. Section 3 presents the research design and data collection and section 4 follows the data analysis and interpretation. We discuss our findings in section 5 and identify the implications for research and practice. In the last section, we draw our conclusions and outline future work.

2. Forecasting the trajectory of disruptive technologies

Identifying new potentially disruptive technologies and/or new disruptive trends and applications is a challenge that may be met by anticipating change and preparing for it by way of understanding the dynamics of innovation, identifying the drivers of the future and collecting intelligence (Paap & Katz, 2004). Dissatisfied with plain empirical evidence and ex-post success verification, researchers in the field have debated the predictive use of the theory of technological disruption (Danneels, 2004; Christensen 2006). Models and methods proposed include diffusion forecasting which takes into account the servicing of multiple markets (Linton, 2002), measures of disruptiveness for predicting the disruptive innovation potential of incumbent firms (Govindarajan & Kopalle, 2005; Govindarajan & Kopalle, 2006) and research on R&D strategies for the purposeful creation of technologies with high disruptive potential (Yu & Hang, 2011).

Disruptive innovations are mostly identified ex post (Christensen & Raynor, 2003). Ex ante prediction frameworks are not well established. Adapting existing technology forecasting methods can help with forecasting potentially disruptive technologies (Danneels, 2004) while ex ante predictions about companies with potential to develop disruptive innovations can be made through the disruptive innovation framework (Govindarajan & Kopalle 2006). Technology roadmapping is often used for the forecasting of disruptive technologies (Vojak & Chambers, 2004; Phaal et al, 2004). Use of scenarios can be successfully applied to aid analysis that particularly suits disruptive innovation (Drew, 2006). Approaches to identifying disruptive technologies are discussed in existing roadmaps. Literature-based discovery is recognised as a starting point which leads to better results when combined with a roadmap development process (Kostoff et al., 2004) but not as a method in its own right. Obstacles include the frequent lack of standards, dominant designs and the potential presence of competing and/or complementary manufacturing technologies (Walsh, 2004) as well as a variety of uncertainty factors including technological, market, regulatory/institutional and social/political uncertainty (Jalonen, 2011).

Perspective is critical in understanding and untangling competing terminology issues. For identification and classification purposes it is important to consider marketing, technological, macro- and micro-level perspectives (Garcia & Calantone, 2002). Within a business setting, innovation is managed differently within large companies than it is in small firms (Dotsika & Patrick, 2013).

Existing approaches of ex ante identification of disruptive innovation can be grouped into three categories depending on the focus and analysis position (Keller & Husig, 2009). Scoring models analyse the disruptive potential of new innovations (Rafii and Kampas, 2002; Christensen et al., 2004; Hüsig et al., 2005; Govindarajan & Kopalle, 2006; Sainio & Puumalainen, 2007; Ganguly et al., 2008; Keller and Hüsig, 2009; Hang et al., 2011). The other two groups use scenario analysis, simulating a potential entry and distribution. Economic models focus on an economic perspective (Adner & Zemsky, 2001; Adner, 2002; Schmidt, 2008) and situational models focus on other aspects (Kostoff, 2004, Paap and Katz, 2004 and Vojak and Chambers, 2004).

Continuous monitoring of the technology landscape in one's own industry to identify technologies that are better performance drivers is a necessity (Paap & Katz, 2004). Integrating the literature in technology forecasting is one way to deal with this and help to reveal trends, identify technology or product candidates for potential disruptive innovation (Young et al. 2008; Yu & Hang, 2010).

Literature-based detection of disruptive technologies and, in particular, disruptive trends within existing disruptive domains, is recognised among the studies on disruptive innovation forecasting (Kostoff et al., 2004; Fageberg, 2004; Young et al. 2008). Keyword co-occurrence and network analysis methods have been used for bibliometric analysis in the area to identify technological trends (Choi et al., 2011a; Li et al., 2016; Wu, 2016), analyse research topics (Wang et al., 2016), follow their evolution (Ye et al., 2015) and track the development of innovation system research (Liu et al., 2015). Similar methods have been implemented on patent analysis for the identification of appropriate technology opportunities (Lee et al., 2014; Kim et al., 2014), the detection of technology trends, significant patents and novel technologies that enable strategic technology planning (Park et al., 2013) and the improvement of technology development efficiency (Choi & Hwang, 2014). Social network analysis methods focusing on centrality measures have been successfully employed to identify dominant areas of operations management research (Behara et al., 2014) while visualisation methods have been found effective in creating knowledge maps exploring research themes, monitor research trends and discover interdependencies between research areas (Yoon et al., 2010; Lee & Su, 2010; Yang et al., 2016). Forecasting research has employed keyword network analysis focusing on clustering and distribution to identify and predict research trends (Choi et al., 2011b) and visualisation to understand advances of emerging technologies (Kim et al., 2008)

The research presented in this paper extends the use of network analysis in forecasting by employing positional influence metrics and visualisation to complement distribution and clustering and by applying it in the domain of disruptive technologies. Our contribution is a literature-based method and resulting framework for *the identification and forecasting of emergent technologies within disruptive domains*. We assume known disruptive domains and existing publications on these domains.

3. Research design

The study adopts a network analysis approach and applies it to bibliometric data of publications on selected disruptive technologies. Network analysis methods are best known for their application in social environments (social network analysis) where they are applied to the study of social relations among a set of actors (Borgatti et al., 2002; Wasserman & Faust, 1994).

Here, authors' keywords, considered representative of the core concept and focus of an article, are used as the network's focal entities. A group of publications in a thematic domain is represented as a conceptual network of keywords (nodes) and keyword relationships (edges). Keywords are considered related if they are co-occurring in a publication. The network conceptualises structure as themes and patterns of relations among these themes. By substituting keywords for actors, the method investigates, maps and analyses the thematic relationships, trend distribution and thematic flows within a semantic domain.

Keyword network analysis has been explored in similar research in language topologies and cognitive science (Motter et al., 2002), bibliometrics (Chiu & Ho, 2007; Lee & Su 2010) and trend discovery (Duvvuru et al, 2012). The method's suitability is based on the fact that it is distinctive from other perspectives in that it uses relational information to study structural properties and, as such, it focuses on the resulting structures, their impact and evolution (Wasserman & Faust, 1994).

The publications selected were on the top seven technologies identified as disruptive, current and influential by business research on top trends (Gartner 2015; Forbes 2015; Frost & Sullivan 2015; McKinsey, 2015; KPMG, 2015) and government policy publications (EUCommissionA 2015; EUCommissionB 2015; NESTA 2015; OECD, 2015). The technologies considered were:

1. 3D Printing
2. Big Data (the potential of their analysis as opposed to the data itself)
3. Bitcoin
4. Cloud Technologies
5. Internet of Things (IoT)
6. Massive Online Open Courses (MOOCs)
7. Social Media

From the above there was general consensus on the disruptive nature of 3D printing, big data, cloud technologies, IoT and social media. Bitcoin and MOOCs appear in all reports but there is partial agreement about their disruptive character. However, they were included due to their relative newness and idiosyncratic nature: MOOCs (Kop & Hill 2008) for bringing forth a disruptive education model (Gasevic et al., 2014; Bulfin, 2014; Kovanovic et al., 2015; Salmon et al., 2015) and bitcoin (Nakamoto, 2008) for its particularly challenging nature in terms of regulatory issues and monetary policy (Grinberg 2012; Dwyer, 2015; Liping, 2013).

Bibliographic data was collected from scholarly articles and business publications. Each disruptive technology corresponds to a semantic domain. Each domain comprises all articles published on the specific technology from its onset to the end of June 2015, conceptualised as a network of keywords.

Academic publications' data were collected through the Thompson Reuters' Web of Science API. The authors' keyword data formed seven groups, one for every thematic domain. The keywords used in the search can be seen in Table 1. The wild character "*" substitutes one or more characters. *Keywords* correspond to the Web of Science *TS* (Topic) field. In the case of MOOC we cleaned the data to exclude the *Multiple Optical Orthogonal Codes* (also MOOC). All keywords were collected and cleaned for synonyms (e.g. some authors use "IOT", others "internet of things"), homonyms (e.g. internet of things, internet-of-things), discrepancies in granularity (i.e. IOT, IOTs) and possible misspellings. Keywords appearing on the same article formed linked pairs. Information on each domain was modelled as a graph providing a means of representing relations between nodes and quantifying structural properties (Wasserman & Faust, 1994). For each group of publications, a network (graph G) was created with the keywords serving as nodes (V) and their links as edges (E). All resulting networks are undirected (i.e. the edges are bidirectional, that is if keyword A is co-occurring with B, the opposite also holds) and weighted (multiple possible occurrences of keyword pairs).

