INFORMATION SYSTEMS RESEARCH REPORT

S Bisetty (2874956), FJ Van Wyk (24880159), C Twaddle (23560444)

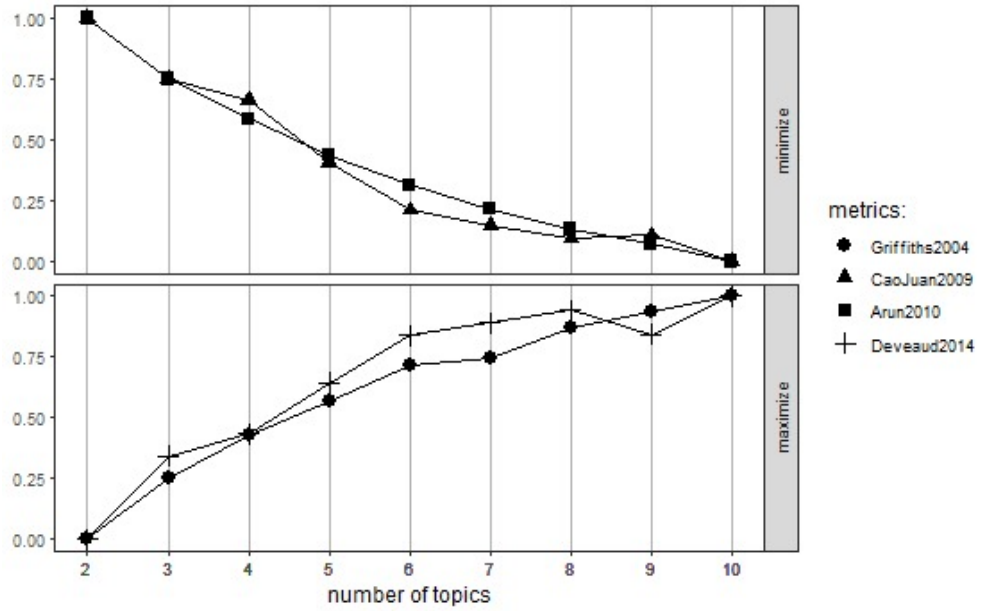
Introduction

Information System (IS) research landscape has evolved and changed over the past decade. This evolution is apparent in the shifting themes and topics that were the focus of research. By utilising text-based techniques, we uncovered prevalent trends, recurring themes and topic relationships within IS literature.

1. Most Common Topics and Themes in IS Research (2011-2020)

Our analysis centred on a dataset comprising information on a variety of IS research articles, such as the title, abstract, authors, year of publication and keywords, that have been published over the past decade. To ensure meaningful insights were extracted, the textual data had been pre-processed to eliminate noise and retain only relevant information. After which a Latent Dirichlet Allocation (LDA) model, a topic modelling algorithm, was applied to identify the most common themes and topics embedded within the corpus.

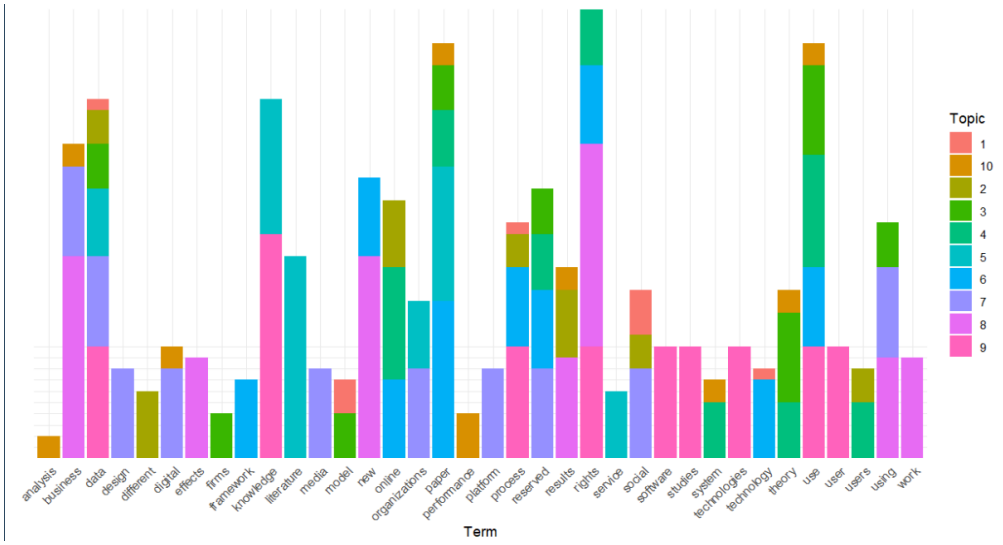
The K Value was calculated and selected after generating the following graph that finds the topic number:



Applying LDA revealed ten distinct topics that are characterised by a set of keywords that represent the focal points of discussion. These keywords can be found below:

Results:

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
social	process	social	model	model	data	technology	model	social	social
users	online	results	results	online	social	different	different	data	process
reserved	use	theory	paper	firms	model	use	using	theory	data
rights	reserved	use	system	paper	use	theory	users	online	online
organizations	literature	data	literature	service	knowledge	paper	literature	knowledge	paper
use	paper	online	technology	new	process	framework	reserved	paper	rights
business	design	digital	using	data	organizations	platform	social	media	reserved
results	effects	rights	new	business	business	new	work	rights	using
process	studies	user	data	use	rights	software	knowledge	technologies	knowledge
paper	theory	system	use	results	performance	performance	analysis	business	digital

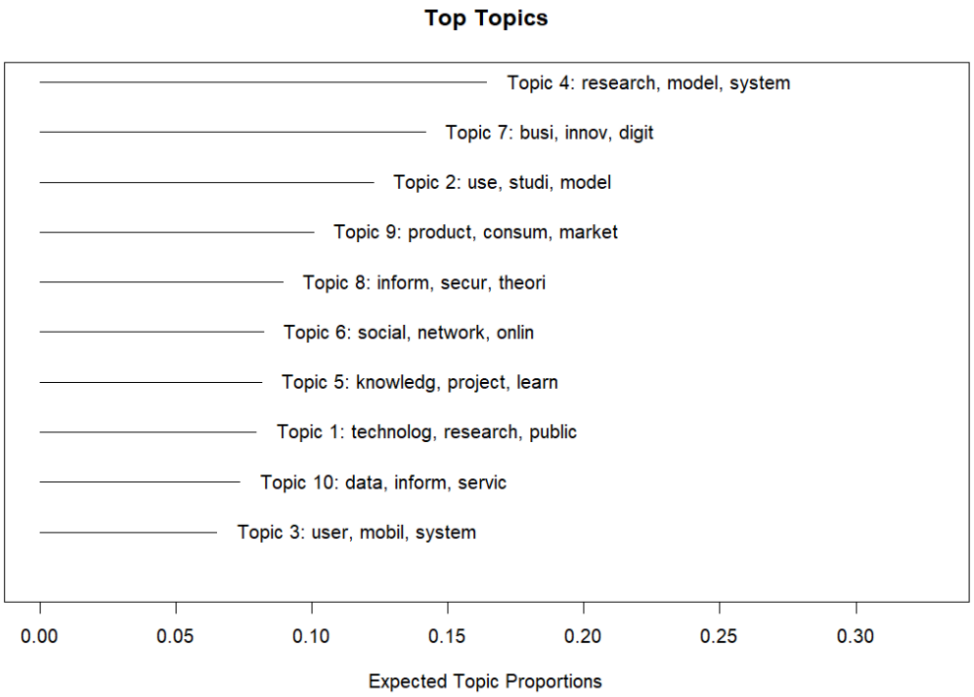


Topics such as data analytics, social impact, knowledge management and user behaviour are prominent, with terms such as “online”, “model” and “rights” offering valuable insight into the ongoing debates and areas being explored within the IS field.

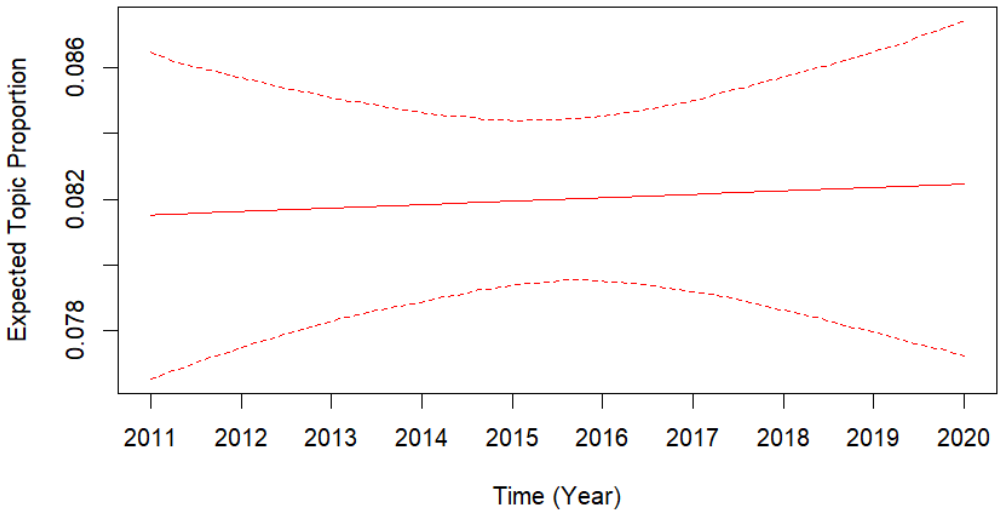
2. Evolution of IS Research Topics and Themes (2011-2020)

Using Structural Topic Modelling (STM) we were able to track the evolution of prevalent topics over the past decade.

An overview of these topics and their overall expected topic proportions can be seen below.



Topic 1: Technology and Digitization

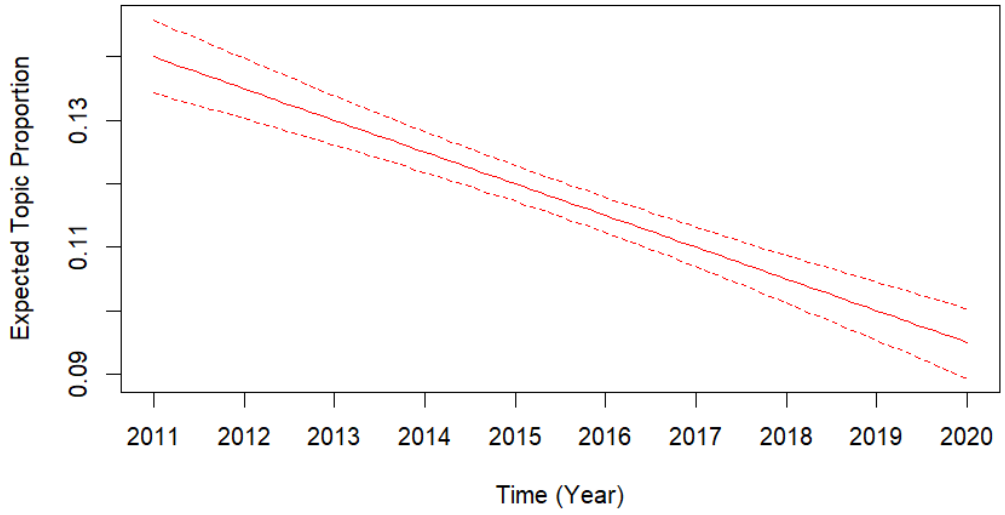


This topic entails discussions surrounding technology adoption, digital transformation and ICT initiatives in a variety of sectors including public services.

Keywords include technology, research, public, digit, develop and ICT are prevalent in this topic.

This topic has had a slight increase in prevalence from 2011 to 2020.

