information

MDPI

Article

Phishing and the Human Factor: Insights from a Bibliometric Analysis

Meltem Mutlutürk [1], **Martin Wynn** [2,*] and **Bilgin Metin** [1]

[1] Department of Management Information Systems, Bogazici University, Istanbul 34342, Turkey; meltem.mutluturk@bogazici.edu.tr (M.M.); bilgin.metin@bogazici.edu.tr (B.M.)
[2] School of Business, Computing and Social Sciences, University of Gloucestershire, Cheltenham GL50 2RH, UK
* Correspondence: mwynn@glos.ac.uk

Abstract: Academic research on the human element in phishing attacks is essential for developing effective prevention and detection strategies and guiding policymakers to protect individuals and organizations from cyber threats. This bibliometric study offers a comprehensive overview of international research on phishing and human factors from 2006 to 2024. Analysing 308 articles from the Web of Science database, a significant increase in publications since 2015 was identified, highlighting the growing importance of this field. The study revealed influential authors such as Vishwanath and Rao, leading journals like *Computers & Security*, and key contributing institutions including Carnegie Mellon University. The analysis uncovered strong collaborations between institutions and countries, with the USA being the most prolific and collaborative. Emerging research themes focus on psychological factors influencing phishing susceptibility, user-centric security measures, and the integration of technological solutions with human behaviour insights. The findings highlight the need for increased collaboration between academia and non-academic organizations and the exploration of industry-specific challenges. These insights offer valuable guidance for researchers, practitioners, and policymakers to advance their understanding of phishing attacks, human factors, and resource allocation in this critical aspect of digitalisation, which continues to have significant impacts across business and society at large.

Keywords: phishing; human factor; bibliometric; collaboration; VOSviewer

check for updates

Citation: Mutlutürk, M.; Wynn, M.; Metin, B. Phishing and the Human Factor: Insights from a Bibliometric Analysis. *Information* **2024**, *15*, 643. https://doi.org/10.3390/info15100643

Academic Editor: Robin Haunschild

Received: 11 August 2024
Revised: 10 October 2024
Accepted: 11 October 2024
Published: 15 October 2024

1. Introduction

Phishing remains a pervasive cybersecurity threat, exploiting human psychological weaknesses to obtain sensitive information and infiltrate systems [1,2]. This form of cyber-attack seeks to deceive users into disclosing sensitive information or performing malicious actions by impersonating legitimate entities or individuals [3]. Phishing techniques have evolved over time, including email phishing, spear-phishing, whaling, vishing (voice phishing), and smishing (SMS phishing), making it increasingly difficult for users to recognize malicious communications [4,5]. The consequences can be severe for both individuals and organizations, including identity theft, financial losses, data breaches, and reputational damage. Throughout 2023, the Anti-Phishing Working Group (APWG) observed nearly five million phishing attacks, marking it a record-breaking year. [6]. As technology advances, so do the techniques employed by malicious actors, creating a constant arms race between security professionals and cybercriminals [4].

Understanding the factors influencing phishing success or failure as well as effective countermeasures is of increasing significance as digitalisation continues apace in the business community and across global society. The human factor is a critical component in cybersecurity, as users are often considered the weakest link in the security chain [1]. Human factors encompass psychological, cognitive, social, and emotional aspects of human interaction with technology and play a key role in users' vulnerability or resilience to phishing attacks [5]. Studies have shown that cognitive biases, lack of security awareness, and individual differences such as personality traits can significantly contribute to phishing susceptibility [7,8]. Phishing typically involves deceptive emails, text messages, or websites designed to resemble legitimate communications from trusted entities [9,10]. These messages attempt to manipulate users into revealing personal information or unwittingly installing malware that facilitates unauthorized access to systems or data [11]. Despite sophisticated technological countermeasures, phishing remains a significant threat, primarily due to the human vulnerabilities it exploits [12].

Academic studies on phishing attacks play a crucial role in combating cyber threats. These studies enable experts to analyse phishing methods, vulnerabilities, and patterns, which can then be used to develop effective prevention and detection techniques. Additionally, academic research helps raise awareness about the dangers of phishing attacks and informs policy decisions related to cybersecurity. In recent years, the academic literature on phishing attacks and the human factors contributing to their success has grown significantly [2]. Researchers have focused on various aspects, including cognitive biases, user awareness, personality traits, and the effectiveness of training programs. For instance, the cognitive processes involved in the decision to trust or distrust phishing emails have been examined [8], and the impact of cue utilization and cognitive reflection on email users' ability to distinguish between phishing and genuine emails has also been researched [13]. The role of personality traits in phishing susceptibility has been explored [14–16], with findings indicating that certain traits, such as conscientiousness and neuroticism, may influence an individual's likelihood of falling for phishing attempts. Furthermore, various training approaches to improve phishing detection have been developed and evaluated [17,18], demonstrating the potential for targeted education to enhance users' resilience to these attacks.

Bibliometric analysis is a valuable tool for mapping the research landscape, identifying trends, and highlighting gaps in the literature [19]. In the context of phishing and human factors, bibliometric analysis can provide a comprehensive overview of the field, revealing key contributors, influential publications, and emerging research themes. Despite the growing body of research on phishing and the human factors contributing to their success, there is a noticeable lack of bibliometric studies specifically focusing on this area [20]. By employing bibliometric methods, existing knowledge can be synthesized, collaboration facilitated, and future research directions enlightened. This approach is particularly important in a field where interdisciplinary collaboration between cybersecurity, psychology, and human–computer interaction is essential. In a previous study [21], emerging research fronts were identified by analysing co-occurring keywords in the literature, providing a comprehensive taxonomy of phishing research. This taxonomy consisted of seven categories: human factors and user behaviour, detection and prevention techniques, security measures and authentication, cyber threats and crime, online platforms and social media, emerging technologies and systems, and economics and protection strategies. Among these, human factors and user behaviour emerged as a critical area requiring further in-depth investigation. By emphasizing these emerging research fronts, the study suggested that researchers should focus on areas such as human factors to address the most pressing issues and advance the state of the art in phishing research.

Building upon these findings, the current study aimed to fill this gap by conducting a focused bibliometric analysis on phishing and human factors, providing valuable insights for both researchers and practitioners in the field of cybersecurity.