Domain	**Keyword** (Web of Science Topic, or TS field)	Cleaning filter
Bitcoin	"bitcoin*" OR "bit-coin"	N/A
MOOCs	"MOOC*" OR "massive online open course*"	-"multiple optical orthogonal codes"
3D Printing	"3D print*" OR "three dimensional printing"	N/A
Internet of Things	"internet*of*things" OR "iot" OR "internet of things"	N/A
Big Data	"big data" OR "big*data"	N/A
Social Media	"social media" OR "social*media"	N/A
Cloud	"cloud computing" OR "cloud security" OR "cloud technolog*" OR "cloud serv*" OR "cloud networking" OR "cloud infrastructure" OR "cloud solution*" OR "private cloud" OR "virtuali*ation"	N/A

Table 1. Web of Science keyword search

Business reports were sourced from the top five industry analyst sources chosen due to (a) their quality standing and capacity as a key source of knowledge about technological innovations (Wang & Ramiller, 2009) and (b) their influence on customers worldwide (Ikeler, 2007; Bernard & Gallupe, 2013). The sources were Forrester, Frost & Sullivan, Gartner, IDC and Ovum. The data collected were compared to the academic publications to establish disparities in volume and time and ascertain whether business/practitioners, researching emerging technologies, would be better off focusing their research sources accordingly (Ware, 2009; Hughes et al., 2008). Business publications, however, are proprietary and the data collected was limited to keyword searches returning the number of publications per subject per year. Without subscription access the data is of lesser quality, mainly due to search disparity, inconsistency in the type of publications across providers and lack of reserved keywords. Business publications' constraints on data (heterogeneous searching facilities, accessibility issues and lack of author keywords) disqualified their use in network analysis.

4. Data analysis

The first part of the analysis (*Distribution of academic and business publication*) compares the distribution of the identified disruptive technologies in the chosen business reports and academic publications. The rest of the analysis applies network analysis techniques to investigate and demonstrate the publications' thematic and temporal relationships in terms of domain, influence and popularity.

The keyword networks were analysed in three parts. The first part analyses the networks' general characteristics (*The keyword networks*). These are of interest because of what they imply about network structural cohesion, thematic sub-groups and communities and what they reveal about thematic relationships between domains. The second and third parts carry out analysis on centrality measures which are the indicators of positional influence within sets. Mature sets (each set comprising all publications on a given domain to date) were analysed first (*Positional influence* section). Positional influences of forecasting nature were tested upon earlier sets for validation (*Pre-maturity positional influence*).

4.1. Distribution of academic and business publications

The entire dataset comprises 61,308 business and 31,386 academic publications (Table 2).

	Gartner	Forrester	IDC	Frost & Sullivan	Ovum	Web Of Science
3D printing	345	39	349	325	3	1250
Big data	3281	2038	4636	997	646	3834
Bitcoin	62	30	36	12	7	87
Cloud	15297	3903	5827	1967	1827	15149
Internet of Things	1807	472	2118	572	183	2937
MOOCs	52	2	15	8	18	285
Social media	4613	5319	2746	1444	312	7844
Total	25457	11803	15727	5325	2996	31386

Table 2. Publications by subject and source

The relationships and contributions of the individual items are depicted in the stacked column chart in Figure 1.

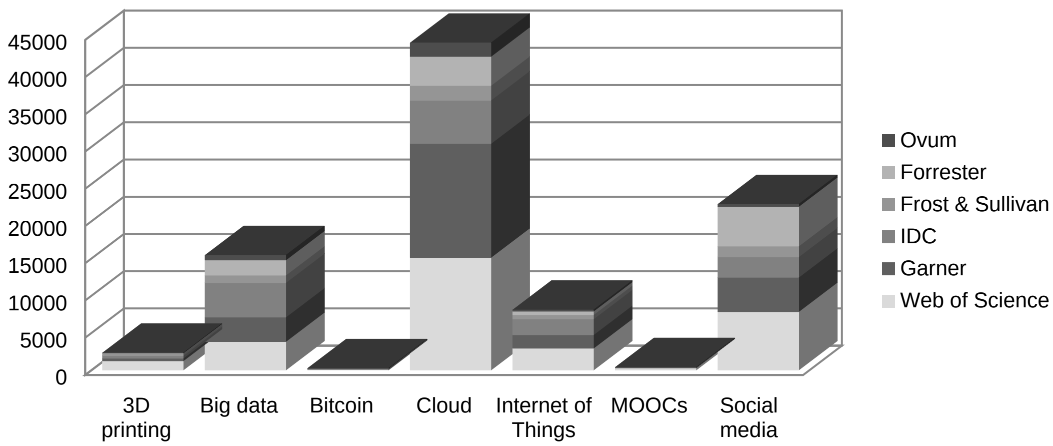


Figure 1. Overall publications by subject area

The chronological distribution of publications by subject area is presented in Figure 2. Contrary to popular belief that expects business publications to be ahead of scholarly articles in covering innovation and technology trends, ignoring the volume difference (which may well be attributed to noise in the results), there is no overall 'lagging behind' evidence (time-wise) in the scholarly publications.

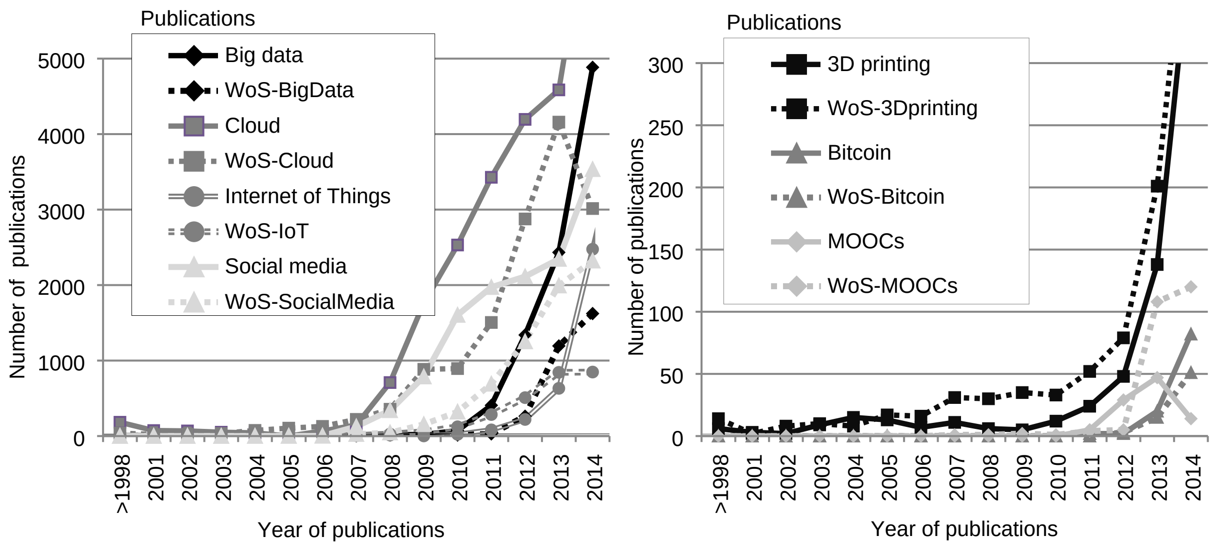


Figure 2. Chronological distribution by subject area

4.2. The keyword networks

For the academic articles we followed a graph-based network approach through a statistical and visual analysis of the seven networks' characteristics. This was carried out by means of UCINET (Borgatti et al., 2002) and Gephi (Bastian et al., 2009).

Network analysis formally expresses concepts of structure and position (Borgatti et al, 2013). Basic network properties were calculated as the first step towards understanding each network's structure. The next step was to compare the metrics with those of existing models and match the networks to topologies of similar statistical properties, thus obtaining a platform on which we could investigate and analyse thematic relationships within domains.

The metrics included the network size (thematic richness), density (ratio of the number of keyword co-occurrences to the number of all possible co-occurrences), diameter (the longest of all the calculated shortest paths) and average degree (degree of a keyword is the number of edges connected to it). Therefore, for graph G=(V,E), where V is the set of nodes and E□V×V the set of links, the size is n=|V|, the number of links m=|E|, the average degree k=(2m/n) and the density δ(G)=(2m/n(n−1)).

Clustering and sub-network metrics were applied to analyse the formation of thematic clusters and the integration of themes within them. The metrics calculated were clustering coefficient (clustering coefficient of a node is the ratio of existing links connecting a node's neighbours to each other to the maximum possible number of such links), Erdös number (distance of a node to a given Erdös node; here the main keyword), average path length (average number of steps along the shortest paths for all possible pairs of keywords), average embeddedness (embeddedness of an edge is the number of common neighbours the edge's endpoints have), modularity (the strength of clustering calculated as the fraction of edges falling within the given groups minus the expected such fraction if edges were randomly distributed). Community detection was done by means of the Louvain method (Blondel et al., 2008). Cliques were not considered, as they are a very strict definition of a cohesive group (absence of a single edge disqualifies an otherwise fully connected subgroup) and do not provide much information, especially in large data sets (Wasserman & Faust, 1994). An additional constraint here was the existence of papers with a large number of keywords that create instant pseudo-cliques (bitcoin: 10; MOOCs: 13; 3D printing: 24; internet of things: 50; big data: 41; social media: 65; cloud: 37). The results can be seen in Table 3. The network visualisations are indicative of the corresponding network populations. Due to network size and tool constraints they are neither standardised nor comparable.