Topic 2: User Influence and Effects

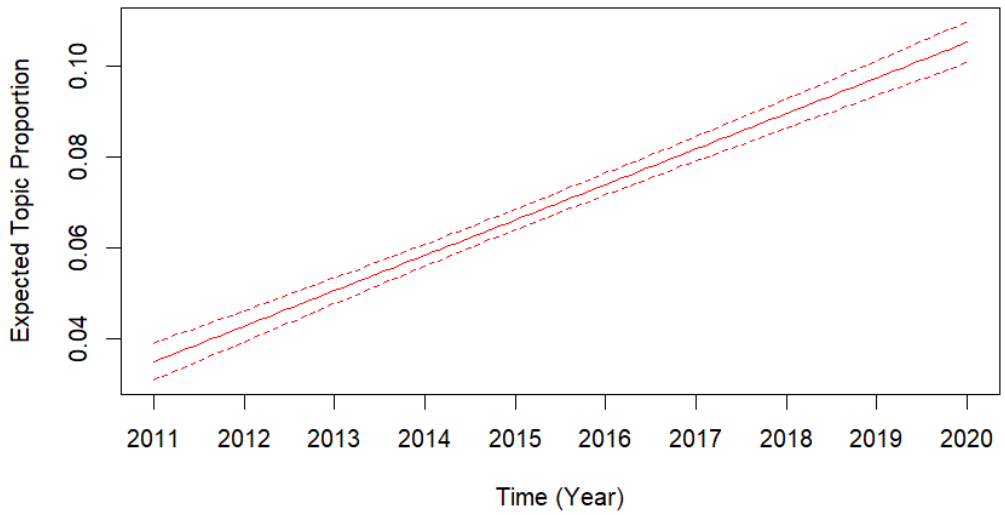


This topic focuses on understanding user behaviour, the factors influencing technology adoption and the effect of technology use on individuals and organisations.

Keywords include use, study, model, research, influence and factor.

This topic has had a steep decrease in prevalence from 2011 to 2020.

Topic 3: Mobile and App Design

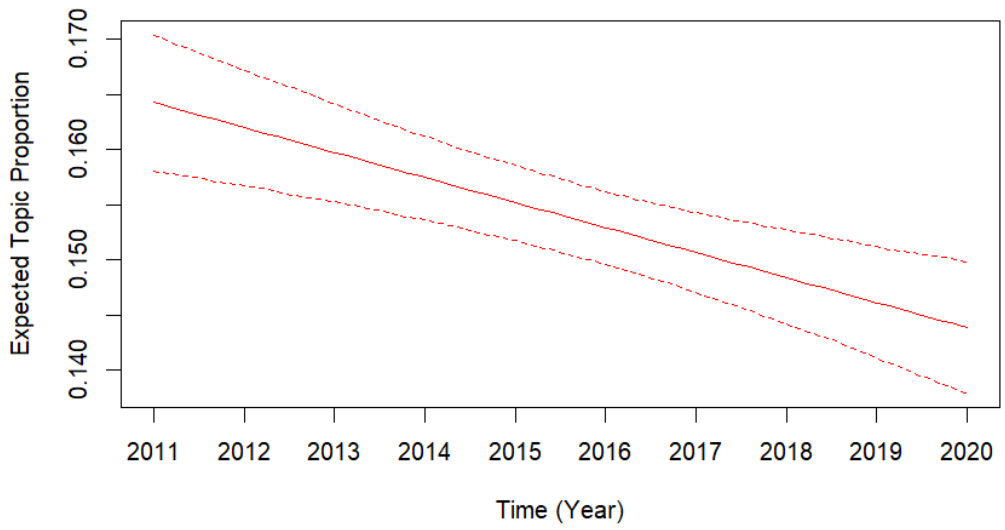


This topic explores mobile technology, app design, user experience and the impact of mobile applications on user behavior and interaction.

Keywords include user, mobile, system, design, experience and app.

This topic has had a steep increase in prevalence from 2011 to 2020.

Topic 4: Research Methodologies and Approaches

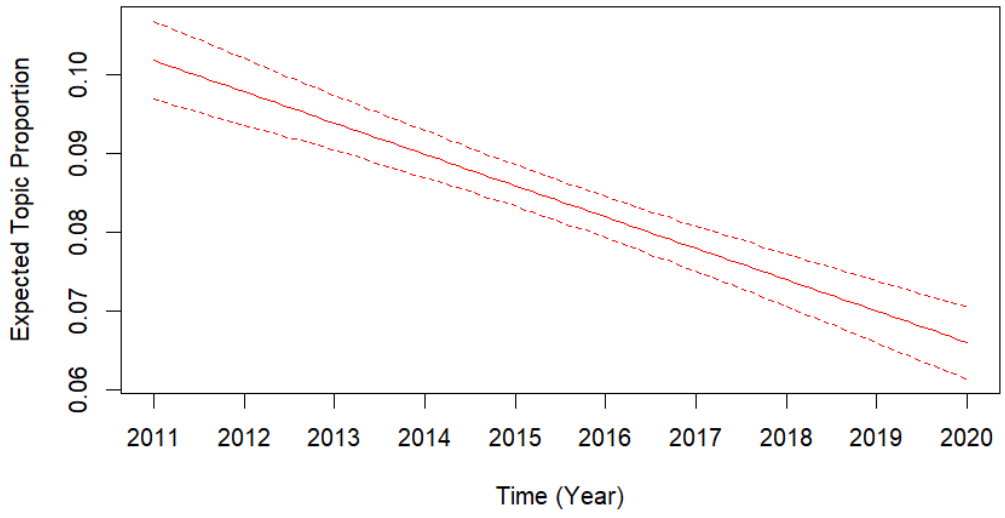


This topic revolves around research methodologies, system design approaches and process modeling techniques used in IS research.

Keywords include research, model, system, design, approach and process.

This topic has had a steep decrease in prevalence from 2011 to 2020.

Topic 5: Knowledge Management and Collaboration

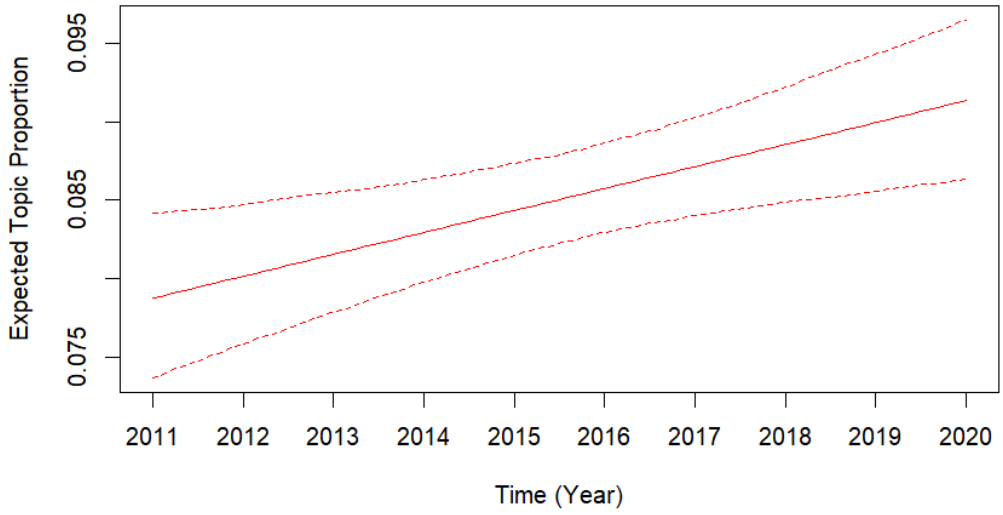


This topic emphasises knowledge management practices, collaborative project development, software engineering and team dynamics in IS projects.

Keywords include knowledge, project, learn, develop, software and team.

This topic has had a steep decrease in prevalence from 2011 to 2020.

Topic 6: Social Media and Online Communities

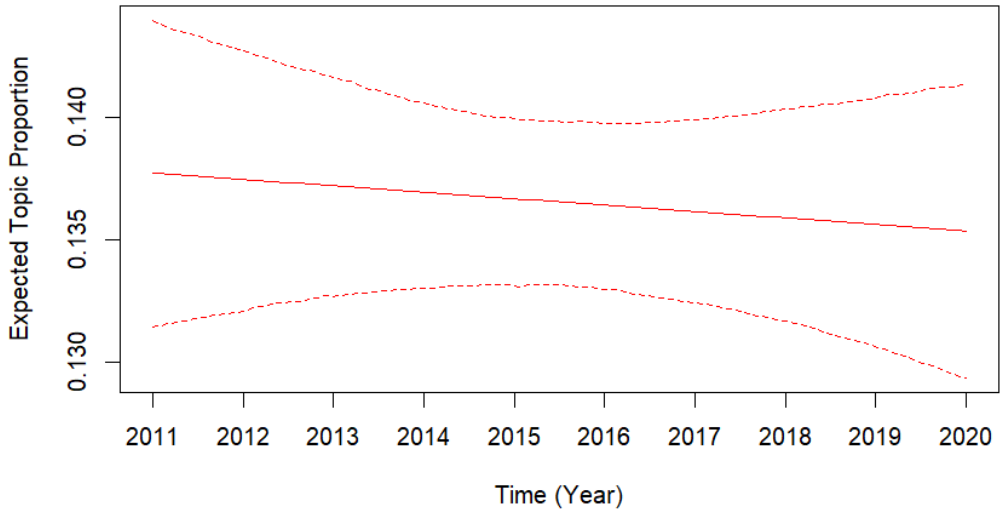


This topic focuses on social media platforms, online networks, user interactions and the spread of information on said platforms.

Keywords include social, network, online, media, user and information.

This topic has had a steep increase in prevalence from 2011 to 2020.

Topic 7: Business Innovation and Digital Transformation

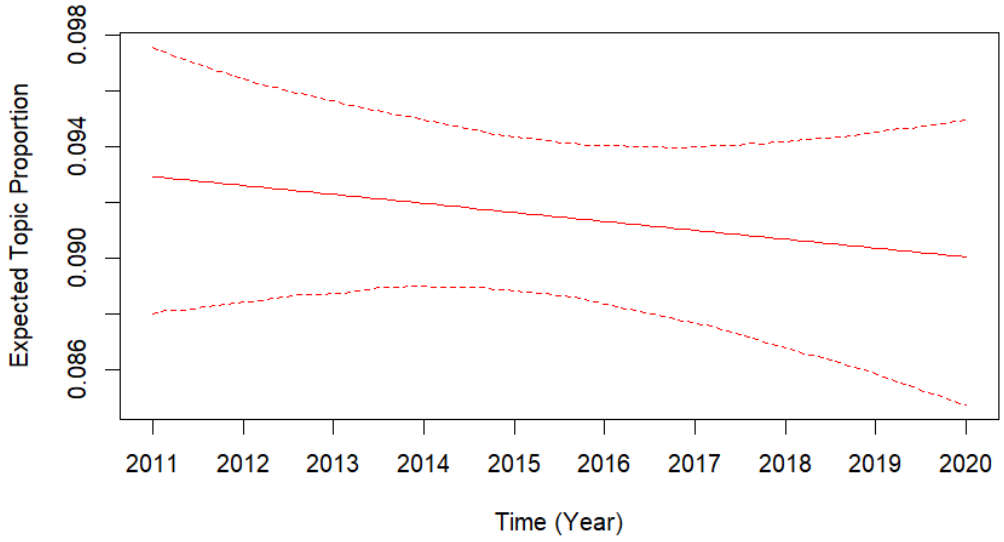


This topic explores business innovation, digital strategies and organisational capabilities and value creation through digital transformation initiatives.