The objectives of this study include the following:

1. Identify key themes, trends, and gaps in the literature on phishing attacks and human factors;
2. Analyse publication patterns and collaboration networks among researchers in this field;
3. Assess the growth and evolution of research on phishing attacks and human factors over time;

4. Evaluate the impact of published studies through citation analysis and identify seminal works in the field.

By achieving these objectives, this study aimed to highlight the importance of human factors in phishing research and demonstrate the utility of bibliometric analysis in uncovering insights that can inform effective countermeasures and policy decisions. Bibliometric analysis is particularly suited for this purpose because it enables the systematic and quantitative assessment of large volumes of academic literature, uncovering patterns that may not be apparent through traditional literature reviews.

2. Materials and Methods

2.1. Bibliometric Analysis

Bibliometric analysis is a commonly employed method in scientometrics that quantitatively examines scientific literature by assessing publication trends, citation networks, and collaborative relationships among researchers and institutions. This approach allows scholars to map the intellectual landscape, identify key themes and emerging trends, and evaluate the impact of individual publications and authors within a particular discipline. In the context of phishing and human factors, bibliometric analysis holds particular significance, as it can aid in consolidating the expanding literature in this field, delivering a systematic understanding of the current stage of knowledge [21]. Moreover, it can reveal the degree of research collaboration and interdisciplinary connections between cybersecurity, human behaviour, psychology, and other related fields, promoting interdisciplinary research.

One of the main advantages of bibliometric studies is their ability to provide a thorough, objective, and data-driven overview of a research area, enabling researchers and policymakers to identify gaps in the literature, emerging topics, and potential avenues for future research. The bibliometric approach involves applying quantitative methods such as citation analysis to bibliometric data like publication counts and citation metrics [19]. Researchers utilize scientometric tools and techniques to evaluate the productivity of scientists, forecast their career trajectories, and assess how funding decisions influence the structure of the academic community. These methods depend on scholarly bibliographic data and essential scientometric tools to create knowledge domain maps [22].

Science mapping examines the relationships among research components [23], exploring the intellectual interactions and structural connections between them. Several methods used in science mapping include citation analysis, co-citation analysis, bibliographic coupling, co-word analysis, and co-authorship analysis [24].

2.2. Data Collection

The Web of Science (WoS) database was used for the bibliometric analysis to acquire relevant publications. WoS was selected as the primary data source due to several factors. Firstly, it is a widely acknowledged and comprehensive repository of scientific literature [25], encompassing a diverse range of research areas, including cybersecurity and human behaviour; secondly, it offers extensive citation data, enabling the examination of citation networks and the assessment of the influence of individual publications and authors; and thirdly, its sophisticated search and filtering capabilities facilitate the efficient retrieval of relevant articles based on keywords, publication years, and other criteria. The search was conducted within the Web of Science Core Collection, including the following editions:

- Science Citation Index Expanded (SCI-Expanded);
- Social Sciences Citation Index (SSCI);
- Arts & Humanities Citation Index (A&HCI);
- Conference Proceedings Citation Index—Science (CPCI-S);
- Conference Proceedings Citation Index—Social Science & Humanities (CPCI-SSH);
- Emerging Sources Citation —Index (ESCI).

The search terms were applied within the "Topic" field, which includes titles, abstracts, author keywords, and KeyWords Plus. A search query was formulated utilizing the following relevant keywords: "phishing", "behavio(u)r", "human", "user", "awareness", "cogniti*", "personality", and "susceptibility". The keyword search within the "Topic" field yielded a total of 1700 articles. The retrieved records were further filtered based on their relevance and quality. This filtering process excluded publications that lacked sufficient focus on the topic. Consequently, a final set of 308 articles addressing human factors in phishing was retained for analysis (see Supplementary Materials Table S1). Notably, a considerable number of articles were excluded due to their emphasis on algorithms and classifiers employed to predict susceptibility, which fell outside the scope of this study centred on the human factors of phishing attacks. This exclusion accounts for the significant decrease in the remaining articles for analysis. The inclusion and exclusion criteria can be found in Table 1.

Table 1. Exclusion and Inclusion Criteria.

Exclusion	Inclusion
Studies focused only on determining phishing attacks using machine learning algorithms	Studies including human factors that increase susceptibility to phishing attacks
	Studies that investigate the relationship between human factors and phishing susceptibility
	Studies that use phishing simulations or real-life phishing incidents as a measure of phishing susceptibility

2.3. Data Analysis and Visualization

VOSviewer 1.6.20 is a widely used software tool for data analysis and visualization, specifically designed for constructing and displaying bibliometric networks. This tool enables the creation of various types of networks, such as co-authorship, co-citation, and keyword co-occurrence networks, by utilizing advanced clustering algorithms [26]. These algorithms help uncover the underlying structure and main themes within a research domain. VOSviewer also offers an interactive and user-friendly interface, facilitating the exploration and interpretation of the visualized networks.

The methodological approach involved conducting citation analysis to identify influential publications and authors, co-citation analysis to explore the intellectual structure of the field, co-authorship analysis to examine collaboration patterns, and keyword co-occurrence analysis to uncover emerging research themes.

This study employed VOSviewer to construct and analyse networks based on publication and citation data as well as author and institution collaborations and keyword co-occurrences. The outcomes were visualized as network maps, which highlight key clusters, trends, and relationships within the research domain of phishing and human factors.

Potential limitations of the methodology include reliance of the Web of Science database, which may exclude relevant articles indexed elsewhere, and the possibility of missing pertinent studies due to the specificity of the search terms. Additionally, bibliometric analyses may be influenced by citation practices and publication biases. To enhance reproducibility, detailed descriptions of the search strategy and data collection process are provided, and the dataset used within the analysis is also provided as Supplementary Material.

3. Results

3.1. Publications

A total of 308 articles on phishing and the human factor were published in the Web of Science database from 2006–2024. Figure 1 illustrates a rising trend in scholarly publications addressing this topic. A noteworthy surge in publications has been observed from 2015 onwards. This escalation may be attributed to the Sony Pictures Entertainment Hack, which occurred at the end of 2014, involving a sophisticated phishing campaign and malware deployment leading to the theft of vast amounts of data, including unreleased films, employee personal information, internal emails, and other sensitive documents. The incident caused significant financial and reputational damage to Sony and highlighted the risks of phishing in compromising organizational security. This steady increase in publications can also be attributed to the onset of the pandemic in early 2020, which resulted in individuals being confined to their homes and, consequently, generating numerous attack vectors and vulnerabilities. The 2020 Annual Cybersecurity Report [27] states that phishing continued to be a popular threat during the pandemic due to its simplicity and high success rate; it is thus understandable that the number of articles on this topic has risen since then.