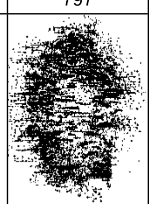
		Bitcoin	MOOCs	3D Printing	Internet of Things	Big Data	Social Media	Cloud
Basic	Nodes (keywords)	116	601	2657	5686	7480	12452	20342
	Edges (keyword co-occurrence)	310	2110	9959	23534	31844	62017	89552
	Density	0.046	0.012	0.003	0.001	0.001	0.001	<0.001
	Diameter	4	5	6	8	8	8	9
	Average degree	5.345	7.002	7.496	8.278	8.514	9.961	8.805
Clustering & sub-networks	Clustering coefficient	0.945	0.922	0.923	0.868	0.908	0.874	0.879
	Erdos number	1.135	1.166	1.319	1.495	1.399	1.449	1.534
	Weakly connected components	4	7	63	114	179	181	283
	Average path length	2.162	2.286	2.571	2.857	2.759	2.803	2.903
	Average embeddedness	4.026	5.079	6.369	10.305	7.879	8.962	7.098
	Modularity	0.646	0.486	0.624	0.547	0.566	0.478	0.470
	Communities	14	63	214	451	417	460	797
	Graphical representation							

Table 3. Network metrics

Basic and clustering network metrics show that all networks share several nontrivial properties of large real-world networks which distinguish them from random networks (Wasserman & Faust, 1994; Newman, 2003; Latapy et al., 2008). These include low density (most pairs of keywords are not co-occurring) and high clustering coefficients (there is high overall interconnectivity between keywords). So, while most pairs of keywords are not directly linked due to low density, they have a common neighbour. This topology enables shortcuts between clusters connecting varied thematic domains by a short average path length. Another shared characteristic, in line with real-world networks, is a highly skewed degree distribution, indicating the existence of a main body of keywords with a large number of direct links (main theme), followed by a long tail of keywords with a small number of connections (related, weakly-connected subjects within thematic domains). Marginal and emerging trends belong in the latter category. The degree distributions suggest scale-free network formations: they are highly heterogeneous and follow approximately a power law $p_k \propto k^{-\alpha}$ where α is a constant (Faloutsos et al., 1999). The individual distributions can be seen in Figure 3 (log-log plot).

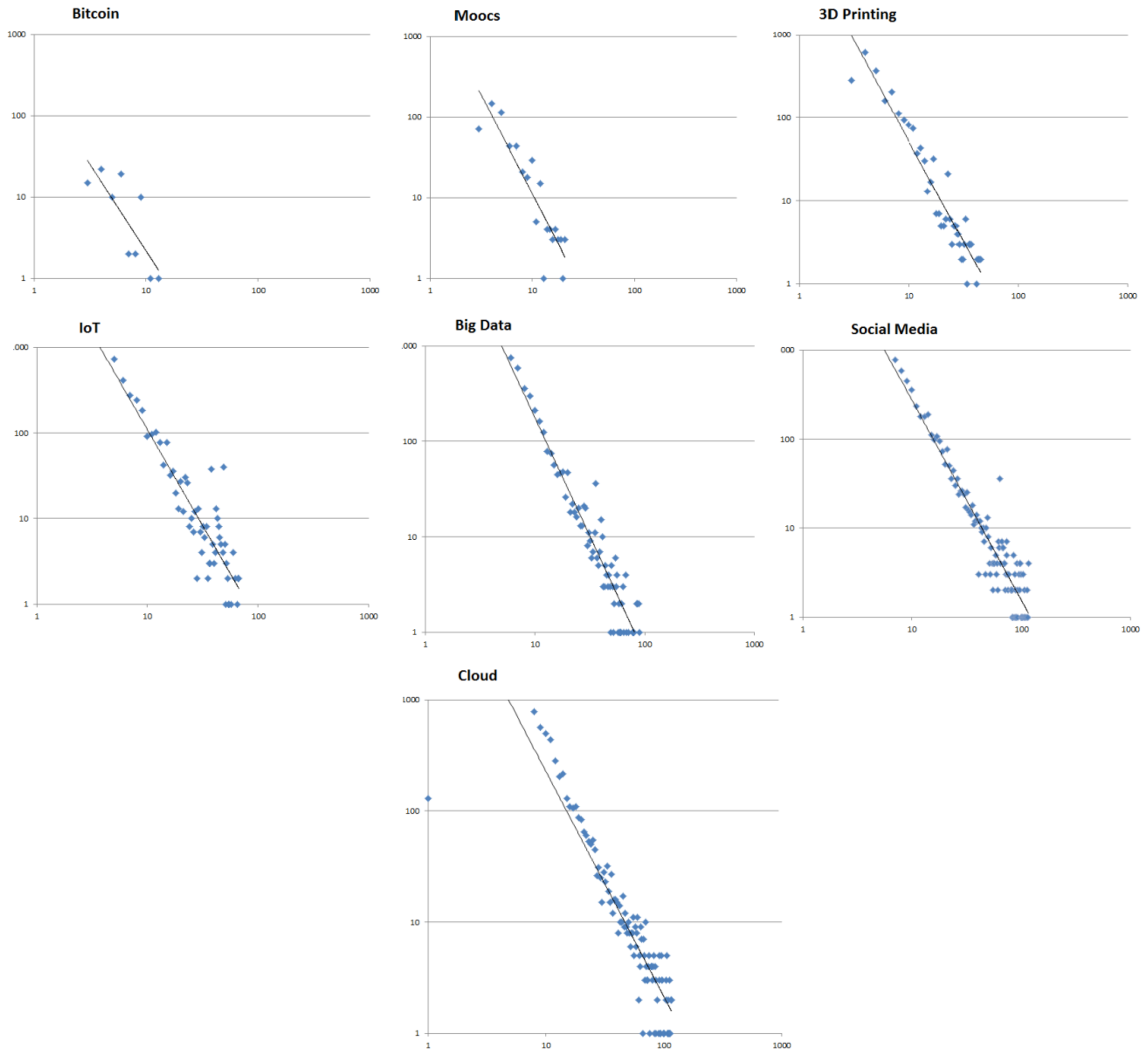


Figure 3. Degree distributions for the seven networks

All seven networks are highly modular in nature with positive modularity values (possible values are between -1 and 1). The *Bitcoin* and *3D Printing* groups have the highest modularity scores with *Cloud* coming last (0.470). They all fit the small-world type characterised by short average path length and high clustering coefficient. The modularity metric that measures the strength of clustering is high among all groups (*Cloud* has the lowest modularity of 0.470). Each network comprises tightly-knit, loosely connected thematic communities with short distance between keywords. These combine together to form larger but less cohesive communities.

By design all networks have a giant component (centred on the theme key node) where most nodes belong. The Erdös number (the average distance to the Erdös node, i.e. main keyword in each network) is low in all networks, ranging from 1.135 to 1.534.

Embeddedness implies thematic shared context and quantifies how well nodes/keywords are connected to the rest of their community/theme. The *Internet of Things* group has the highest average embeddedness (10.305). The group has the lowest clustering coefficient (0.868). However, the value is still high and not considerably lower to the rest of the group to justify alternative interpretation.

4.3. Positional influence

Positional analysis of the networks was carried out by means of centrality metrics. Centrality measures indicate which nodes occupy significant positions in the network (Freeman, 1979; Wasserman & Faust, 1994). Degree centrality (number of ties of a node), eigenvector centrality (relative score indicating how well a node is connected to other well-connected nodes), betweenness centrality (how many times a node acts as a bridge along the shortest path between two other nodes), closeness centrality (a node's distance to all other nodes) and eccentricity (distance from a node to the farthest node) were considered. Centrality metrics here correspond to the popularity of keywords/themes and the influence their position holds in bridging trends and controlling thematic flows.

Degree centrality of node *i* is defined as $C_D(i) = \sum_{j=1}^{n} x_{ij}$ where x_{ij} is a tie from node *i* to node *j.* The higher the centrality, the more central in terms of semantic importance the node/keyword is in the network. The nodes with the highest degree and eigenvector centralities are the same in each group. This is not surprising because eigenvector centrality (calculated here with power iteration of 100) expands on the notion of degree centrality, as it is the sum of a node's ties to other nodes weighted by their degree centrality. The nodes with the highest betweenness centrality are more varied though still overlapping. Betweenness centrality of node *i* is defined as $C_B(i) = \sum g_{jk}(i)/g_{jk}, i \neq j \neq k$ where g_{jk} is the number of geodesics (i.e. shortest paths) connecting *j* and *k* and $g_{jk}(i)$ is the number of geodesics connecting *j* and *k* that pass through node *i*. In the current setting, the higher the betweenness centrality, the more times the keyword acts as a bridge between other nodes and therefore the more influential it is in the flow of information in the network. Closeness centrality is defined as $C_c(i) = \sum_{j=1}^{n} d_{ij}$ whered_{ij} is the distance connecting node *i* to node *j.* It captures the degree to which a keyword/theme is near all other keywords in a network. It seems a misnomer as in fact the calculation represents the farness score of the node: the higher the score the greater the distance.

Notable results identifying central thematic trends within groups can be seen in Table 4 below. For degree, eigenvector, betweenness and closeness centralities the two keywords with the highest values were recorded in the table, excluding the groups' core (i.e. main keyword). In the case of multiple keywords with the same centrality, the two with the higher other centralities are recorded in the table.