Keywords include business, innovation, digital, firm, value and process.

This topic has had a gradual decrease in prevalence from 2011 to 2020.

Topic 8: Information Security and Risk Management

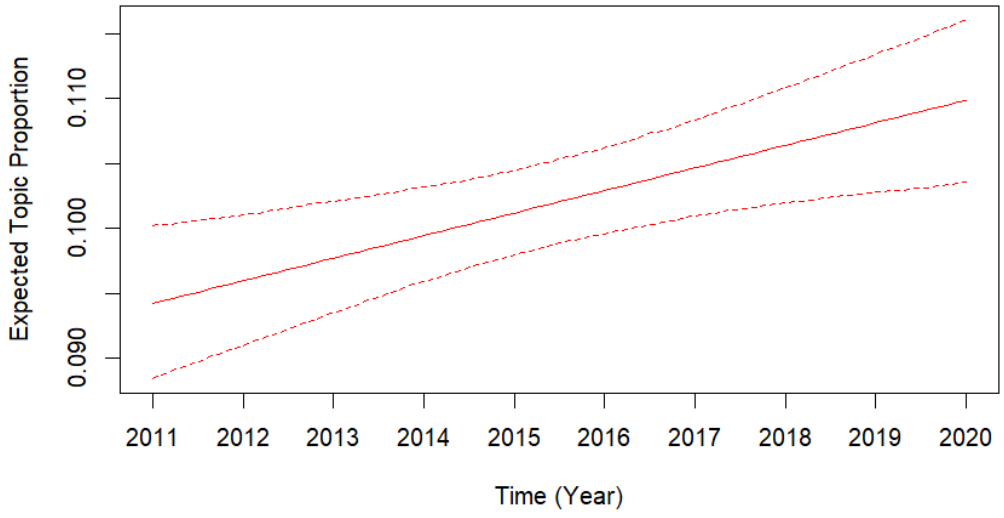


This topic addresses information security theories, practices, risk management strategies and organisational responses to security threats and challenges.

Keywords include information, security, practice, work and organisation.

This topic has had a gradual decrease in prevalence from 2011 to 2020.

Topic 9: E-commerce and Consumer Behaviour

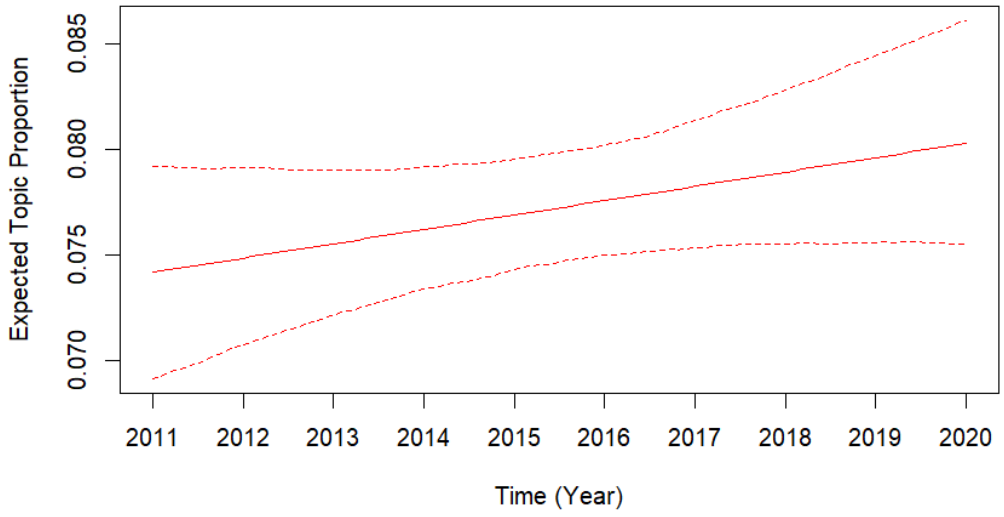


This topic examines e-commerce trends, consumer behaviour, market dynamics, online product reviews and the effects of online shopping on businesses and consumers.

Keywords include product, consumer, market, custom, online and effect.

This topic has had a steep increase in prevalence from 2011 to 2020.

Topic 10: Data Management and Healthcare



This topic covers data management practices, information services, risk assessment and healthcare informatics.

Keywords include data, information, service, system, risk and health.

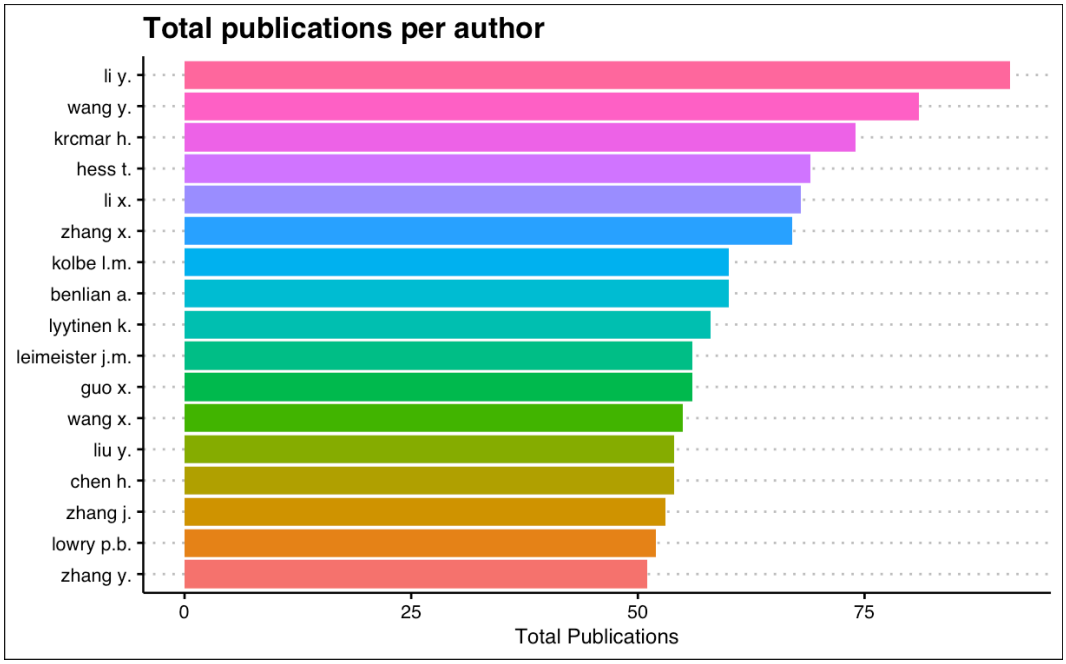
This topic has had a gradual increase in prevalence from 2011 to 2020.

3. Most Productive Authors and Journals in the Field of IS

Multiple metrics can be used to measure Author productivity. This includes

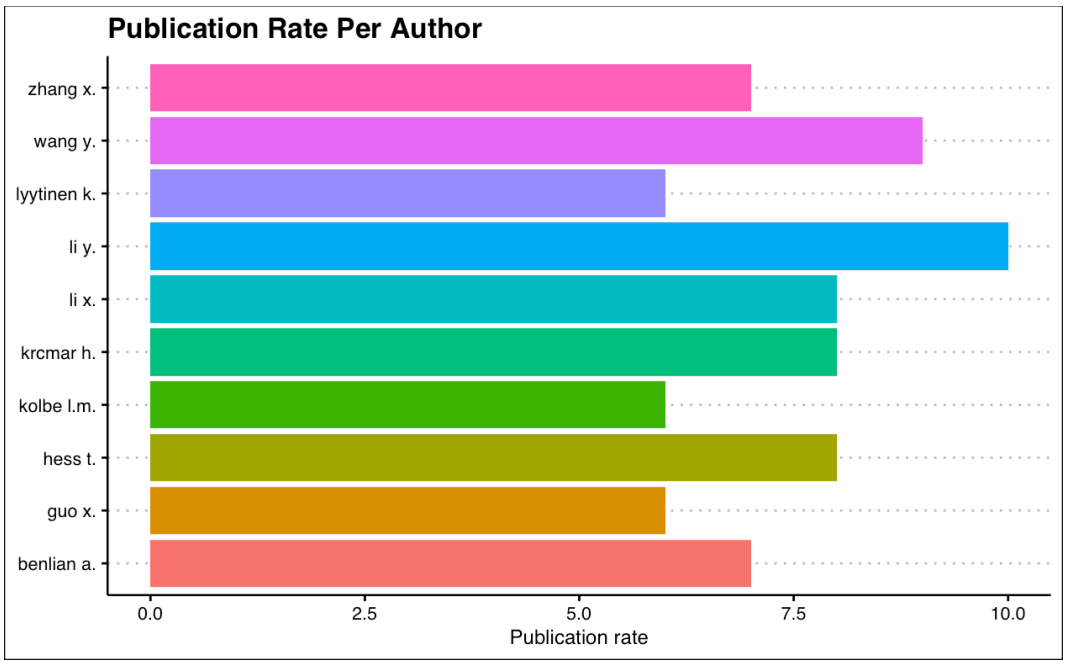
Publications per Author

The following graph displays the total count of publications written per author for the top authors (>50 publications). This indicates how many articles were published by each respective author. There are a total of 45,345 publications written by 16,693 authors. This allows for an average of 10.32 total publications per author. Yuan Li has published the most articles, having been involved in the publication of 91 articles.



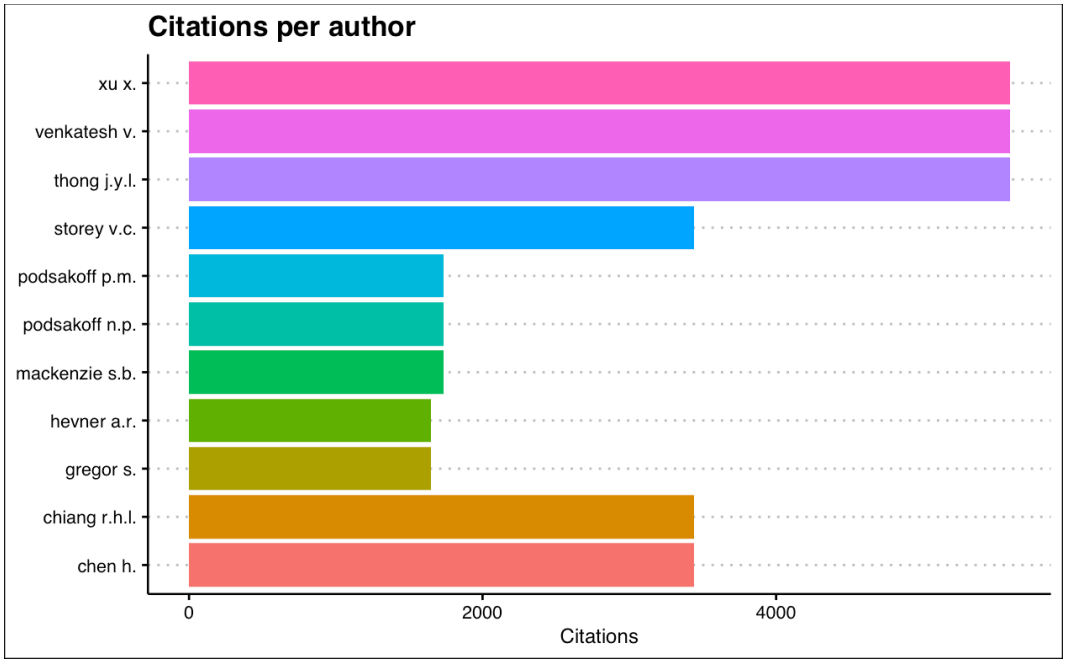
Publication Rate per Author

The publication rate for each author was calculated by determining for how long they were active (last publication - first publication), followed by dividing their total publications. This shows us how productive each author is/was throughout their academic career. The average publication rate was 1.513, with the highest being Yuan Li at ~10 publications per year.