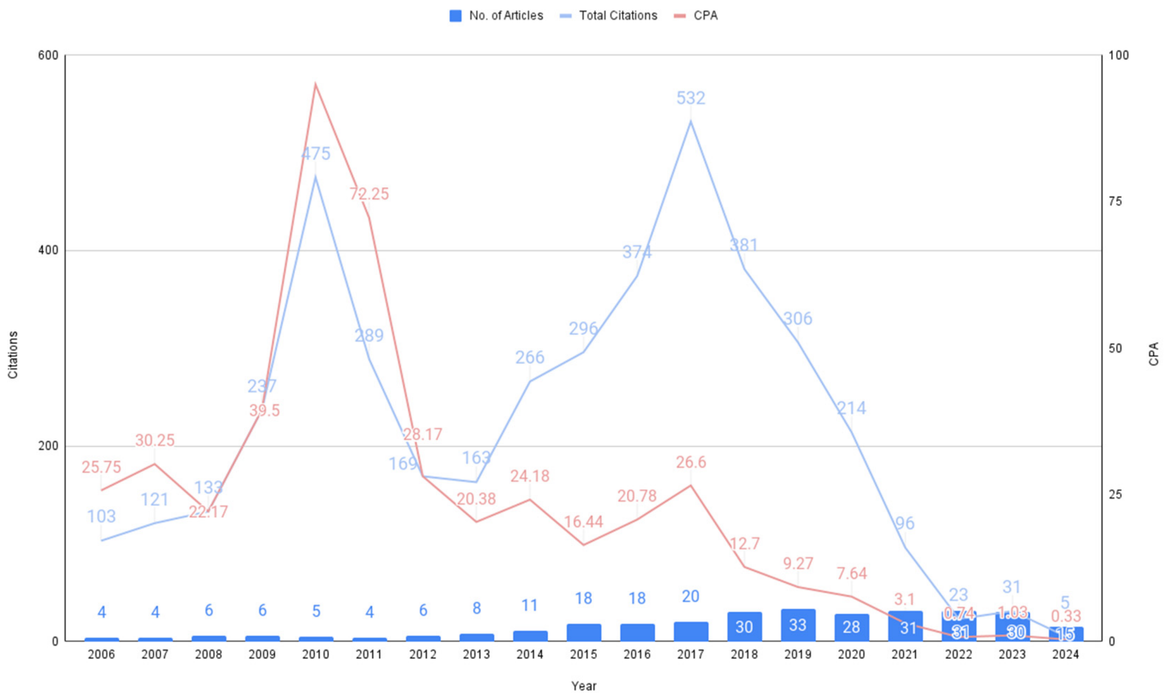

Figure 1. Frequency of publications, corresponding citations, and citations per article (CPA).

The bibliometric analysis contributes to the existing body of literature by providing a comprehensive overview of research trends, key publications, and influential authors in the domain of phishing and human factors. By identifying the most impactful works and examining collaboration patterns, this study offers valuable insights that can guide future research and policy development in cybersecurity.

The 308 articles received a total of 4214 citations, with an average of 13.68 citations per article. A dip in citations from the year 2011 was observed. This may be due to a combination of factors, including citation lag, where newly published articles take time to accumulate citations. Additionally, a shift in research trends during that period may have reduced the relevance of earlier articles, resulting in fewer citations. The sharp increase in 2009–2010 may reflect a temporary surge in interest or highly impactful articles, followed by a return to average citation rates in subsequent years.

Table 2 showcases the top ten articles on phishing and the human factor, based on citations per year (CPY). These studies are among the most influential in the field, shaping current research. The decision to rank articles based on CPY was made to account for the inherent advantage older articles have in accumulating citations over time. CPY provides a more balanced view by normalising citation counts according to the number of years since the article's publication, thereby highlighting the annual impact of each work. This metric is especially useful in revealing the influence of newer articles, which may not have had enough time to gather as many total citations but still have significant annual contributions to the field.

Table 2. The top ten articles on phishing and the human factor, based on citations per year.

Ref.	Author(s)	Year	Title	Source	Affiliation(s)	Total Citations	CPY
[28]	Sheng, S; Holbrook, M; Kumaraguru, P; Cranor, L; Downs, J	2010	Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions	CHI2010: Proceedings Of The 28th Annual Chi Conference on Human Factors in Computing Systems, Vols 1–4	Carnegie Mellon University; Indraprastha Institute of Information Technology Delhi	264	18.86
[29]	Parsons, K; Calic, D; Pattinson, M; Butavicius, M; McCormac, A; Zwaans, T	2017	The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies	*Computers & Security*	Defence Science & Technology; University of Adelaide; University of Adelaide	102	14.57
[30]	Vishwanath, A; Herath, T; Chen, R; Wang, JG; Rao, HR	2011	Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model	*Decision Support Systems*	State University of New York (SUNY) System; State University of New York (SUNY) Buffalo; Brock University; Ball State University; University of Texas Arlington	188	14.46
[31]	Herley, C	2009	So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users	New Security Paradigms Workshop 2009, Proceedings	Microsoft Research	186	12.40
[32]	Alsharnouby, M; Alaca, F; Chiasson, S	2015	Why phishing still works: User strategies for combating phishing attacks	*International Journal of Human-Computer Studies*	Carleton University	107	11.89
[5]	Goel, S; Williams, K; Dincelli, E	2017	Got Phished? Internet Security and Human Vulnerability	*Journal Of the Association for Information Systems*	State University of New York (SUNY) System; State University of New York (SUNY) Albany	80	11.43
[33]	Williams, EJ; Hinds, J; Joinson, AN	2018	Exploring susceptibility to phishing in the workplace	*International Journal of Human-Computer Studies*	University of Bath	63	10.50
[34]	Caputo, DD; Pfleeger, SL; Freeman, JD; Johnson, ME	2014	Going Spear Phishing: Exploring Embedded Training and Awareness	*IEEE Security & Privacy*	Dartmouth College; Vanderbilt University	100	10.00
[35]	Vishwanath, A; Harrison, B; Ng, YJ	2018	Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility	*Communication Research*	State University of New York (SUNY)	59	9.83
[36]	Wang, JG; Herath, T; Chen, R; Vishwanath, A; Rao, HR	2012	Phishing Susceptibility: An Investigation into the Processing of a Targeted Spear Phishing Email	*IEEE Transactions on Professional Communication*	University of Texas Arlington; Brock University; Ball State University; State University of New York (SUNY) System; State University of New York (SUNY) Buffalo; Sogang University	102	8.50

CPY: citations per year.