Metric	Bitcoin	MOOCs	3D Printing	Internet of Things	Big Data	Social Media	Cloud
Degree centrality	security silk road	e-learning connectivism	rapid prototyping additive manufacturing	cloud computing wireless sensor networks	cloud computing social media	twitter facebook	virtualization
Eigenvector centrality	security silk road	e-learning connectivism	rapid prototyping additive manufacturing	cloud computing wireless sensor networks	cloud computing social media	twitter facebook	virtualization security
Betweenness centrality	security trust	e-learning online education	rapid prototyping additive manufacturing	wireless sensor networks rfid	cloud computing mapreduce	twitter facebook	virtualization network

Closeness centrality	cryptograp hy privacy	higher education e-learning	hybrid cell-biomaterial printing computer aided	security saving energy	event recognition data-driven learning	adult learning politicalfor	health cloud security

Table 4. Central trends within groups

The overall results show certain strong thematic relationships between groups and especially among the four ‘older’ domains (*Internet of Things, Big Data, Social Media* and *Cloud*). The *Big Data* group nodes with the highest degree and eigenvector centrality are ‘social media’ and ‘cloud’ (‘internet of things’ comes 7th), and the highest number of ties in the *Internet of Things* group belongs to ‘cloud computing’. High degree and betweenness centralities belong to popular *and* bridging keywords, a relationship that suggests recurring and/or established themes. High betweenness and closeness centralities suggest central keywords that act as hubs to thematically related flows of information.

The highest centrality common (and recurring with various modifiers) keyword in most groups is ‘security’. Other keywords with relatively high degree and/or closeness centralities are ‘quality’ (without or with various modifiers such as ‘- control’, ‘- of service’, etc.), ‘privacy’ and ‘learning’/‘education’. Keywords associated with profit (e.g. ‘profit maximization’, ‘revenue’, ‘revenue maximization’) have high centrality in *Cloud* but are almost non-existent in the other groups. Keywords associated with finance (e.g. ‘finance’, ‘financial performance’) have very low centrality metrics in all groups. Despite expectations, the keyword ‘innovation’ (with or without descriptors, e.g. ‘- management’, ‘- adoption’, etc.) is not an obvious link term in any group (i.e. has relatively low betweenness centrality). A collection of common recurring keywords of influence in all groups is presented in Table 5.

Keyword		**Bitcoin**	**MOOCs**	**3D Printing**	**IoT**	**Big Data**	**SM**	**Cloud**
education/ learning	C'_D	-	0.15	0.01	0.01	0	0.02	0.01
	C'_C	-	0.52	0.62	0.42	0.73	0.75	0.50
	C'_B	-	0.04	0	0	0	0	0
internet	C'_D	0.06	0.01	0.01	0.02	0.01	0.08	0.01
	C'_C	0.51	0.46	0.43	0.44	0.43	0.48	0.44
	C'_B	0.02	0	0	0.01	0	0.03	0
performan ce	C'_D	0.08	0.01	0	0.01	0.01	0.01	0.04
	C'_C	0.36	0.46	0.44	0.43	0.48	0.44	0.51
	C'_B	0	0	0	0	0	0	0.01
privacy	C'_D	0.05	-	-	0.03	0.03	0.02	0.03
	C'_C	0.75	-	-	0.45	0.48	0.49	0.48
	C'_B	0	-	-	0.01	0	0	0.01
quality	C'_D	-	0.03	0	0	0.01	0	0.02
	C'_C	-	0.48	0.44	0.43	0.43	0.43	0.47
	C'_B	-	0	0	0	0	0	0.01
security	C'_D	0.11	-	-	0.16	0.08	0.02	0.05
	C'_C	0.53	-	-	0.71	0.43	0.42	0.71
	C'_B	0.12	-	-	0.02	0.01	0	0.02
big data	C'_D	-	-	0.01	-	0.57	0.02	0
	C'_C	-	-	0.44	-	0.71	0.44	0.42
	C'_B	-	-	0	-	0.67	0.01	0
cloud	C'_D	0.03	0.03	0	0.10	0.15	0.01	0.48
	C'_C	0.48	0.48	0.43	0.49	0.50	0.42	0.65
	C'_B	0	0	0	0.07	0.04	0	0.61
IoT	C'_D	-	-	0.01	0.50	0.03	-	0
	C'_C	-	-	0.44	0.67	0.44	-	0.41

	C'_B	-	-	0	0.65	0	-	0
social media	C'_D	0.03	0.07	0	0	0.07	0.54	0
	C'_C	0.48	0.49	0.43	0.42	0.47	0.69	0.41
	C'_B	0	0	0	0	0.01	0.67	0

Table 5. Common influential keywords

Single occurrences have been recorded here, leaving out composite terms (e.g. ‘quality’ but not ‘quality of service’). The degree, betweenness and closeness centralities have been normalised. Normalised degree centrality is $C'_D(i) = C_D(i)/n - 1$ where n is the number of nodes. Normalised betweenness is $C'_B(i) = C_B(i)/[(n - 1)(n - 2)/2]$. Normalised closeness is $C'_C(i) = [C_C(i)]^{-1}(n - 1)$.

Closeness (here in its original definition of “farness”) together with eccentricity yield some further interesting results. Eccentricity of node *i* is a complementary metric, defined as $E(i) = \max\{g_{ij}, \forall j \in V\}$ and is the largest geodesic distance between *i* and any other node *j* (diameter is the maximum eccentricity). It captures the distance between a node/theme and the node that is furthest away from it. Low scores of both measures (indicating high centrality of the node) give results not dissimilar to the other closeness metrics. Low farness indicates central keywords/themes that have a short average distance to all other keywords in the network. Low eccentricity similarly captures well-connected themes, with low distance to the farthest keywords in the network. Unsurprisingly, the most notable examples of lowest farness and eccentricity were the keywords ‘cryptography’ in *Bitcoin,* ‘e-learning’ *in MOOCs* and ‘security’ *Internet of Things*. High scores however indicate niche themes within groups. The highest scores for closeness centrality and eccentricity belong to special connected but not central keywords which may correspond to specialised publications or represent emerging trends within a given semantic domain. Nevertheless, there is no obvious way to distinguish between specialist keywords and emerging trends. Noteworthy examples of the highest scores can be seen in Table 6.

Metric	Bitcoin	MOOCs	3D Printing	Internet of Things	Big Data	Social Media	Cloud
Closeness (farness) centrality	governance leon4processor	persistence non-traditional students	pneumatic diaphragm actuator micro-droplet jetting	design for the environment massive machine communicatio	surveillance length of stay	url pattern learning itf regex forum crawling	kishino knot bracket polynomial minimal surface
Eccentricity	governance drug markets passive resource discovery leon4processor	persistence non-traditional students open virtual worlds 3D web	pneumatic diaphragm actuator micro-droplet jetting orthodontic appliances cnc milling machine	design for the environment ultra-reliable communications massive machine communications future mobile	demand-driven improvement nursing assessment computerised medical records surveillance multi-query optimisation	itf regex url pattern learning forum crawling canonical correlation analysis heterogeneous feature spaces	kishino knot bracket polynomial minimal surface representation complexity analysis gustafson’s law

Table 6. Specialised themes within groups

Graphic representations of the MOOCs network with nodes ranked according to their (a) closeness centrality and (b) eccentricity can be seen in Figure 4. The closest and least eccentric nodes are light grey. Higher centrality (farness) or eccentricity nodes are larger and deeper grey: the higher the score the deeper the colour. These suggest marginal or emerging trends. The MOOCs network was chosen because of its size which is small enough to achieve useful visualisation and large enough to convey interesting information.

The least central and most eccentric keywords can be seen on the right of each diagram connected to the rest of the network through the keyword *retention*. The diagrams are similar apart from differences in the range (denoted by the size and shade of the nodes) and the “lack of eccentricity” of certain nodes that stand out in the second graph.

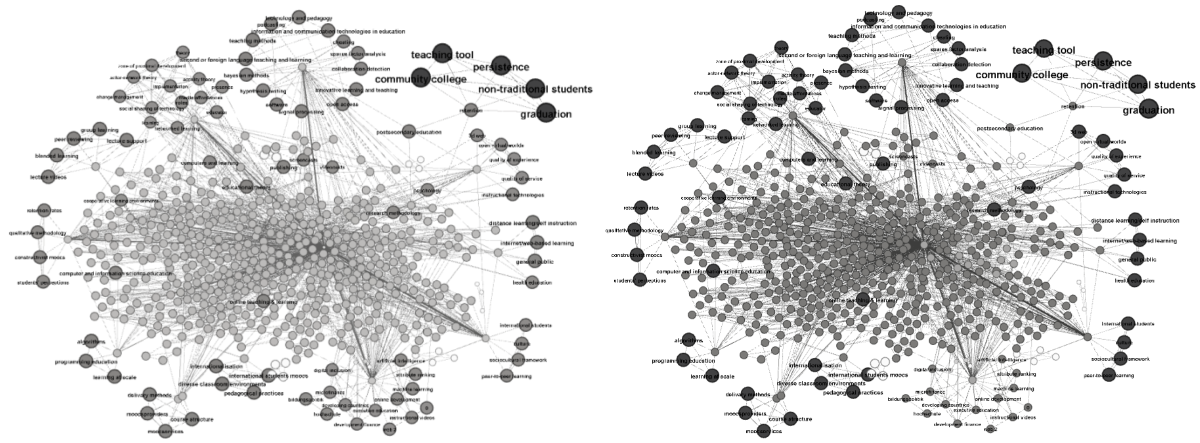


Figure 4. Closeness (here Farness) and Eccentricity as an identifier for marginal trends

Keywords with degree metrics at odds, that is, keywords whose degree, betweenness and closeness were not positively related, provided a different insight to the network structure (Liang & Chen, 2011). Table 7 interprets the findings and presents selected results.