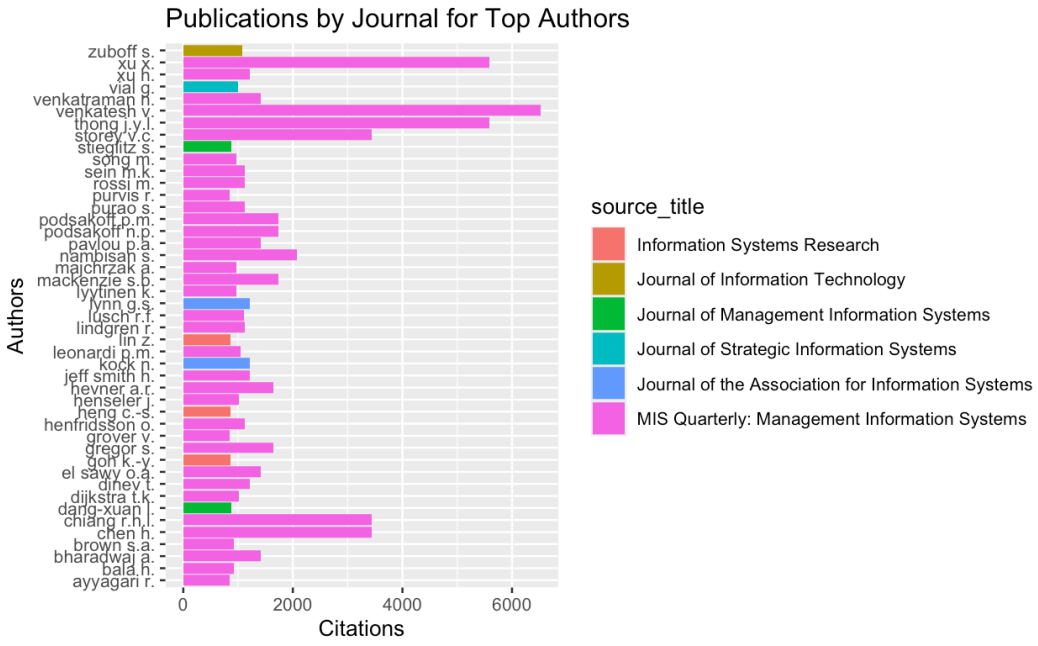
Citations Per author

The visualisation shows us which authors have been cited the most. With Venkatesh, V., Morris, M. G., and Davis, G. B. having the most citations. With the average author being cited ~ 21 times.



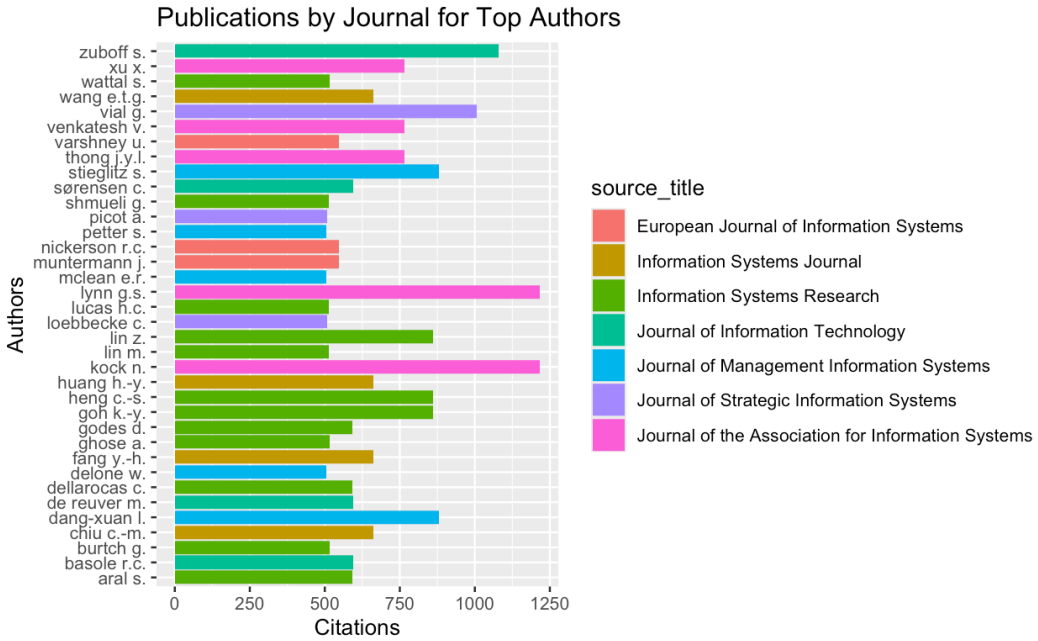
Publications by author for respective Journals

As seen in the graph below, The MIS Quarterly has garnered the most citations (74,687) despite only having the second highest publications (555). The Information Systems Research was published the most articles (567) but only received the second most citations (32,130). On average the various journals had a mean publications of 365, and the average total citations was 25,342. There is a strong positive correlation between total publications and total citations, with a positive correlation value of 0.772.



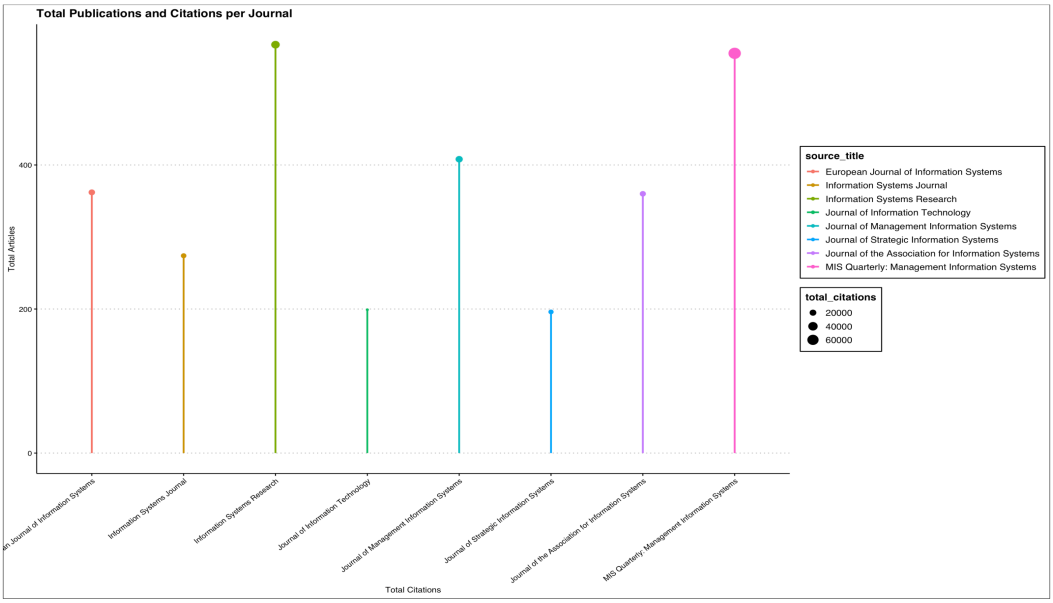
Excluding MIS Quarterly

When excluding the MIS Quarterly from the visualisation, it provides valuable insight into the number of publications, authors and citations for the remaining journals. This allows us to compare these journals without skewing the graph with data from the MIS Quarterly. Furthermore, this shows us the top authors for these journals respectively.



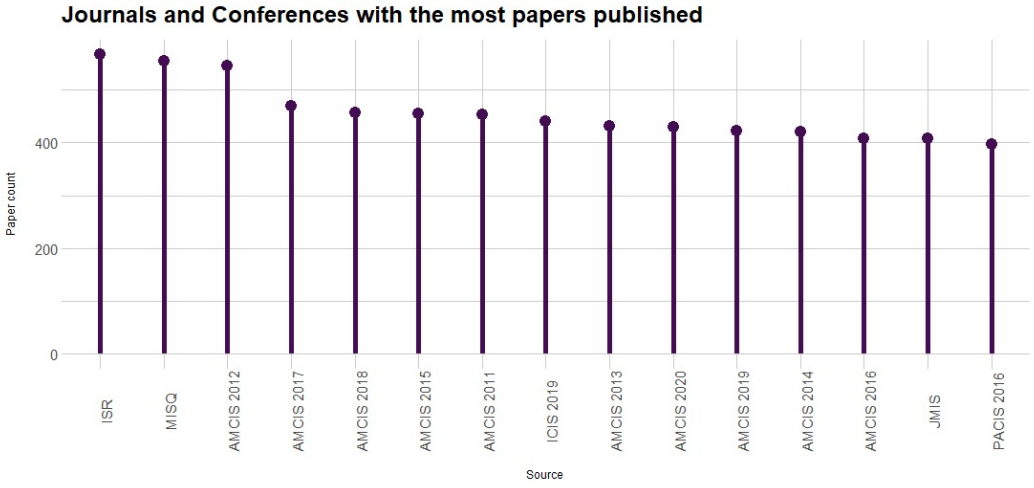
Total Publications and Citations per Journal

This graph displays the Total Publications per journal as well as the total citations. Again, showcasing that MIS Quarterly has been cited the most, and the Information Systems Research having the most publications.



Most frequent sources

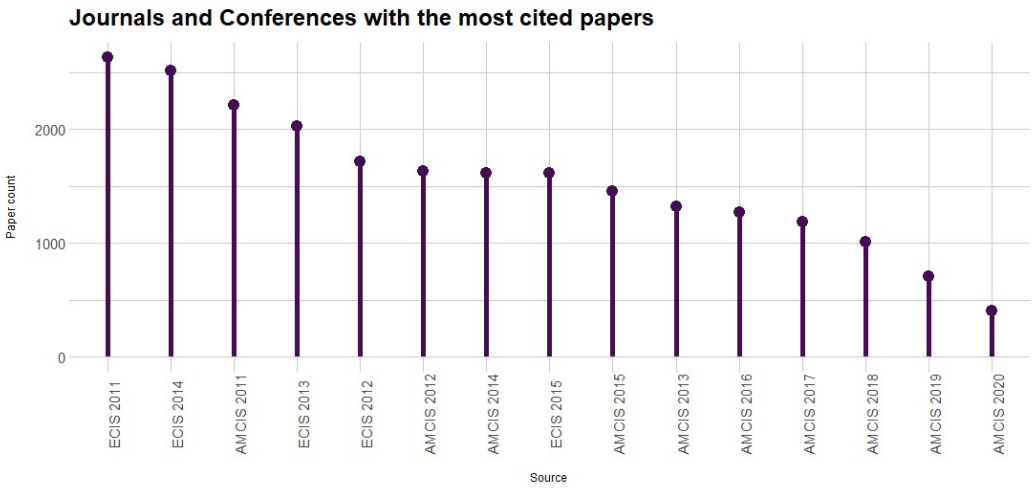
The following graph plots the top 15 most sources in the data. These are the journals and conferences that have the most number of papers published under there name. The graph shows that the yearly Americas Conference of Information Systems seems to produce the most papers.