The most cited article (264 citations) is by Sheng et al. [28], which explores the connection between demographics and phishing susceptibility and evaluates the efficacy of diverse anti-phishing educational materials through a role-play survey with 1001 online participants.

The listed articles cover a range of aspects related to phishing vulnerability. For example, Vishwanath et al. [30] delved into the psychological processing of phishing attempts, while Williams et al. [33] examined factors that influence employee susceptibility to phishing attacks in the workplace settings. Alsharnouby et al. [32] addressed the persistent effectiveness of phishing, highlighting ongoing challenges in user education and cybersecurity implementation.

These top-cited works demonstrate the field's interdisciplinary nature, combining insights from psychology, information systems, and human–computer interaction. This approach is evident in both empirical investigations (e.g., Parsons et al. [29]) and theoretical explorations (e.g., Herley [31]). Such interdisciplinary focus is crucial for developing comprehensive models of phishing susceptibility that incorporate cognitive, emotional, social, and technological factors. This bibliometric analysis builds upon these studies by synthesizing their contributions and identifying key themes and trends that have emerged over time.

The citations also highlight significant advancements, such as the effectiveness of embedded training programs in increasing awareness and reducing susceptibility to spear-phishing attacks [34] and the development of models that examine how suspicion, cognition, and automatic responses contribute to phishing vulnerability [35]. However, they also point to gaps in research, particularly the need for further exploration into how individuals process targeted spear-phishing emails [36] and into the psychological mechanisms that underpin phishing susceptibility. These gaps reveal the importance of refining both educational strategies and cognitive models to improve long-term phishing prevention efforts. By highlighting these gaps, the present study contributes to the existing literature by pinpointing underexplored areas that warrant further investigation. Specifically, the analysis of keyword trends and collaboration networks provides insights into emerging topics and potential interdisciplinary collaborations that can advance the field.

Consequently, from the standpoint of human factors, key studies on phishing point to the impact of psychological characteristics, demographic factors, and individual experiences on user susceptibility. They bring to light the ongoing difficulties in training and awareness initiatives, underlining the necessity for a deeper comprehension of trust dynamics and the development of enduring prevention tactics to efficiently counteract phishing threats. These studies emphasize the critical need to incorporate knowledge from psychology and human behaviour into cybersecurity measures.

3.2. Sources

In terms of journals that publish articles in this field, *Computers & Security* leads with 14 articles on the topic. Table 3 displays the top ten journals with the highest number of articles related to this subject. Examining the titles of the journals in this table reveals that the majority focus on human–computer interaction and human behaviour. This interdisciplinary focus is critical for developing comprehensive models of phishing susceptibility that incorporate cognitive, emotional, social, and technological factors [37]. The findings align with previous bibliometric analyses in cybersecurity, which emphasize the importance of integrating technical and human-centred approaches [38]. The present study adds to the literature by specifically mapping the journals that are central to research on phishing and its human factors, thus providing valuable information for researchers seeking appropriate venues for their work.

The bibliographic coupling of journals, as visualized in the VOSviewer map in Figure 2, reveals significant insights into the research landscape of phishing and human factors. This map was generated using bibliographic coupling as the type of analysis, focusing on journals and the unit of analysis and highlighting connections based on shared citations. To ensure a balanced comparison of journals, normalized citation scores were used as weights, accounting for differences in citation patterns across disciplines. The default clustering algorithm in VOSviewer was applied to group journals into distinct clusters based on the strength of their bibliographic coupling, and a minimum document and citation threshold was set to include journals with substantial contributions to the field.

Table 3. Top 10 journals publishing articles on phishing and the human factor.

Source	Articles	Citations	CPA	JIF_{2023}
Computers & Security	14	134	9.57	4.8
Information and Computer Security	7	42	6.00	1.6
Behaviour & Information Technology	6	141	23.50	2.9
International Journal of Human-Computer Studies	6	213	35.50	5.3
Computers in Human Behavior	5	117	23.40	9.0
Decision Support Systems	5	243	48.60	6.7
European Journal of Information Systems	5	67	13.40	7.3
Human Factors	5	95	19.00	2.9
Financial Cryptography and Data Security	4	154	38.50	-
PLOS ONE	4	67	16.75	-

CPA: citations per article; JIF: journal impact factor.