Not all keywords with centralities at odds are recorded here. Common keywords of general, non-domain-specific nature (e.g. digital technologies, manufacturing, classification) were omitted. Keywords acting as descriptors or modifiers (e.g critical, distributed, allocation) were perceived as 'incomplete' and were similarly excluded.

	Low Degree	Low Betweenness	Low Closeness
High Degree		Popular mature keyword. Thematic links bypass it. tissue engineering (3DP) social media (BD) anonymity (Bitcoin) security (Cloud) zigbee (IoT)	Popular niche keyword. Embedded in thematically linked faraway cluster. nitinol wire (3DP) silk road (Bitcoin)
High Betweenness	Bridging infrequent keyword. Its few ties are crucial to thematic network flow. exoskeleton (3DP) [relative oddness] mapreduce(BD)		Bridging niche keyword. Monopolises thematic ties between mainstream and marginal trends. None found
High Closeness	Central and infrequent keyword. It's linked to key themes alginate-hyaluronic acid scaffold (3DP) proteomics (BD) stacked autoencoders (BD) distributed denial-of-service (Bitcoin) block chain (Bitcoin) organic computing (IoT)	Central and mature keyword. Well established along with many others. saas, iaas (Cloud) cloud (BD) bioinformatics (BD) money (Bitcoin) cryptography (Bitcoin) wireless sensor networks (IoT) m2m (IoT) connectivism (MOOCs) higher education (MOOCs) adult learning (SM)	

Key

3DP: 3D Printing

BD: Big Data

IoT: Internet of Things

SM: Social Media

Table 7. Centralities and trend interpretation

4.4. Pre-maturity positional influence

The thematic trends identified in the previous section correspond to a level of disruptive maturity: the disruptive technologies have been about for a number of years and are well established. The last part examines temporal variations by identifying earlier thematic trends within groups. This allows for some retrospective validation of the findings. The cut-off points were chosen to correspond with the end of linear distribution of the number of publications. This corresponds to publications before 2009 for Cloud, 2010 for IoT and Social Media, 2012 for 3D Printing and Big Data and 2013 for MOOCs. The Bitcoin publications are still following a linear distribution so the whole sample was considered.

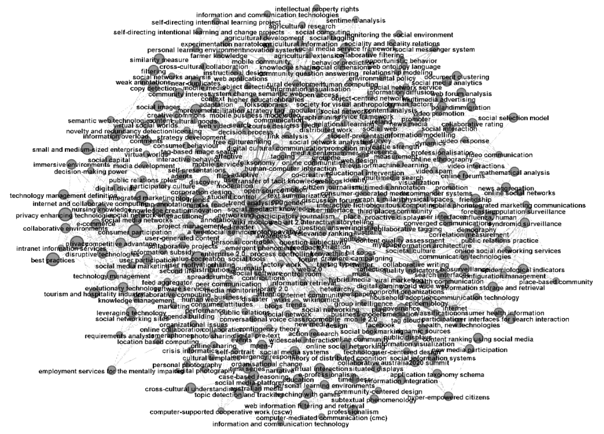
The central trends within groups and influential keywords were identified by high centrality metrics and found not dissimilar to those studied in the whole samples. The groups are thematically linked though the common threads are not as well-established as in the mature groups. As such, these central themes were not studied further. Likewise, farness and eccentricity identified marginal trends for the selected time period. Some of these trends could be classified as low–end potentially disruptive (e.g. mesh networks, bio-/syndromic- surveillance) but there is no way to distinguish between them and niche topics.

Similarly to the mature sets considered in the previous section, keywords whose degree, betweenness and closeness were not positively related, provided again the most interesting results (Table 8).

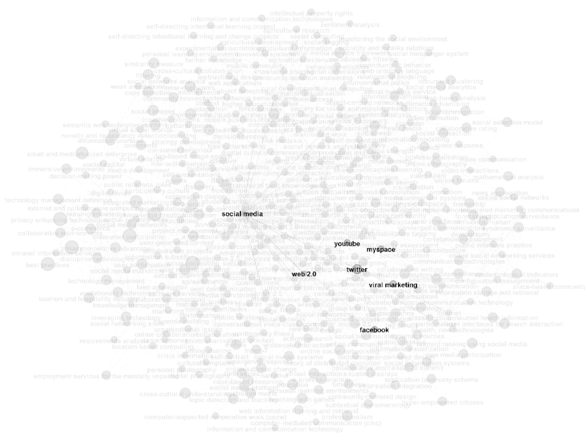
	Low Degree	Low Betweenness	Low Closeness
High Degree		Popular mature keyword. Thematic links bypass it. cloud computing (BD) web 2.0 (BD) performance (Cloud) biometrics (IoT) social networks (SM) anonymity(Bitcoin)	Popular niche keyword. Embedded in thematically linked faraway cluster. data grid (BD) storage visualisation (Cloud) future internet (IoT) sensors (IoT) silk road (Bitcoin)
High Betweenness	Bridging infrequent keyword. Its few ties are crucial to thematic network flow. digital divide (SM) [relative oddness]		Bridging niche keyword. Monopolises thematic ties between mainstream and marginal trends. None found
High Closeness	Central and infrequent keyword. It's linked to key themes biological applications of polymers (3DP) programmable networks (Cloud) fuzzy clustering (BD) communication system security (IoT) rfid technology (IoT) twitter (SM) youtube (SM) facebook (SM) e-campaigning(SM) distributed denial-of-service (Bitcoin)	Central and mature keyword. Well established along with many others. artificial organs (3DP) rapid tooling(3DP) prototype(3DP) visualisation (BD) data grid (Cloud) intelligent storage(Cloud) SOA (IoT) diversity (MOOCs) openness (MOOCs) web 2.0 (SM) money (Bitcoin) cryptography (Bitcoin)	

Table 8. Trend interpretation

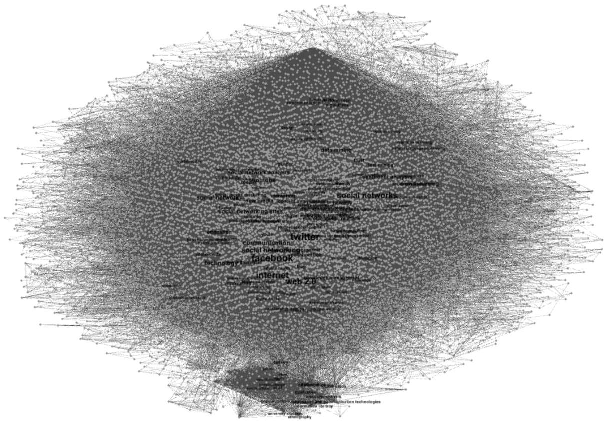
From the above, the keywords with high closeness and low degree are the most noteworthy. They represent central but infrequent trends at the time. Twitter's popularity makes it an obvious example. Looking at the whole collection of publications (up to July 2015), Twitter has the highest degree, eigenvector and betweenness centrality scores. A snapshot of Twitter's ego network is highlighted in Figure 5 (2010 top right; 2015 bottom right; corresponding whole network on left).



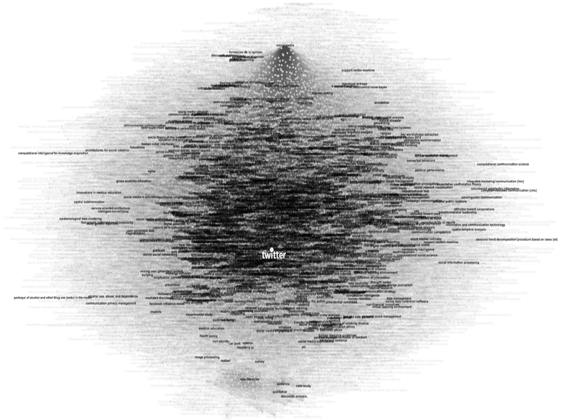
Twitter ego network 2010



SM network (Twitter-related nodes in dark) 2010



Twitter ego network 2015



SM network (Twitter-related nodes in dark) 2015

Figure 5. Twitter trend within Social Media (SM) 2010 & 2015

Twitter and the rest of the social media applications (Facebook, Youtube and e-Campaigning) are not the only examples. Most of the keywords/trends identified with high closeness and low degree have become disruptive trends in their own right in their respected groups:

- *Distributed Denial of Service* (DDoS - attacks launched from multiple connected devices disrupting intended users' access to services) which is one of the most significant current concerns for security professionals (Alomari et al., 2012; Zargar et al., 2013);
- *Fuzzy clustering* (a clustering method which allows data to belong to more than one clusters) with current applications in the fields of pattern recognition and image processing, parameter estimation to medical diagnostics, weather forecast and time series predication (Gao & Xie, 2010; Zhao, 2013);
- Ground-breaking *biological applications of polymers* (Blanazs et al., 2009; Jagur□Grodzinski, 2010);
- The idea of *programmable networks* which is gaining considerable momentum at present with the emergence of the radical concept of Software-defined Networking (SDN) (Xie et al., 2009; Nunes et al., 2014; Feamster et al., 2014);
- And *RFID technology* (Radio Frequency Identification) applications on supply-chain management, communications and bioinformatics (Ustundag & Tanyas, 2009; Zhu & Kurata, 2012; Ohashi et al., 2010).