Source Abbreviation	Full Source Name
ISR	Information Systems Research
JMIS	Journal of Management Information Systems
MISQ	MIS Quarterly: Management Information Systems
AMCIS 2020	26th Americas Conference on Information Systems, AMCIS 2020
AMCIS 2019	25th Americas Conference on Information Systems, AMCIS 2019
AMCIS 2018	Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018
AMCIS 2017	AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation
AMCIS 2016	AMCIS 2016: Surfing the IT Innovation Wave - 22nd Americas Conference on Information Systems
AMCIS 2015	2015 Americas Conference on Information Systems, AMCIS 2015
AMCIS 2014	20th Americas Conference on Information Systems, AMCIS 2014
AMCIS 2013	19th Americas Conference on Information Systems, AMCIS 2013 - Hyperconnected World: Anything, Anywhere, Anytime
AMCIS 2012	18th Americas Conference on Information Systems 2012, AMCIS 2012
AMCIS 2011	17th Americas Conference on Information Systems 2011, AMCIS 2011
ICIS 2019	40th International Conference on Information Systems, ICIS 2019
PACIS 2016	Pacific Asia Conference on Information Systems, PACIS 2016 - Proceedings

Most influential sources

The graph below shows the sources that produce the most cited papers. The total citations per article in the source was tallied to calculate the influence per source.



Source Abbreviation	Full Source Name
AMCIS 2020	26th Americas Conference on Information Systems, AMCIS 2020
AMCIS 2019	25th Americas Conference on Information Systems, AMCIS 2019
AMCIS 2018	Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018
AMCIS 2017	AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation
AMCIS 2016	AMCIS 2016: Surfing the IT Innovation Wave - 22nd Americas Conference on Information Systems
AMCIS 2015	2015 Americas Conference on Information Systems, AMCIS 2015
AMCIS 2014	20th Americas Conference on Information Systems, AMCIS 2014
AMCIS 2013	19th Americas Conference on Information Systems, AMCIS 2013 - Hyperconnected World: Anything, Anywhere, Anytime
AMCIS 2012	18th Americas Conference on Information Systems 2012, AMCIS 2012
AMCIS 2011	17th Americas Conference on Information Systems 2011, AMCIS 2011
ECIS 2015	23rd European Conference on Information Systems, ECIS 2015
ECIS 2014	ECIS 2014 Proceedings - 22nd European Conference on Information Systems
ECIS 2013	ECIS 2013 - Proceedings of the 21st European Conference on Information Systems
ECIS 2012	ECIS 2012 - Proceedings of the 20th European Conference on Information Systems
ECIS 2011	19th European Conference on Information Systems, ECIS 2011

4. Impact of IS research (Measured by Citation Counts)

Top cited overall

The table below shows the top 10 most cited papers over the ten years. The top most cited paper has 5 593 citations, and was published in 2012, showing it has had a clear impact since its release. The paper explores the theory of acceptance and use of information technology, which is often used to explore new developments within the information systems field. The other papers in the top 10 invlove a mixture of exploring the practical implications of technology as well as contributing to theoretical research.

Title	Authors	Year	Cited By
Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology	Venkatesh V., Thong J.Y.L., Xu X.	2012	5 593
Business intelligence and analytics: From big data to big impact	Chen H., Chiang R.H.L., Storey V.C.	2012	3 441
Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques	MacKenzie S.B., Podsakoff P.M., Podsakoff N.P.	2011	1 732
Positioning and presenting design science research for maximum impact	Gregor S., Hevner A.R.	2013	1 651
Digital business strategy: Toward a next generation of insights	Bharadwaj A., El Sawy O.A., Pavlou P.A., Venkatraman N.	2013	1 418
Information privacy research: An interdisciplinary review	Jeff Smith H., Dinev T., Xu H.	2011	1 222
Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations	Kock N., Lynn G.S.	2012	1 217
Action design research	Sein M.K., Henfridsson O., Purao S., Rossi M., Lindgren R.	2011	1 119
Service innovation: A service-dominant logic perspective	Lusch R.F., Nambisan S.	2015	1 106
Big other: Surveillance capitalism and the prospects of an information civilization	Zuboff S.	2015	1 080

Top cited papers for each year

The table below displays the paper from each year with the most citations. This table shows that generally older papers tend to have more citations than newer ones, but this can be attributed to them being around for longer. Their is an overall trend of research about how technology and people interact, with a focus on digitisation in more recent years.

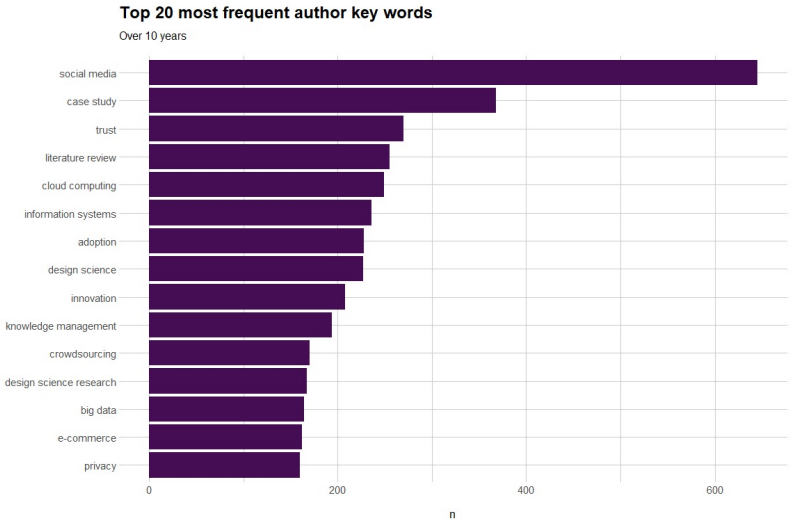
Title	Authors	Year	Cited By
An affordance perspective of team collaboration and enforced working from home during COVID-19	Waizenegger L., McKenna B., Cai W., Bendz T.	2020	208
Understanding digital transformation: A review and a research agenda	Vial G.	2019	1007
The digital platform: A research agenda	De Reuver M., Sørensen C., Basole R.C.	2018	594
Digital innovation management: Reinventing innovation management research in a digital world	Nambisan S., Lyytinen K., Majchrzak A., Song M.	2017	971
Unified theory of acceptance and use of technology: A synthesis and the road ahead	Venkatesh V., Thong J.Y.L., Xu X.	2016	766
Service innovation: A service-dominant logic perspective	Lusch R.F., Nambisan S.	2015	1106
Understanding customers' repeat purchase intentions in B2C e-commerce: The roles of utilitarian value, hedonic value and perceived risk	Chiu C.-M., Wang E.T.G., Fang Y.-H., Huang H.-Y.	2014	662
Positioning and presenting design science research for maximum impact	Gregor S., Hevner A.R.	2013	1651
Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology	Venkatesh V., Thong J.Y.L., Xu X.	2012	5593
Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques	MacKenzie S.B., Podsakoff P.M., Podsakoff N.P.	2011	1732

5. Most Frequent Keywords in IS and How They Have Changed Over Time

There are two type of keywords that will be analysed from the data source: author keywords and index keywords. Author keywords are the terms that the author chooses to represent their paper. Index keywords are terms that terms that are chosen by indexers to categorise the paper witin the database and aid in database organisation. In this section, we will compare the two types amongst them and against each other.

Author Keywords

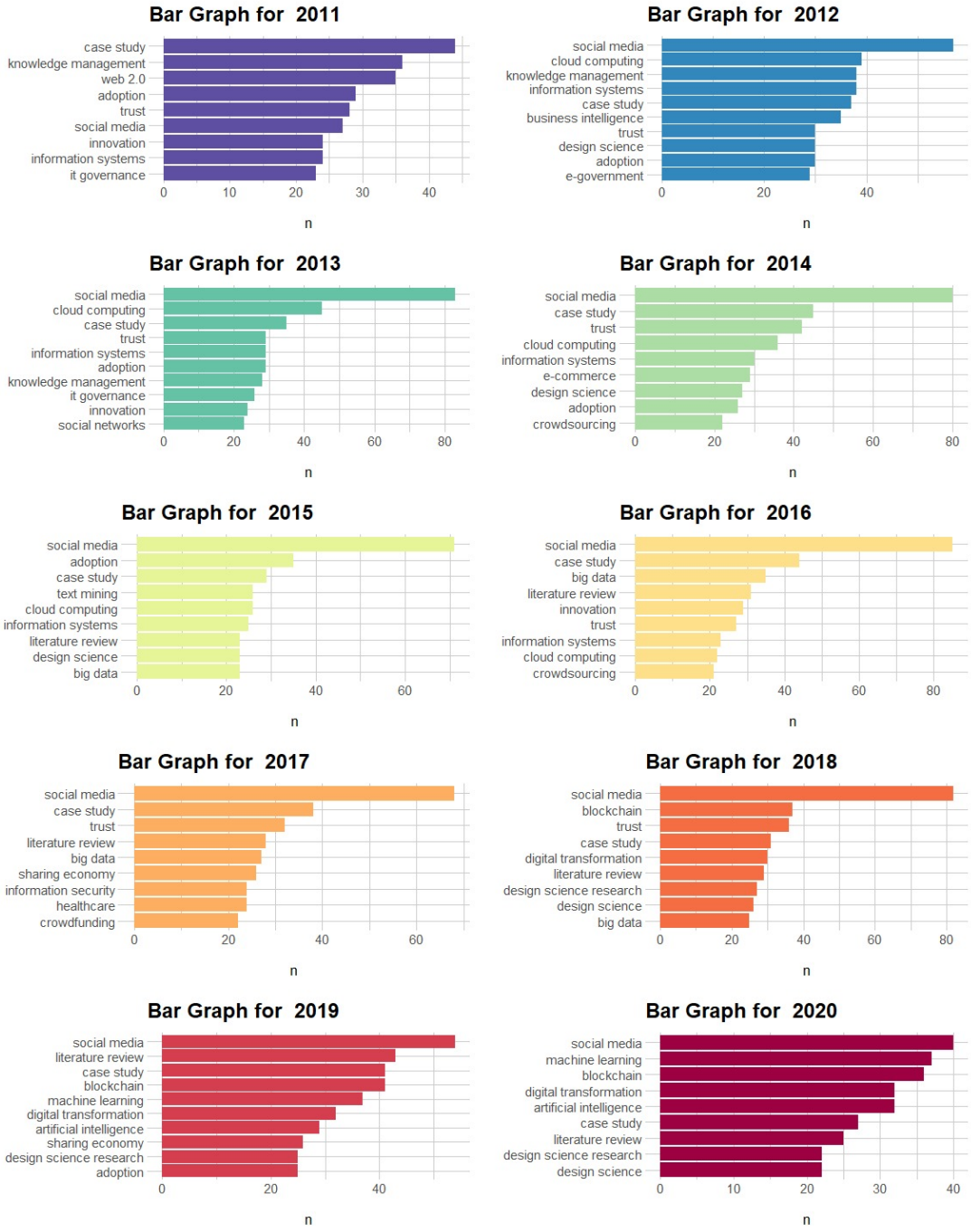
Over 10 years



World Cloud of Author Keywords

The above graph and word cloud shows the most frequent author keywords in IS ver 10 years. The clear leader is social media and this can be attributed to the increased in interest in social media itself and social media analysis throughout the 2010s.The word cloud visualises the keywords in a way that is easier to see which were most prevalent over the years, as well as showing some of the other keywords studied, but not as frequently.

By year



The above graphs show the top author keywords for each year. Social media has stayed the most popular across all years, except for 2011, contributing the fact that it is also the most frequent keyword over all. Each year the keywords shift to include more topics that were prevalent in that year, from knowledge management and cloud computing in 2011 and 2012, to machine learning, blockchain and artificial intelligence in 2019 and 2020. This shows the IS field is constantly evolving in the research that is done and the topics that are covered.

Open Access vs Limited Access

The above word cloud compared the keywords that are most found in open access papers versus papers with limited access. The graph shows that many of the case studies and research methods are found in open access papers, whereas the specific trends, such as cloud computing, blockchain and social media are more often found in limited access papers.