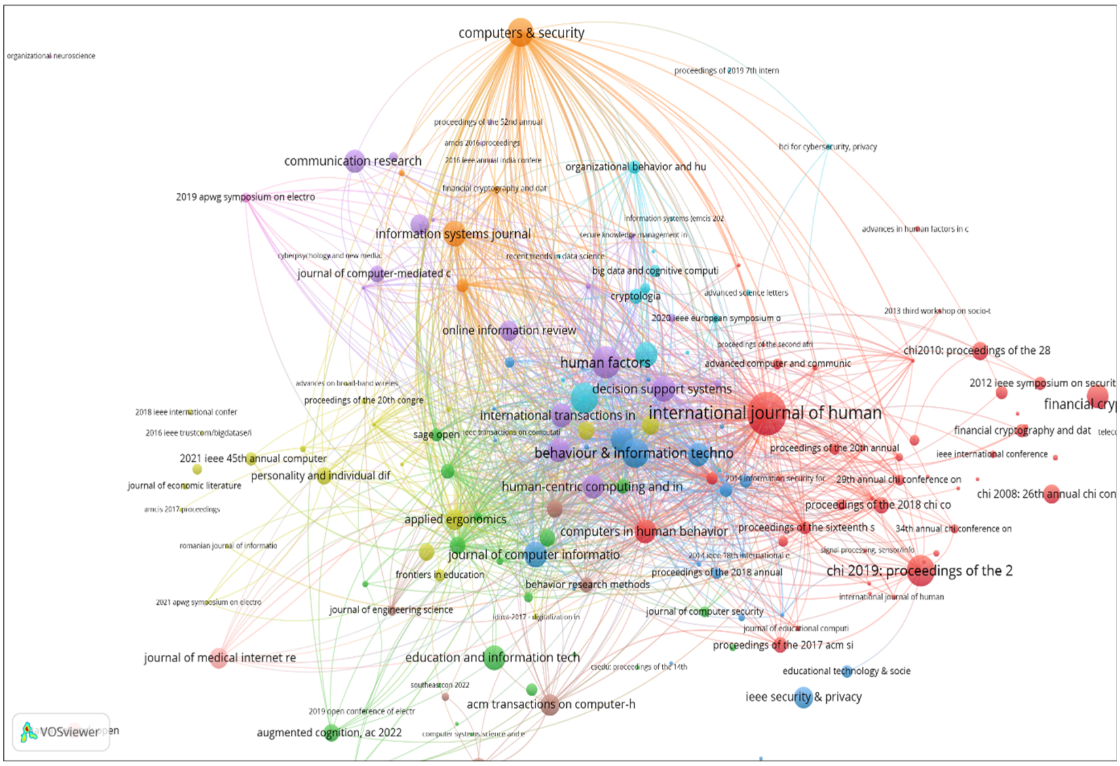
computers & security
organizational neuroscience
proceedings of 2019 7th intern
proceedings of the 52nd annual
amcis 2016 proceedings
2016 ieee annual india confere
hci for cybersecurity, privacy
communication research
organizational behavior and hu
financial cryptography and dat
2019 apwg symposium on electro
information systems iemcis 202
advances in human factors in c
secure knowledge management in
information systems journal
cyberpsychology and new media:
recent trends in data science
journal of computer-mediated c
big data and cognitive computi
advanced science letters
cryptologia
2013 third workshop on socio-t
2020 ieee european symposium o
online information review
proceedings of the second afri
chi2010: proceedings of the 28
human factors
advanced computer and communic
advances on broad-band wireles
2012 ieee symposium on securit
decision support systems
proceedings of the 20th congre
financial cryp
2018 ieee international confer
international transactions in
international journal of human
2016 ieee trustcom/bigdatase/i
sage open
ieee transactions on computati
financial cryptography and dat
teleco
proceedings of the 20th annual
2021 ieee 45th annual computer
behaviour & information techno
ieee international conference
personality and individual dif
journal of economic literature
2014 information security for
29th annual chi conference on
human-centric computing and in
amcis 2017 proceedings
chi 2008: 26th annual chi con
proceedings of the 2018 chi co
applied ergonomics
computers in human behavior
proceedings of the sixteenth s
34th annual chi conference on
romanian journal of informatio
signal processing, sensor/info
journal of computer informatio
2014 ieee 18th international c
frontiers in education
proceedings of the 2018 annual
chi 2019: proceedings of the 2
behavior research methods
international journal of human
2021 apwg symposium on electro
journal of engineering science
journal of computer security
idimt-2017 - digitalization in
journal of educational computi
proceedings of the 2017 acm si
journal of medical internet re
education and information tech
csedu: proceedings of the 14th
educational technology & socie
southeastcon 2022
acm transactions on computer-h
ieee security & privacy
2019 open conference of electr
VOSviewer
augmented cognition, ac 2022
computer systems science and e

Figure 2. Bibliographic coupling analysis of sources.

Computers & Security stands out as a central journal with a high number of citations and strong bibliographic coupling links, highlighting its pivotal role in disseminating key research findings related to cybersecurity. Similarly, *International Journal of Human-Computer Studies* and *Computers in Human Behavior* are influential, reflecting their critical contributions to understanding human–computer interaction and behavioural aspects in cybersecurity. These clusters illustrate the multidisciplinary nature of research in this area and highlight significant interdisciplinary connections, such as the strong coupling between *Computers & Security* and journals regarding the human–computer interaction, indicating a growing cross-disciplinary interest in securing human interfaces and understanding user vulnerabilities.

Furthermore, journals like *Decision Support Systems* and *Journal of Management Information Systems* connect with both technical cybersecurity journals and those focused on human factors, suggesting integrated approaches to managing cybersecurity risks within organizational contexts. Despite having fewer documents, some journals exhibit strong coupling strength, indicating emerging influence. The *APWG Symposium on Electronic Crime Research* is notable for its strong connections, emphasizing its growing role in research on electronic crime and phishing. This analysis of sources contributes to the existing literature by identifying not only established journals but also emerging platforms that are shaping the discourse on phishing and human factors. This information is valuable for both new and established researchers in the field.

Technical and behavioural insights are evident from the strong bibliographic coupling of *IEEE Symposium on Security and Privacy* and *ACM Transactions on Computer-Human Interaction* as well as *Journal of Medical Internet Research* and *Cyberpsychology, Behavior, and Social Networking*. These connections indicate a robust interest in understanding how human behaviour impacts security outcomes.

3.3. Authors

Figure 3 presents the quantity of articles and the extent of author collaboration. This does not include solo authors since collaboration can only be assessed with multiple authors. The "No collaboration" category indicates the number of articles with authors from the same institution. "National collaboration" represents authors from the same country but different institutions. Finally, "International collaboration" refers to the number of articles written by authors from various countries.

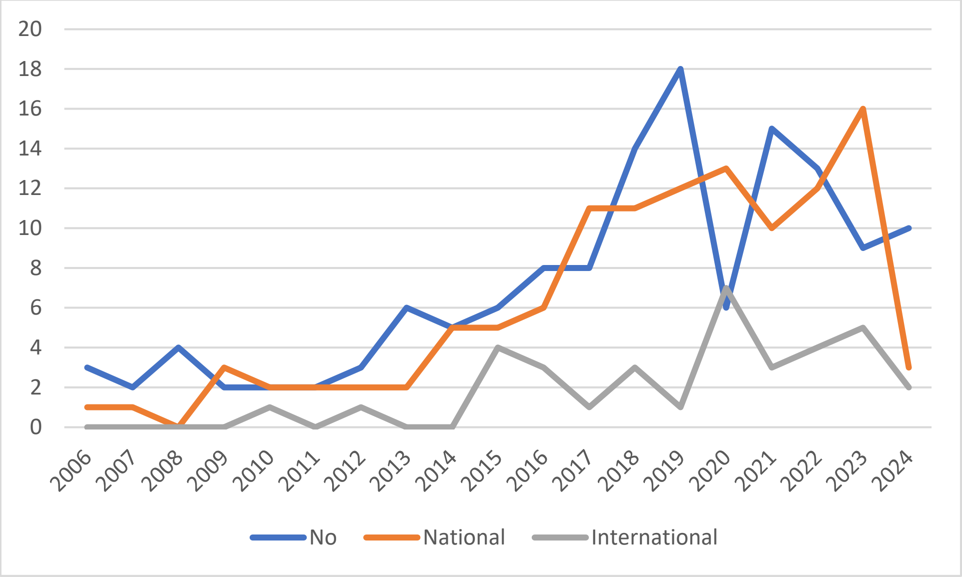

Figure 3. Author collaboration.