An equally interesting combination is high betweenness paired with low closeness. This keyword would identify a trend both bridging and niche. However, there are no examples of such theme in any of the groups.

5. Discussion and implications for practice

Identifying potentially disruptive technologies is crucial to safeguarding competitive advantage by enabling stakeholders to assign resources in a manner that increases the chances of exploiting the disruption and/or mitigating the ensuing risks. This paper contributes to technology forecasting by proposing a new method for the identification of emergent technologies with disruptive potential by means of keyword network analysis. Here we discuss our findings in line with the paper's research objectives.

Seven major domains were identified in the field of disruptive technologies: 3D printing, big data, bitcoin, cloud technologies, internet of things, MOOCs and social media. Data from business (Gartner, Forrester, IDC, Frost & Sullivan and Ovum) and academic publications (Reuters' Web of Science) were collected on the identified technologies from onset to June 2015. A total of 61,308 business and 31,386 academic publications were considered. No temporal distribution differences were detected between business and academic publications on the disruptive technologies investigated. This promotes an understanding of the role of "timing" in publishing and represents an important finding for practitioners/companies researching disruptive trends. Providing there are no access barriers, information on emergent technologies appears in academic and business publications without significant temporal differences.

Social network analysis techniques were employed to investigate and demonstrate the thematic and temporal relationships in the academic publications and identify the nature of important and/or emerging trends.

Based on basic network metrics we found that thematic networks within the academic publications fit well the small-world type and share its properties. Most keywords are not neighbours but can be reached from every other by a small number of steps. Having small world properties means that links between themes and therefore information flows are not constrained to spatial or temporal proximity and there is no impediment to hub formation. As a result, there is an abundance of (identifiable) thematic hubs mediating thematic network flows. It is the existence and topological characteristics of these thematic hubs that enable further analysis and identification of potentially disruptive trends.

Six influential common subjects were identified by means of high degree, betweenness and closeness centrality scores. The corresponding keywords (*education/learning*, *internet, performance, privacy, quality,* and *security*) denote central popular themes which, while generic, are influential in all groups and reflect a level of alignment as an idea, or concept (here identified as a *disruptive trend*) reaches a level of maturity. Gauging the level of maturity of a disruptive technology helps to determine the stage of adoption and leverage it appropriately. Emerging technologies correspond to pioneer adopters and high transformational potential while mature technologies are associated with extensive adoption and moderate differentiation of competitive advantage. However, the identified common trends signify an indicator of maturity only. The actual state of maturity would need to be triangulated by other measures (e.g. sales).

Niche and potentially emerging trends within groups can be detected (but remain undistinguished) by means of eccentricity and farness scores. In smaller networks ($V < 500$) visualisation is an effective tool as it depicts these marginal trends, their thematic relationships and their corresponding neighbourhoods with relative precision and clarity.

Temporal comparison between developing and maturing networks detected distinctive thematic flows associated with disparity in centrality scores. Among the combinations, high closeness paired with low degree centrality is associated with disruptive trends within groups. Apart from the well-known and equally well-documented disruptive nature of the social media applications identified (Twitter, Facebook, Youtube and e-Campaigning), this method detected also the following trends which have since proved to be of disruptive nature: Distributed Denial of Service, fuzzy clustering, biological applications of polymers, programmable networks and RFID technology applications.

Not all keywords identified by high closeness and low degree centrality can be associated with disruptive trends, even after discarding non-domain specific keywords, those acting as modifiers/descriptors and those covering a broad area (e.g. communication system security/IoT, health cloud service/cloud, theoretical chemistry/3d printing, etc.). From the remaining keywords, there evidence suggesting that some may come to be identified as disruptive trends. The above findings were grouped in the framework depicted in Figure 6. The framework can be used for assessing the level of maturity of disruptive trends and for assisting the forecasting of emergent technologies with disruptive potential.

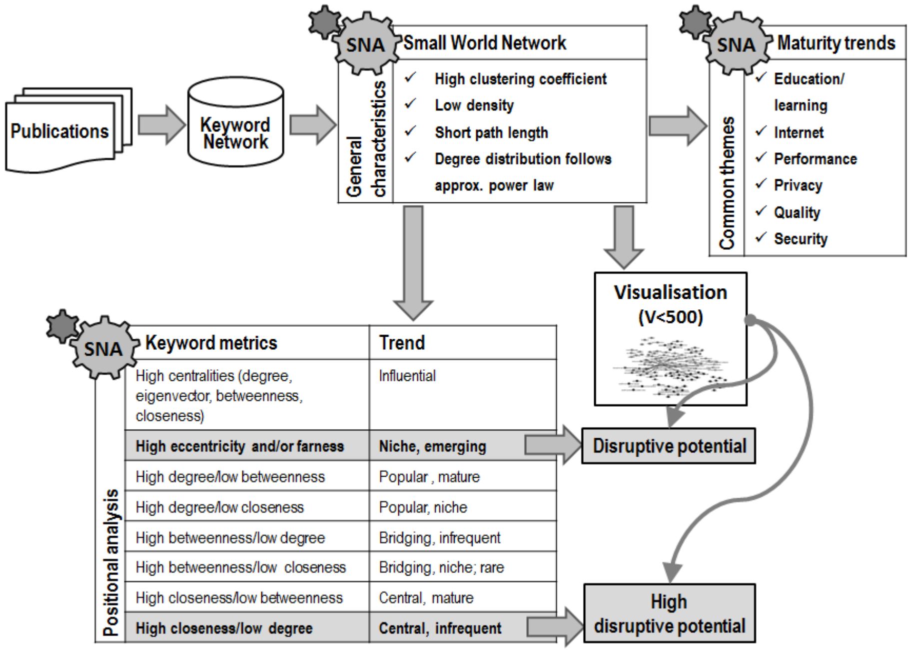


Figure 6. Potentially disruptive trend detection framework.

6. Conclusions and future work

The paper presents a new, literature-driven method for the forecasting of emerging technologies within disruptive domains. We collected data from academic publications on seven technologies identified as disruptive and adopted social network analysis techniques in order to investigate and demonstrate key thematic and temporal relationships among keywords. The research carried out shows that a keyword network approach is well suited to the analysis of the thematic, structural and temporal characteristics of disruptive technologies through their presence in the literature. Furthermore, there is evidence to suggest that the method can be used for the forecasting of potentially disruptive trends. Visualisation is possible and meaningful for smaller sets of data but becomes problematic when it comes to large networks. Based on our findings we proposed a framework for assisting this forecasting which is the contribution of this paper.

The theoretical implications are two-fold. The first relates to the application of social network analysis methods to the domain of disruptive technologies. The results are in line with similar research in other domains (Choi et al., 2011a; Choi et al., 2011a; Behara et al., 2014; Li et al., 2016). The second implication and main contribution of this paper, is the combination, ordering and systematic grouping of network analysis metrics to create discrete modules of reference which are then combined into the proposed framework. The framework's modularity enables the output of intermediate modules to be fed into exit components where the disruptive potential of emerging technologies is ascertained.

The shortcomings encountered can be grouped in two main categories:

1. Keyword networks: there are issues with syntax, misspellings and multiple uses of certain keywords (with identifiers and/or as part of composite topic). This may influence the positional

metrics of certain keywords appearing multiple times with various identifiers. The process is non-trivial and requires synthesis of multiple occurrences and re-positioning of edges. It is therefore difficult to automate and needs further investigation.

2. Business publications: keyword analysis needs to be extended to include business publications. While there is no time difference, academic articles are by nature tightly focused and specialised whereas business publications cover wider areas and may be better suited to include wider area thematic flows and influence considerations.

Future research will aim to validate the proposed framework. In order to do this we intend to test and evaluate the trajectory of current trends which have been identified as potentially disruptive by our method. Examples include the *alginate-hyaluronic acid scaffold* (3D printing), *proteomics* and *stacked auto-encoders* (Big Data), *organic computing* (IoT) and *blockchain* (bitcoin).

Further validation of the framework's robustness in identifying disruption and the overall suitability of keyword network analysis to literature-based forecasting is also being investigated. This will require the application of the method to different contexts and/or industries.

REFERENCES

Adner, R., Zemsky, P., (2001), Disruptive Technologies and the Emergence of Competition, Insead, Fontainebleau

Adner, R., (2002), When are technologies disruptive?. A demand-based view of the emergence of competition, Strategic Management Journal, 23 (2002), pp. 667–688

Alomari, E., Manickam, S., Gupta, B. B., Karuppayah, S., & Alfaris, R. (2012). Botnet-based distributed denial of service (DDoS) attacks on web servers: classification and art. arXiv preprint arXiv:1208.0403.