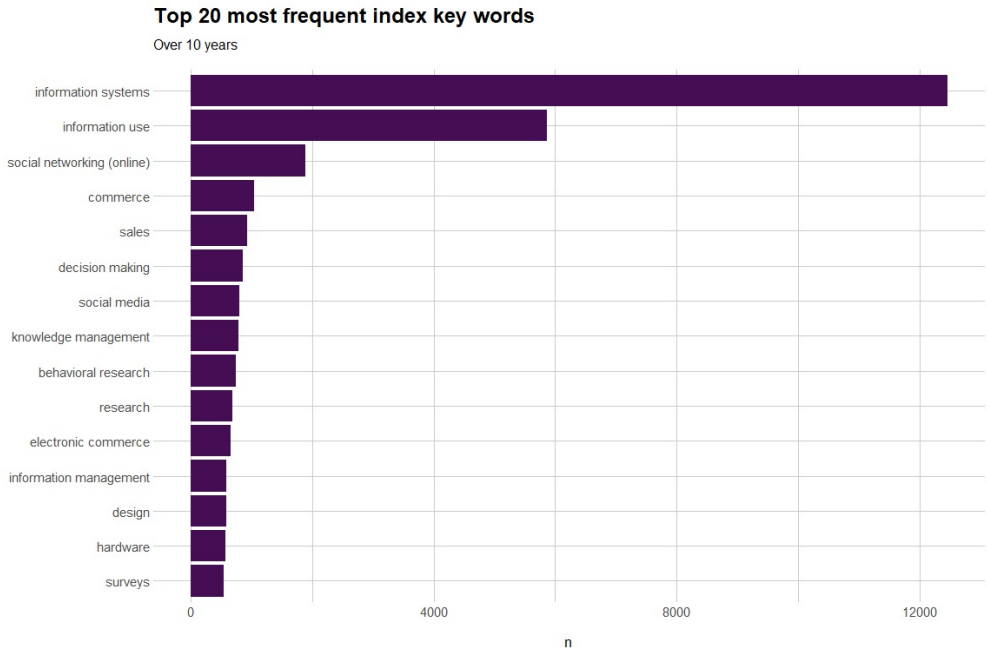
Conference Papers vs Journal Papers

The above word cloud compared the keywords that are most found in papers publicized in a journal versus papers published at a conference. The graph shows that many of the papers related to theory and business and economics related problems are published in journals, whereas the more general topics of scoial media and big data are found more often in conference papers.

Index Key Phrases

This section explores the index keywords in the same way author keywords were explored.

Over 10 years

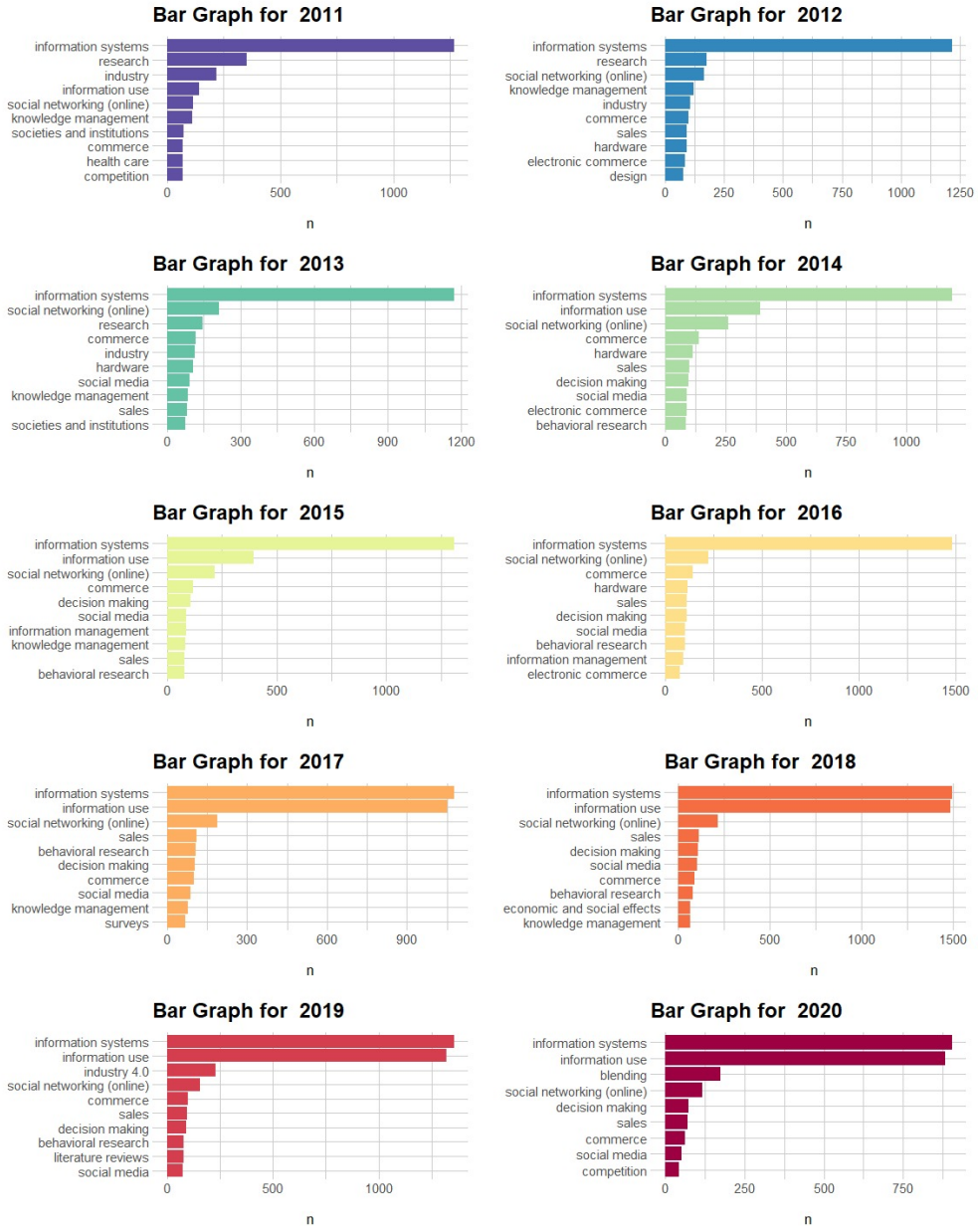


World Cloud of Index Keywords

The above graph and word cloud shows the most frequent index keywords in IS over 10 years. The clear leader is information systems and this is because these key words are used more for paper classification, and all of them fall into the category of information systems. Social media and social networking still feature quite highly in these keywords. The index

keyword word cloud shows that there is a broader variety in index keywords, but there is a clear most frequent few and many only briefly mentioned.

By Year

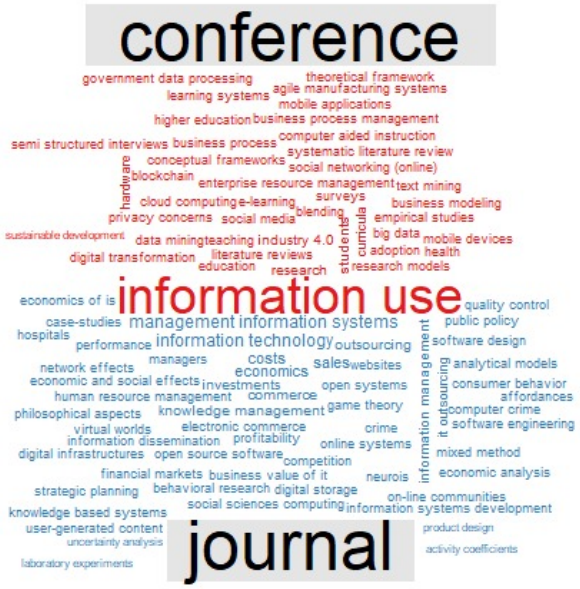


The above graphs show the top index keywords for each year. Information systems has stayed the most popular across all years, because of the reason discussed above. Each year the second most popular key word often shifts between information use and social networking. This graph also highlights the difference in the yearly author keywords and index keywords, as these graphs have more generic keywords that the author keywords.

Open Access vs Limited Access

The above word cloud compared the keywords that are most found in open access papers versus papers with limited access. The figure shows that there is not as clear a difference between the index keywords for each type, versus the author keywords for each type. There is not as much variation in size in each type of paper. Open access seems to feature more business related research, shown through keywords such as management information systems, cost investments, and economics. Limited access papers once again seem to focus on trends, such as social networking and e-learning.

Conference Papers vs Journal Papers

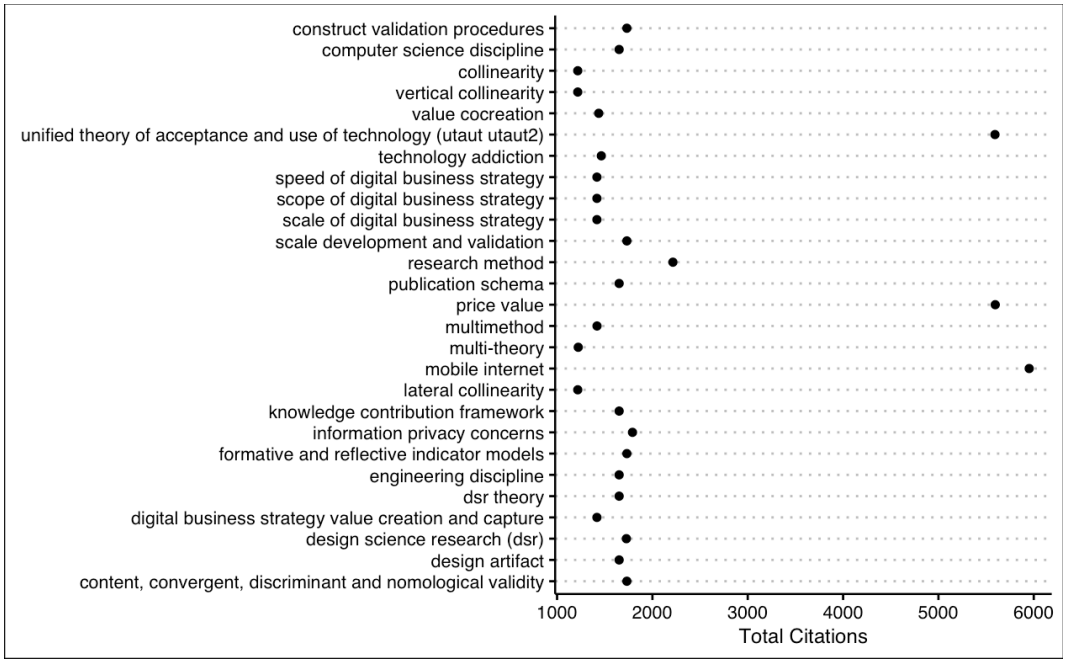


The above word cloud compared the keywords that are most found in papers publicized in a journal versus papers published at a conference. The figure shows a similar trend as the open access vs limited access word cloud, in that there is not as clear a difference in the keywords that are featured in each type of paper.