The collaboration patterns in phishing and human factor research from 2006 to 2024 reveal a dynamic and evolving landscape. Overall, there has been a general increase in all types of collaborations over the years, indicating growing interest and research activity in this field. "No collaboration" peaked in 2019 with 18 publications, showing fluctuations but generally increasing from 2006 to 2019 before declining, possibly indicating a shift towards more collaborative research. "National collaboration" demonstrated the most consistent growth, peaking in 2023 with 16 publications. "International collaboration", while generally lower than the other categories, showed sporadic growth with notable increases in 2015 and 2020, suggesting potential for increased global cooperative efforts.

Figure 3 shows a shift in collaboration patterns over time. The early years (2006–2012) were dominated by no collaboration or national collaboration. The middle years (2013–2019) saw growth across all categories, with solo research often leading. Recent years (2020–2024) show a more balanced distribution between collaboration types, with national collaboration frequently taking the lead. Notably, 2020 saw a significant spike in international collaborations, possibly due to the global focus on cybersecurity during the COVID-19 pandemic. The year 2023 had the highest number of national collaborations, suggesting strong within-country research networks. These trends suggest a maturing research field in phishing and human factors, with a growing emphasis on collaborative work, particularly at the national level, as researchers increasingly work together to address complex challenges in cybersecurity.

It is important to acknowledge that the increase in collaborations observed in phishing and human factors research may be part of a broader trend in scientific research. However, phishing and human factors research exhibit unique aspects that may influence its collaboration patterns differently compared to other fields. The interdisciplinary nature of this research area, which combines elements from psychology, sociology, human–computer interaction, and cybersecurity, necessitates collaboration across diverse disciplines. This requirement for interdisciplinary expertise may lead to distinct collaboration dynamics. For example, researchers may collaborate nationally to leverage shared language, cultural understanding, and easier coordination when integrating insights from social sciences into technical cybersecurity measures.

Additionally, phishing attacks often exploit cultural and psychological factors specific to certain populations. National collaborations may be particularly valuable for understanding local contexts, social norms, and user behaviours that influence phishing susceptibility. Researchers working within the same country can more readily access relevant participant pools and tailor interventions to the specific needs of their population.

International collaborations, while increasingly important due to the global nature of phishing threats, may present challenges unique to research on phishing and human factors. Differences in language, cultural norms, ethical considerations, and legal regulations regarding human-subject research can make international collaboration more complex in this field compared to more technically focused areas like network security. These factors may contribute to the observed prevalence of national collaborations over international ones in the research on phishing and human factors.

Moreover, the need to develop culturally sensitive anti-phishing strategies underscores the importance of national research efforts. While technical solutions in network security may be broadly applicable across different contexts, interventions addressing human factors in phishing often require a deep understanding of local cultural nuances.

In summary, while the increase in collaborations in research on phishing and human factors aligns with general trends in scientific research, the unique interdisciplinary and cultural aspects of this field shape its specific collaboration patterns. Recognizing these nuances is crucial for fostering effective collaborations that can address the complex challenges posed by phishing attacks.

This study distinguishes between citation, co-citation, and bibliographic coupling networks to highlight their differences. Both bibliographic coupling and co-citation represent indirect relationships and might offer less precise information regarding the relatedness of articles [39]. A citation link connects two items when one cites the other. It is important to note that VOSviewer treats citation links as undirected, meaning that it does not differentiate between a citation from item A to item B and one from item B to item A. A bibliographic coupling link exists between two items when they both reference the same document. In contrast, a co-citation link occurs when two items are both cited by the same document [37].

Figure 4 visualizes the co-authorship network of key researchers in the field of phishing and human factors, weighted by normalized citations. Each node represents an author, with larger nodes indicating more normalized citations received by that author's co-authored work. The closer the nodes are, the more frequently these authors collaborate, and the different colours represent distinct clusters of collaboration, signifying various research communities.

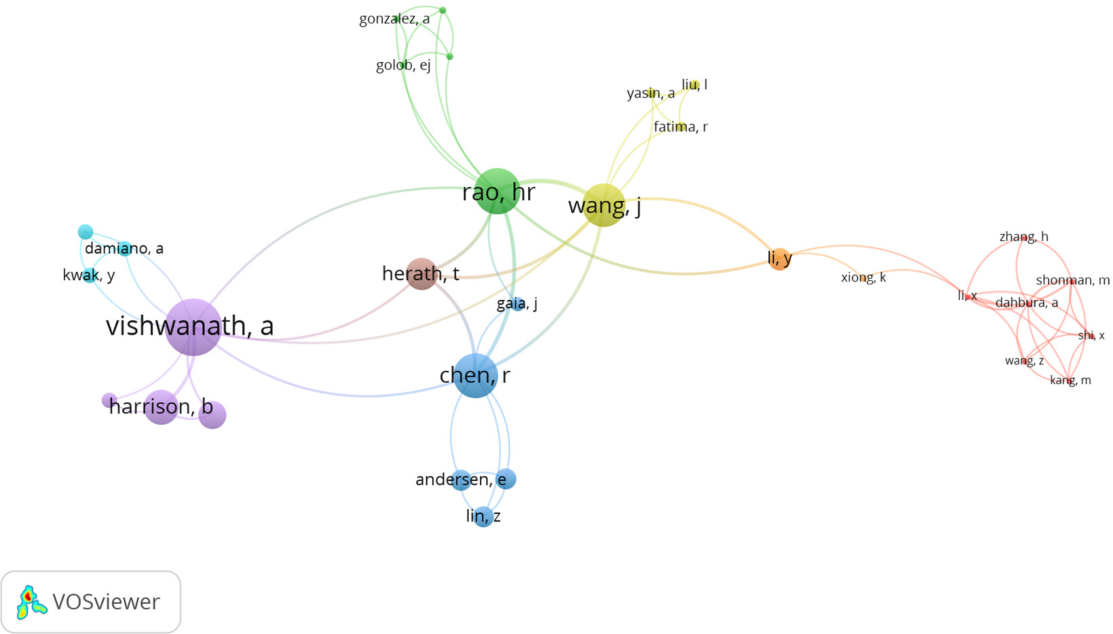

Figure 4. Co-authorship relation of authors.