Amara, N., Landry, R., Becheikh, N., & Ouimet, M. (2008). Learning and novelty of innovation in established manufacturing SMEs. *Technovation*, *28*(7), 450-463.

Assink, M. (2006). Inhibitors of disruptive innovation capability: a conceptual model. European Journal of Innovation Management, 9(2), 215-233.

Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

Behara, S., Babbar, R. S., & Andrew Smart, P. (2014). Leadership in OM research: a social network analysis of European researchers. *International Journal of Operations & Production Management*, *34*(12), 1537-1563.

Bernard, J. G., & Gallupe, R. B. (2013). IT Industry Analysts: A Review and Two Research Agendas. Communications Of The Association For Information Systems, 33(1), 16.

Blanazs, A., Armes, S. P., & Ryan, A. J. (2009). Self□assembled block copolymer aggregates: from micelles to vesicles and their biological applications. Macromolecular Rapid Communications, 30(4□5), 267-277.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002), Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. SAGE Publications Limited.

Bulfin, S., Pangrazio, L., & Selwyn, N. (2014). Making 'MOOCs': The construction of a new digital higher education within news media discourse. *The International Review of Research in Open and Distributed Learning*, *15*(5).

Christensen, C. M. (1997), *The innovators dilemma: When new technologies cause great firms to fail*, Boston, MA: Harvard Business School Press.

Christensen, C. M. and Raynor, M. E. (2003) *The innovator's solution: Creating and sustaining successful growth*, Boston, MA: Harvard Business Press.

Christensen, C. M., Anthony, S. D. and Roth, E. A. (2004),*Seeing what's next: Using theories of innovation to predict industry change*, Boston, MA: Harvard Business School Press

Christensen, C. M., (2006), The ongoing process of building a theory of disruption, J. Prod. Innov. Manage., 23, pp. 39–55

Chiu, W. T., & Ho, Y. S. (2007). Bibliometric analysis of tsunami research. *Scientometrics*, *73*(1), 3-17.

Choi, S., Yoon, J., Kim, K., Lee, J. Y., & Kim, C. H. (2011a). SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells. *Scientometrics*, *88*(3), 863-883.

Choi, J., Yi, S., & Lee, K. C. (2011b). Analysis of keyword networks in MIS research and implications for predicting knowledge evolution. *Information & Management*, *48*(8), 371-381.

Choi, J., & Hwang, Y. S. (2014). Patent keyword network analysis for improving technology development efficiency. *Technological Forecasting and Social Change*, *83*, 170-182.

Danneels, E. (2004). Disruptive technology reconsidered: A critique and research agenda. *Journal of product innovation management*, *21*(4), 246-258.

Doloreux D, Shearmur R, (2012), *Collaboration, information and the geography of innovation in knowledge intensive business services. Journal of Economic Geography 2012;12:79-105.*

Dotsika F., Patrick K., (2013), Collaborative KM for SMEs: a framework evaluation study, Information Technology and People, Vol 26, No 4, pp.368-382

Drew, S. A. (2006). Building technology foresight: using scenarios to embrace innovation. *European Journal of Innovation Management*, *9*(3), 241-257.

Duvvuru, A., Kamarthi, S., & Sultornsanee, S. (2012, May). Undercovering research trends: Network analysis of keywords in scholarly articles. In *Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference on* (pp. 265-270). IEEE.

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, *17*, 81-91.

EUCommissionA, (2015), [online], Digital Single Market, http://ec.europa.eu/priorities/digital-single-market/docs/dsm-communication_en.pdf, last accessed (16 Feb 2016)

EUCommissionB, (2015), [online], Bitcoin Regulation, http://coinbrief.net/bitcoin-regulation-european-commission/, last accessed (16 Feb 2016)

Fagerberg, J. (2004). Innovation: a guide to the literature.

Faloutsos, M., Faloutsos, P., Faloutsos, C., 1999. On power-law relationships of the Internet topology. SIGCOMM 251–262

Feamster, N., Rexford, J., & Zegura, E. (2014). The road to SDN: an intellectual history of programmable networks. ACM SIGCOMM Computer Communication Review, 44(2), 87-98.

Fichman, R. G. (2004). Going beyond the dominant paradigm for information technology innovation research: Emerging concepts and methods. *Journal of the association for information systems*, *5*(8), 11.

Forbes, (2015), [online], 20 Digital Trends For 2015, http://www.forbes.com/sites/georgedeeb/2015/01/16/the-top-20-digital-trends-of-2015/ , last accessed (16 Feb 2016)

Frankelius, P., (2009), Questioning two myths in innovation literature, The Journal of High Technology Management Research, Volume 20, Issue 1, Pages 40-51

Frost & Sullivan by Marc Einstein, (2015), [online], 2015 Global and Japan Information and Communication Technology (ICT) Outlook, http://www.frost.com/sublib/display-report.do?id=9ABE-00-1A-00-00, last accessed (16 Feb 2016)

Ganguly, A., Nilchiani, R., Farr, J.V., (2008), Using a systems process to access the potential for disruptive technology, Conference on Systems Engineering Research, Los Angeles, CA, April 2008

Gao X. B., Xie W. X., "Advances in theory and applications of fuzzy clustering" Chinese Science Bulletin., vol. 45, no. 11, pp. 961-970, 2010

Gartner, (2015), [online], Top 10 Strategic Technology Trends for 2015, http://www.gartner.com/newsroom/id/2867917 , last accessed (16 Feb 2016)

Gasevic, D., Kovanovic, V., Joksimovic, S., & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *The International Review Of Research In Open And Distributed Learning*, *15*(5).

Garcia, R. and Calantone, R. (2002), A critical look at technological innovation typology and innovativeness terminology: a literature review. Journal of Product Innovation Management, 19: 110

Gertler M S, Levitte Y M, (2005), *Local nodes in global networks: the geography of knowledge flows in biotechnology innovation. Industry and Innovation 2005;12:487-507*

Grinberg, R. (2011). Bitcoin: an innovative alternative digital currency. *Hastings Science & Technology Law Journal*, *4*, 160.

Govindarajan, V., Kopalle, P. K. (2005). Disruptiveness of Innovations: Measurement and an Assessment of Reliability and Validity. Strategic Management Journal, 27 (2) (2005), pp. 189–199.

Govindarajan, V., Kopalle, P.K. (2006), The usefulness of measuring disruptiveness of innovations ex post in making ex ante predictions, Journal of Product Innovation Management, 23 (2006), pp. 12–18

Hamel, G. (1998). The challenge today: Changing the rules of the game. *Business Strategy Review*, *9*(2), 19-26.

Hang, C.C., Chen, J., Yu, D, (2011), An assessment framework for disruptive innovation Foresight, 13 (5) pp. 4–13.

Hausman, A., & Johnston, W. J. (2014). The role of innovation in driving the economy: Lessons from the global financial crisis. *Journal of Business Research*, *67*(1), 2720-2726.

Helal, S. (2015). Sustaining Computer's Impact and Adapting to Change.*IEEE Computer*, (1), 7-9.

Huber F, (2012), *Do clusters really matter for innovation practices in Information Technology? Questioning the significance of technological knowledge spillovers. Journal of Economic Geography 2012;12:107-126.*

Hughes, T., O'Regan, N., & Wornham, D. (2008). The credibility issue: Closing the academic/practitioner gap. *Strategic Change*, *17*(7□8), 215-233.

Hüsig, S., Hipp,C., Dowling. M., (2005),Analysing disruptive potential: the case of wireless local area network and mobile communications network companies, R&D Management, 35 (2005), pp. 17–35

Hyvonen, S., Touminen M.,(2006), Entrepreneurial innovations, market-driven intangibles and learning orientation: critical indicators for performance advantages in SMEs, International Journal of Management and Decision Making, 7 (6) (2006), pp. 643–660

Ikeler, A. (2007). The Under-Examined Public: Making Sense of Industry Analysts and Analyst Relations. *Journal of Promotion Management*, *13*(3-4), 233-260.

Jagur□Grodzinski, J. (2010). Polymeric gels and hydrogels for biomedical and pharmaceutical applications. Polymers for Advanced Technologies, 21(1), 27-47.

Jalonen, H. (2011). The uncertainty of innovation: a systematic review of the literature. *Journal of Management Research*, *4*(1).

Johnson, M. W., Christensen, C. M., & Kagermann, H. (2008). Reinventing your business model. *Harvard business review*, *86*(12), 57-68.

Keller, A., Husig, S (2009), Ex ante identification of disruptive innovations in the software industry applied to web applications: The case of Microsoft's vs Google's office applications Technol. Forecasting Soc. Change, 76 (2009), pp. 1044–1054

Kim, Y. G., Suh, J. H., & Park, S. C. (2008). Visualization of patent analysis for emerging technology. *Expert Systems with Applications*, *34*(3), 1804-1812.

Kim, B., Gazzola, G., Lee, J. M., Kim, D., Kim, K., & Jeong, M. K. (2014). Inter-cluster connectivity analysis for technology opportunity discovery. *Scientometrics*, *98*(3), 1811-1825.