6. Relationship Between Specific Topics and Themes and Citation Counts

Author Keywords Citations Counts

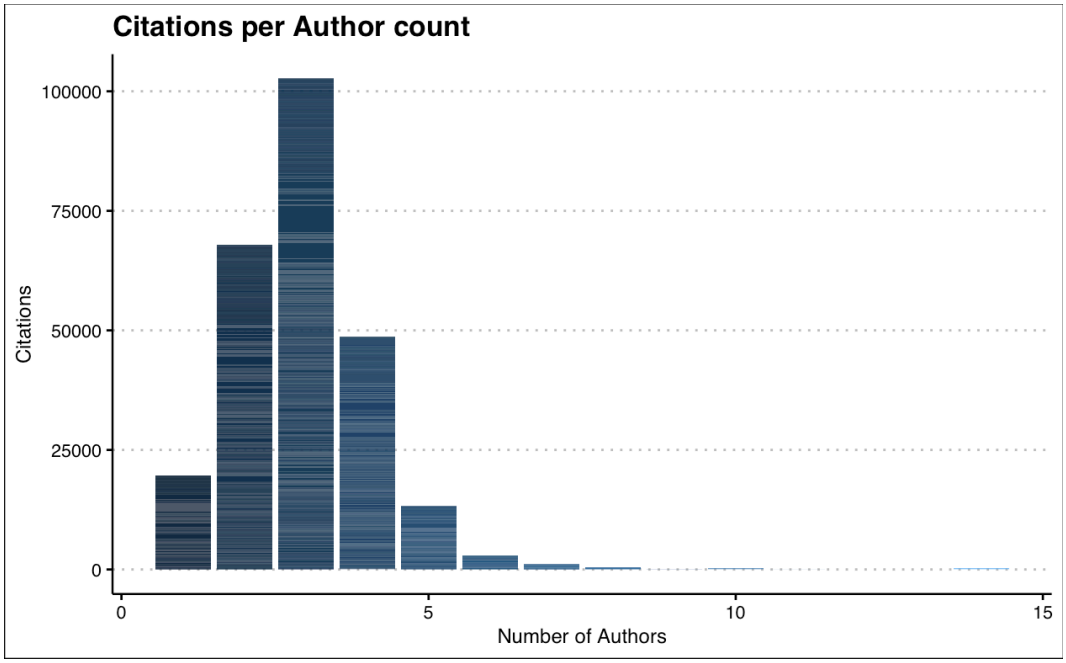
The graph below displays the author keywords from articles which have received more than 1500 citations. Here we can see 3 key words which appears and has been cited dramatically more than others. Mobile Internet as a key word was the most prevalent (~6,000), highlighting its increased relevance in modern society. The Unified theory of acceptance and use of technology (5,400) along with price value (5,300) were also outliers.



7. Relationship Between Authorship Patterns and Research Impact in IS Research

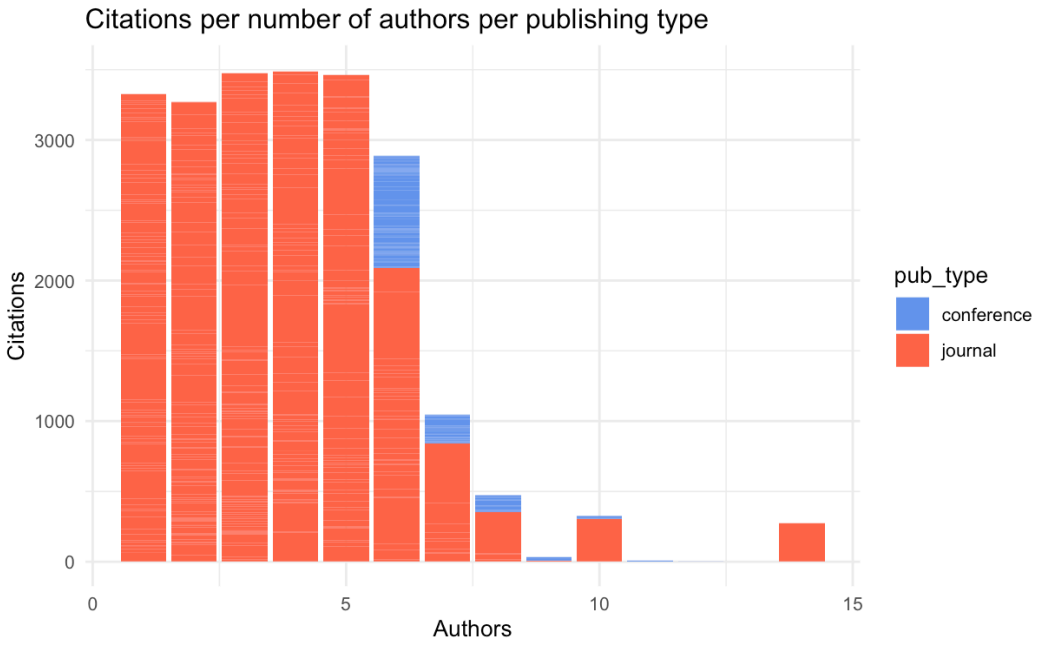
Citations per Author count

The following visualisation displays the amount of authors involved in a publication, along with how many times the publications were cited. This shows that multi-authorship publications tend to be cited more commonly, with publications with 3 authors being cited the most. However, this correlation between authorship and citation does not necessarily result in direct causation. With results being heavily skewed by publications such as the Unified Theory of Acceptance, written by Venkatesh, V., Morris, M. G., and Davis, G. B.



Citations per Publishing type

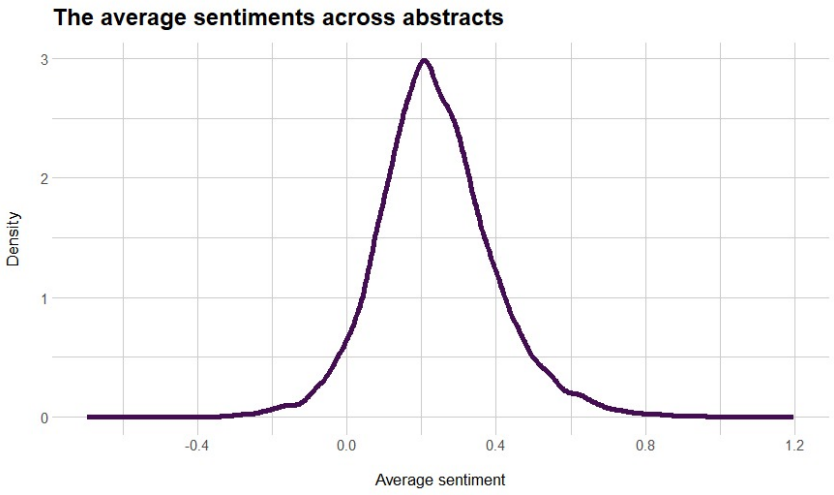
The following graph displays the citations of publications from both journals and conferences. It shows that publications in journals tend to be more commonly cited when compared to conference publications.

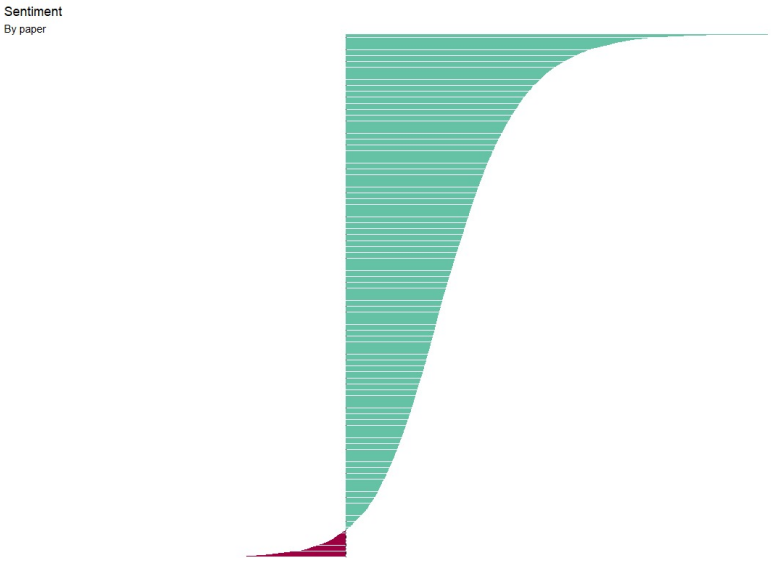

Citations per number of authors per publishing type
Citations
3000
2000
1000
0
0
5
10
15
Authors
pub_type
conference
journal

8. Tonality of IS Abstracts by Topic

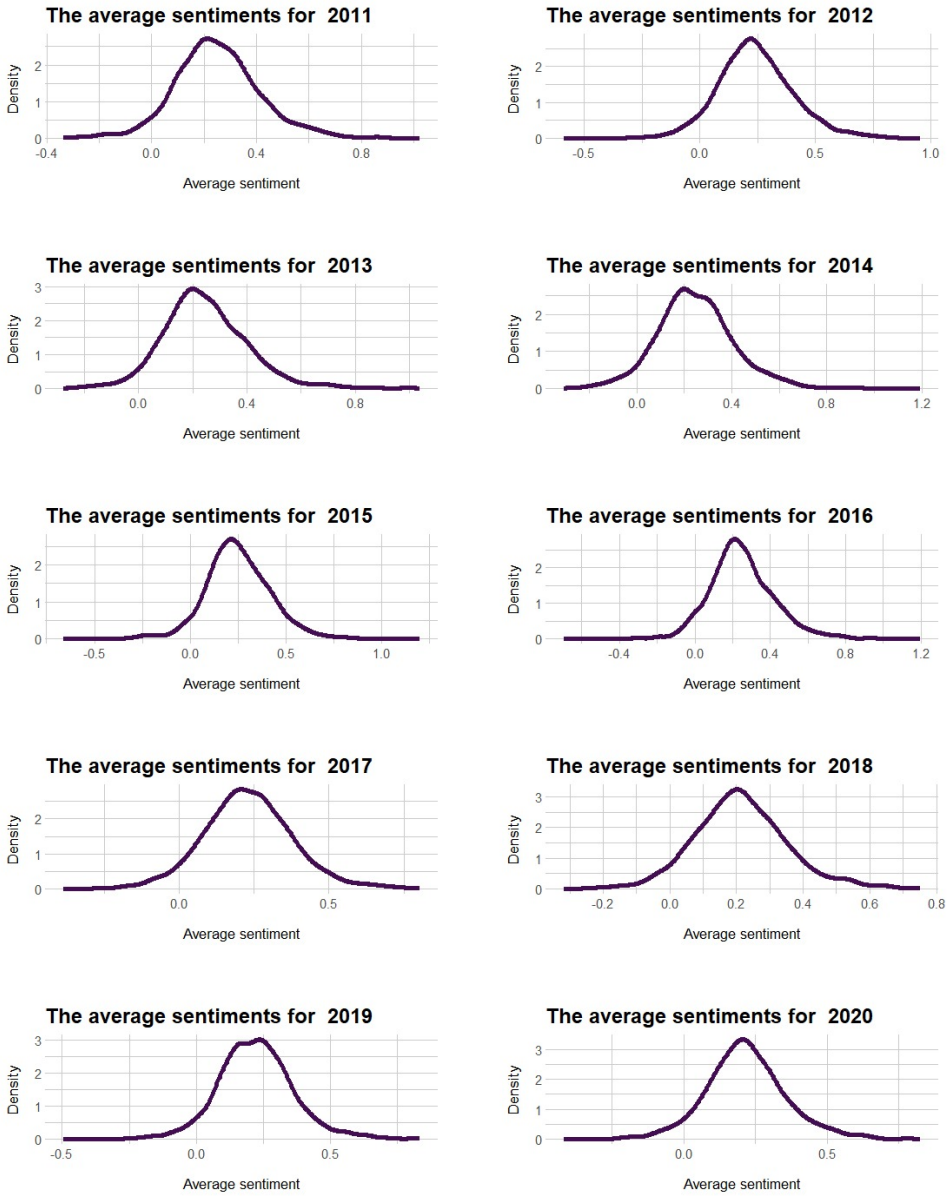
Sentiment refers to the emotional tone behind some kind of text. Sentiment analysis is the process of identifying the tone of text to identify a dominant attitude. The following graphs are related to investigating the sentiment for each topic identified through LDA.