At the centre of the map is Vishwanath, A, whose work on phishing vulnerability, cognitive models of suspicion, and social media deception has had a profound influence in the field. Vishwanath's studies, particularly on how habitual online behaviours and individual psychological differences impact phishing susceptibility, are highly cited, making Vishwanath a key player in the phishing research landscape. Co-authors such as Harrison, B and Ng, YJ are part of the same purple cluster, showing their significant contributions to understanding how cognitive and personality factors, such as suspicion and automaticity, influence users' susceptibility to phishing. The research themes of Vishwanath's group also extend into exploring email habits, media habits, and the role of personality traits like self-regulation and cognitive biases in phishing detection, thus addressing both individual and contextual factors that contribute to susceptibility to phishing.

Rao, HR, appearing in the green cluster, is another major contributor with a focus on information assurance, cybersecurity policies, and privacy. Rao's research, often in collaboration with authors like Wang, J and Chen, R, addresses email authentication, self-efficacy in phishing detection, and information processing models to understand user behaviour in cybersecurity contexts. Their work has pioneered efforts in visual email authentication, integrating cognitive effort and decision aids to reduce phishing vulnerability. This cluster illustrates the interdisciplinary efforts bridging human behaviour, policy compliance, and technical interventions.

Wang, J is another key figure closely connected to Rao, with significant contributions to phishing detection, coping mechanisms in phishing threats, and cybersecurity compliance policies. This collaboration, reflected in the strong ties between their nodes, emphasises the cross-disciplinary approach needed to address phishing from both technical and behavioural perspectives. Their joint work on the extended parallel process model (EPPM) for phishing detection highlights how users process and respond to phishing attacks, emphasising self-efficacy and cognitive strategies.

Chen, R (blue cluster), often a bridge between Rao's group and Vishwanath's, also plays a key role in phishing research. Chen's work emphasises phishing susceptibility and the design of phishing deception indicators, linking technical phishing detection systems with human-centred security strategies. Collaborators like Lin, Z and Andersen, E further this interdisciplinary connection by focusing on human–computer interactions and usability in phishing contexts.

In contrast, Li, Y, highlighted in the orange cluster on the far right of the map, leads a specialised group focused on fraud detection, cybersecurity threat models, and large-scale data analysis. This cluster is relatively isolated from the more behaviourally focused clusters of Vishwanath and Rao, suggesting a more technical research orientation. Li's work, including machine learning for phishing detection and cognitive behaviour modelling, reflects a deep technical expertise in identifying and mitigating phishing attacks on a large scale. Herath, T, positioned in the brown cluster near Rao, represents more niche contributions within user behaviour and phishing detection. Herath's work often emphasises the rational rejection of security advice and user behaviour, as seen in their collaboration with Rao and Vishwanath, and their contributions are critical in understanding how users interact with phishing defences.

When viewed alongside the author collaboration trends in Figure 3, which shows the growth of national and international collaborations over time, Figure 4 offers a detailed view of specific co-authorship relationships driving the field forward. For example, the strong collaboration between Vishwanath, A; Rao, HR; and Wang, J reflects not only the interdisciplinary nature of phishing research but also the critical importance of collaborative networks in addressing both technical and human factors in cybersecurity.

The relative isolation of certain clusters, such as the technical group of Li, Y, from the central clusters indicates potential gaps in collaboration between highly technical and behaviourally focused research. Li, Y and their team could benefit from greater integration with behavioural scientists to develop more holistic solutions to phishing detection. Conversely, behavioural clusters like Vishwanath's may gain from incorporating advanced detection technologies that Li and their team have developed.

In Figure 5, each circle represents an author, with larger circles indicating a higher number of publications. The proximity between two circles (authors) signifies the strength of their relationship based on bibliographic coupling [38]. This means that authors positioned closer together in the visualization tend to cite the same publications, while those situated further apart generally do not share common cited works [40]. This network visualisation was generated by employing the bibliographic coupling method with authors as the unit of analysis. The size of each author node is weighted by citations, with larger circles indicating a higher citation count. The association strength normalisation method was used to normalize the bibliographic coupling links, ensuring that the proximity between authors reflects the strength of their shared citations rather than just the raw number of citations. The clustering resolution was adjusted to highlight distinct research communities, which are represented by different colour clusters, making it easier to identify groups of researchers working on similar topics. The distance between authors in the map reflects the extent to which they cite similar works.

The map highlights various clusters representing distinct research communities, with prominent clusters centred around influential authors like Vishwanath, Wang, and Rao, demonstrating their significant contributions and influence. While bibliographic coupling does not directly indicate co-authorship, it reflects intellectual proximity and the likelihood of thematic alignment in their research. Authors such as Rajivan, Caputo, and Parsons are positioned close to one another, reflecting that they are engaged in similar bodies of literature, even if they may not have collaborated directly. The shared intellectual foundation, as represented by strong links between them, suggests potential for future collaboration and indicates that their work addresses common research problems, which is vital for the evolution of anti-phishing strategies. This network structure is a valuable tool for understanding how research fields evolve and for identifying potential collaborators who share common research interests.