Klenner, P., Husig, S, Dowling M.,Ex-ante evaluation of disruptive susceptibility in established value networks—when are markets ready for disruptive innovations? Res. Policy, 42 (2013), pp. 914–927

Kop, R., & Hill, A. (2008). Connectivism: Learning theory of the future or vestige of the past?. *The International Review of Research in Open and Distributed Learning*, *9*(3).

Kostoff, R. N., Boylan, R., & Simons, G. R. (2004). Disruptive technology roadmaps. *Technological Forecasting and Social Change*, *71*(1), 141-159.

Kostoff, R., (2004), Disruptive technology roadmaps, Technological Forecasting and Social Change, 71, pp. 141–159

Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2015). What public media reveals about MOOCs: A systematic analysis of news reports. *British Journal of Educational Technology*.

Latapy, M., Magnien, C., & Del Vecchio, N. (2008). Basic notions for the analysis of large two-mode networks. *Social Networks*, *30*(1), 31-48.

Lee, P. C., & Su, H. N. (2010). Investigating the structure of regional innovation system research through keyword co-occurrence and social network analysis. *Innovation*, *12*(1), 26-40.

Lee, Y., Kim, S. Y., Song, I., Park, Y., & Shin, J. (2014). Technology opportunity identification customized to the technological capability of SMEs through two-stage patent analysis. *Scientometrics*, *100*(1), 227-244.

Li, H., An, H., Wang, Y., Huang, J., & Gao, X. (2016). Evolutionary features of academic articles co-keyword network and keywords co-occurrence network: Based on two-mode affiliation network. *Physica A: Statistical Mechanics and its Applications*, *450*, 657-669.

Liang, Y., & Chen, J. (2011). Group Network Centrality Analysis of Blogs in Politics. *Communications in Information Science and Management Engineering*.

Linton, J. D. (2002). Forecasting the market diffusion of disruptive and discontinuous innovation. *Engineering Management, IEEE Transactions on*, *49*(4), 365-374.

Liping, J. (2013). Bitcoin: Theory, Practice and Effect. *Studies of International Finance*, *12*, 002.

Liu, Z., Yin, Y., Liu, W., & Dunford, M. (2015). Visualizing the intellectual structure and evolution of innovation systems research: a bibliometric analysis. *Scientometrics*, *103*(1), 135-158.

Markides, C. (2006). Disruptive innovation: In need of better theory*. *Journal of product innovation management*, *23*(1), 19-25.

McKinsey, 2015, [online], Disruptive technologies: Advances that will transform life, business, and the global economy, http://www.mckinsey.com/insights/business_technology/disruptive_technologies, last accessed (16 Feb 2016)

Moore, S. B., & Manring, S. L. (2009). Strategy development in small and medium sized enterprises for sustainability and increased value creation. *Journal of cleaner production*, *17*(2), 276-282.

Motter, A. E., de Moura, A. P. S., Lai, Y. C., & Dasgupta, P. (2002). Topology of the conceptual network of language. Science Phys Rev E, 65, 065102

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Consulted*, *1*(2012), 28.

NESTA, (2015), [online], Growing a digital social innovation ecosystem for Europe, http://www.nesta.org.uk/sites/default/files/dsireport.pdf , last accessed (16 Feb 2016)

Newman, M.E.J. 2003 The structure and function of complex networks. SIAM Review, 45, 2, 167–256.

Nunes, B., Mendonca, M., Nguyen, X. N., Obraczka, K., & Turletti, T. (2014). A survey of software-defined networking: Past, present, and future of programmable networks. Communications Surveys & Tutorials, IEEE, 16(3), 1617-1634.

ODI (2014), Open Disruptive Innovation Scheme, [online] https://ec.europa.eu/digital-agenda/en/open-disruptive-innovation-0 , accessed 10 August 2015

OECD, 2015, [online], Digital Economy Outlook OECD, http://ec.europa.eu/eurostat/documents/42577/3222224/Digital+economy+outlook+2015/dbdec3c6-ca38-432c-82f2-1e330d9d6a24, last accessed (16 Feb 2016)

Ohashi, K., Ota, S., Ohno-Machado, L., & Tanaka, H. (2010). Smart medical environment at the point of care: Auto-tracking clinical interventions at the bed side using RFID technology. Computers in Biology and Medicine, 40(6), 545-554.

Paap,J., Katz, R., (2004), Anticipating disruptive innovation, Research Technology Management, 47, pp. 13–22

Park, H., Kim, K., Choi, S., & Yoon, J. (2013). A patent intelligence system for strategic technology planning. *Expert Systems with Applications*, *40*(7), 2373-2390.

Pei-Chun Lee & Hsin-Ning Su (2010) Investigating the structure of regional innovation system research through keyword co-occurrence and social network analysis, Innovation, 12:1, 26-40

Phaal, R., Farrukh, C. J., & Probert, D. R. (2004). Technology roadmapping—a planning framework for evolution and revolution. *Technological Forecasting and Social Change*, *71*(1), 5-26.

Plassaras, N. A. (2013). Regulating digital currencies: bringing Bitcoin within the reach of IMF. *Chicago Journal of International Law*, *14*, 377.

Rafii,F., Kampa, P.,(2002), How to identify your enemies before they destroy you, Harvard Business Review, 80 (2002), pp. 115–123

Sainio, L.-M. , Puumalainen, K., (2007), Evaluating technology disruptiveness in a strategic corporate context: a case study, Technological Forecasting and Social Change, 74, pp. 1315–1333

Salmon, G., Gregory, J., Lokuge Dona, K., & Ross, B. (2015). Experiential online development for educators: The example of the Carpe Diem MOOC. *British Journal of Educational Technology*.

Schmidt, G., (2008), When is a disruptive innovation disruptive?, The Journal of Product Innovation Management, 25, pp. 347–369

Tappeiner G, Hauser C, Walde J, (2008), *Regional knowledge spillovers: fact or artefact? Research Policy 2008;37:861-874.*

Ustundag, A., & Tanyas, M. (2009). The impacts of radio frequency identification (RFID) technology on supply chain costs. Transportation Research Part E: Logistics and Transportation Review, 45(1), 29-38.

Vojak, B. A., & Chambers, F. A. (2004). Roadmapping disruptive technical threats and opportunities in complex, technology-based subsystems: the SAILS methodology. *Technological Forecasting and Social Change*, *71*(1), 121-139.

Walsh, S. T. (2004). Roadmapping a disruptive technology: A case study: The emerging microsystems and top-down nanosystems industry. *Technological Forecasting and Social Change*, *71*(1), 161-185.

Wang, P., & Ramiller, N. C. (2009). Community learning in information technology innovation. *MIS quarterly*, 709-734.

Wang, Xiao, Xinhu Zheng, Qingpeng Zhang, Tao Wang, and Dayong Shen. "Crowdsourcing in ITS: The state of the work and the networking." *IEEE Transactions on Intelligent Transportation Systems* 17, no. 6 (2016): 1596-1605.

Ware, M. (2009). Access by UK small and medium-sized enterprises to professional and academic literature. *Publishing Research Consortium, Bristol*.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.

Wu, C. C. (2016). Constructing a weighted keyword-based patent network approach to identify technological trends and evolution in a field of green energy: a case of biofuels. *Quality & Quantity*, *50*(1), 213-235.

Xie, L., Smith, P., Banfield, M., Leopold, H., Sterbenz, J. P., & Hutchison, D. (2009). Towards resilient networks using programmable networking technologies. In Active and Programmable Networks (pp. 83-95). Springer Berlin Heidelberg.

Yang, S., Han, R., Wolfram, D., & Zhao, Y. (2016). Visualizing the intellectual structure of information science (2006–2015): Introducing author keyword coupling analysis. *Journal of Informetrics*, *10*(1), 132-150.

Ye, C., Liu, D., Chen, N., & Lin, L. (2015, August). Mapping the topic evolution using citation-topic model and social network analysis. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on* (pp. 2648-2653). IEEE.

Yoon, B., Lee, S., & Lee, G. (2010). Development and application of a keyword-based knowledge map for effective R&D planning. *Scientometrics*, *85*(3), 803-820.

Young, G.K.*,* Jong, H.S. *and* Sang, C.P. *(*2008*).* Visualization of patent analysis for emerging technology*.* Expert Systems with Applications*,* 34*, pp.* 1804–1812*.*

Yu, D., & Hang, C. C. (2010). A reflective review of disruptive innovation theory. *International Journal of Management Reviews*, *12*(4), 435-452.

Yu, D., & Hang, C. C, (2011), Creating technology candidates for disruptive innovations: Generally applicable R&D strategies, Technovation, 31, pp. 401–410

Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. Communications Surveys & Tutorials, IEEE, 15(4), 2046-2069.

Zhao, F. (2013). Fuzzy clustering algorithms with self-tuning non-local spatial information for image segmentation. Neurocomputing, 106, 115-125.

Zhu, X., Mukhopadhyay, S. K., & Kurata, H. (2012). A review of RFID technology and its managerial applications in different industries. Journal of Engineering and Technology Management, 29(1), 152-167.