Abstract and sentiments



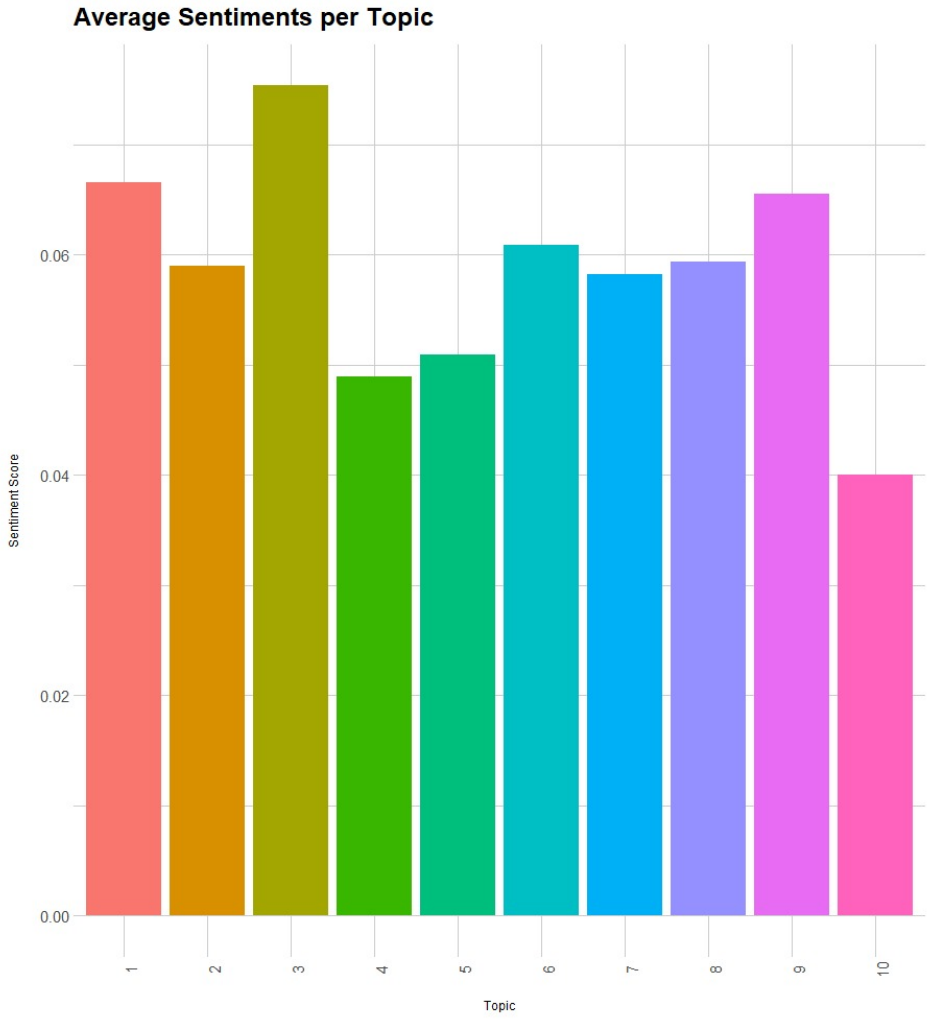


The above two graphs represent the average sentiment for each paper based on its abstract. The first graph shows that most of the papers have a positive sentiment of around 0,2. The send graphs shows that the majority of the papers have a postie average sentiment, the negative papers are a small proportion of the total.



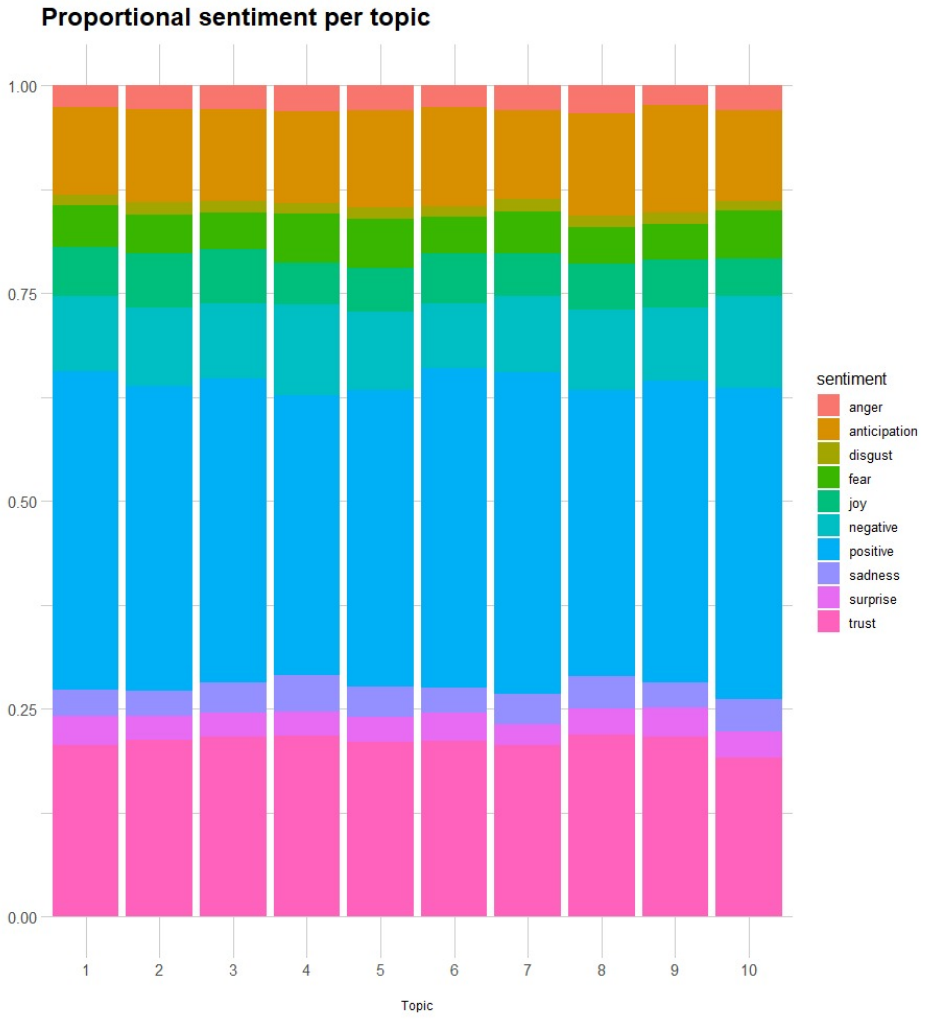
The above graph shows any changes in tonality across abstracts for each year. All the years differ slightly, but are all positive on average.

Average per topic



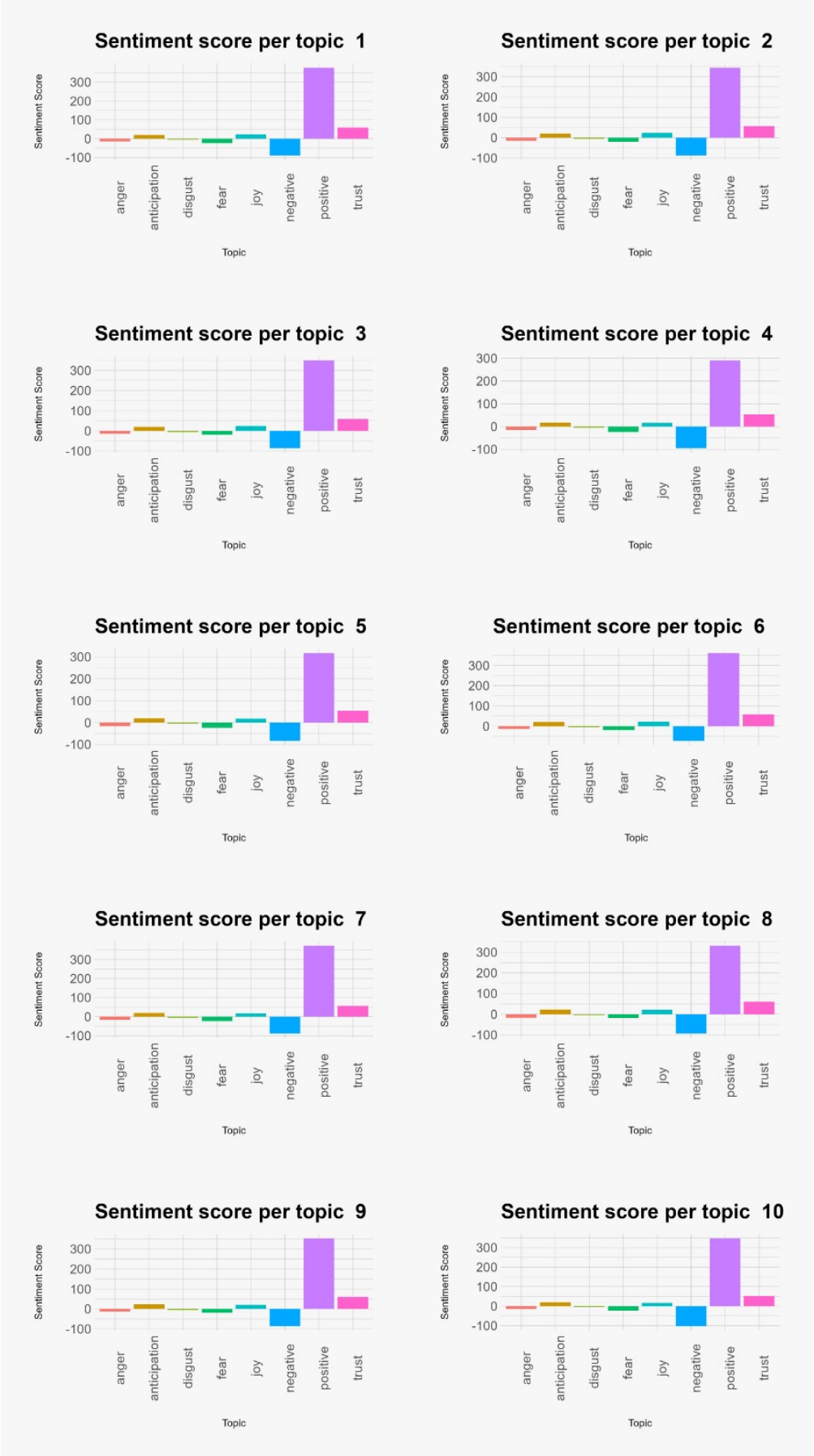
The above graph displays the average sentiment for each topic. This metric is calculated using the beta for the terms of each topic as well as the value of each term from the afinn sentiment database. This shows that most topics are positive, with topic 3 being the most positive.Topic 4 is the least positive of all.

Proportional sentiments



The above graph represents the portion that each sentiment category takes up for each topic. Each topic has similar proportions for each sentiment, with subtle differences in the postive and negative sentiments. As shown, positive is the most prevalent sentiment, contributing to most papers being positive on average.

Sentiments per topic



This above graph shows the sentiment proportions for each topic in more detail. Each

graph shows how much of each sentiment is in each topic, with negative sentiments and emotions being negative on the y-axis.

Conclusion

The field of Information Systems is a dynamic and evolving one. Through this analysis of IS research from 2011 to 2020, we can see a shift in themes and topics over the years. From research on traditional topics such as knowledge management in the early 2010s to focusing more on digital transformation and social media in later years. By investigating the patterns in topics, authors, keywords and sentiments, we found the field of Information Systems is adaptable and constantly evolving, and researches technical issues as well as social issues.