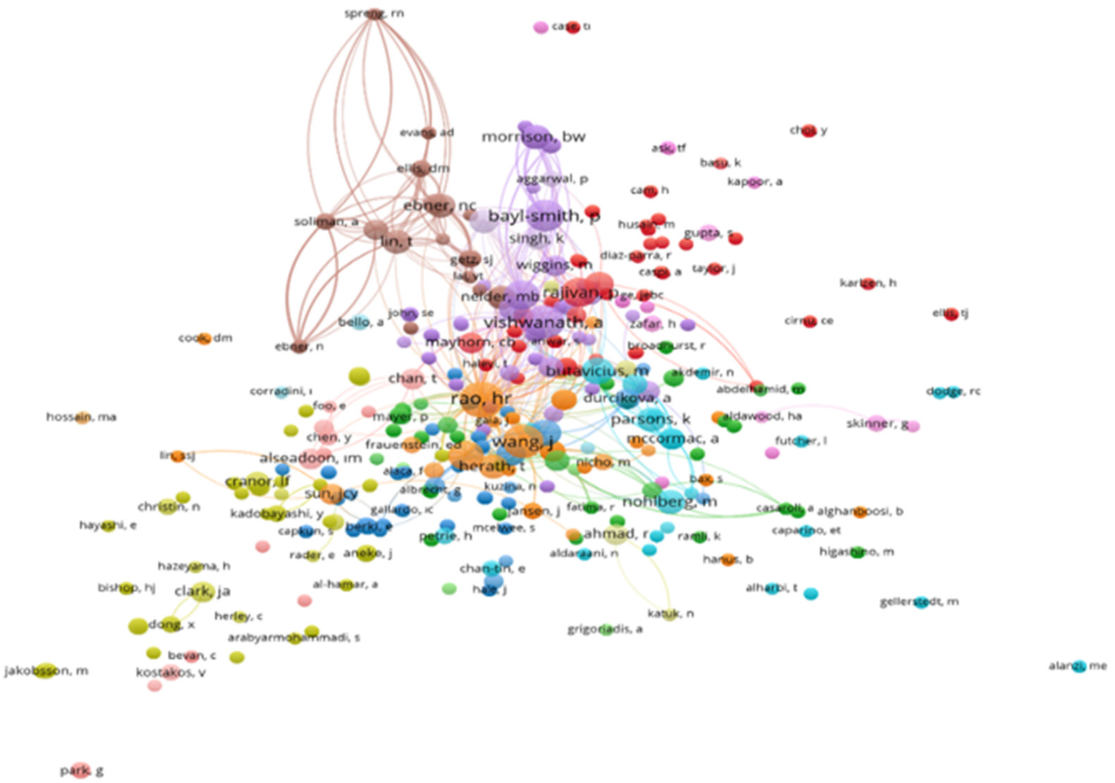

Figure 5. Bibliographic coupling of authors.

Figure 6 displays a co-citation map that visualizes the authors who are most frequently cited together. In this map, authors that are closer to each other and have stronger connecting links have been cited together more often in various publications. Co-citation occurs when two documents are both cited by a common third document [41,42]. The analysis of co-citation operates under the assumption that two papers cited together share a strong relationship and should, therefore, be grouped together. Each circle or node on the map represents an author, and the connections (links) between these nodes signify co-citation relationships among authors. The proximity between two authors on the map roughly indicates the extent of their relatedness based on co-citations [43].

Looking at the clusters in Figure 6, the green cluster is anchored by the author Kumaraguru, the yellow cluster is anchored by Vishwanath, and the blue cluster is anchored by Jakobsson. Kumaraguru's work on user education and awareness programs frequently attracts significant scholarly attention, forming a pivotal point in the green cluster. Vishwanath's centrality in the yellow cluster further reinforces their foundational role in phishing research. Jakobsson has made significant contributions to understanding and mitigating social engineering attacks, anchoring the blue cluster and emphasizing their role in shaping best practices in cybersecurity. As these authors anchor their respective clusters, other authors within these clusters tend to conduct research on the same topics or sub-topics, as they are frequently cited together. This thematic grouping helps identify sub-disciplines within phishing and human factors, providing a clearer picture of the field's landscape and guiding future research directions.

Table 4 shows the top 30 most influential authors in this field of study based on the total number of citations. Vishwanath is the most influential author, with a total of 476 citations from eight publications. Vishwanath's work has significantly advanced the understanding of psychological factors influencing phishing susceptibility, contributing to the development of user-centric security measures [40]. Similarly, Rao's extensive research has provided valuable insights into information assurance and cybersecurity practices. By highlighting these key contributors, the present study helps delineate the intellectual structure of the field and identifies seminal works that have shaped current research trajectories. The most publications of a single author are nine. Rao, having published a total of nine articles, is the most productive.

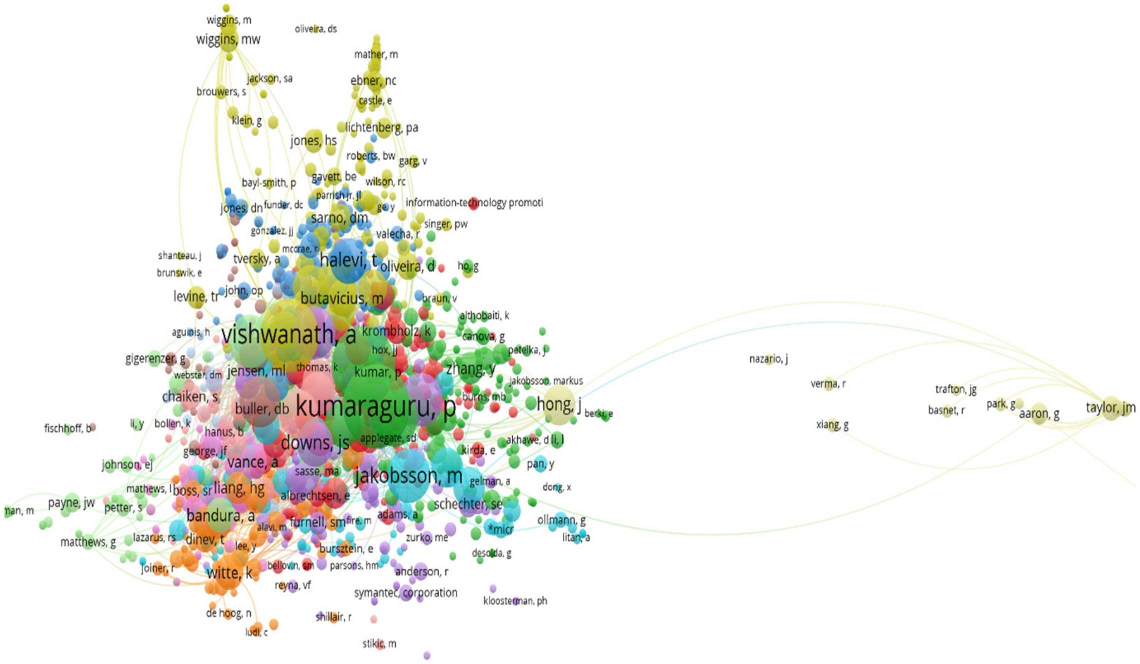

Figure 6. Co-citation analysis of authors.

Table 4. The top 30 most influential authors in this field of study.

Author	Articles	Citations	CPA
Vishwanath, A	8	476	59.5
Rao, Hr	9	425	47.2
Wang, J	7	408	58.28
Chen, R	5	336	67.20
Herath, T	4	323	80.75
Kumaraguru, P	2	287	143.5
Sheng, S	2	287	143.5
Cranor, L	1	264	264
Downs, J	1	264	264
Holbrook, M	1	264	264
Wright, Rt	4	196	49
Marett, K	2	193	96.5
Herley, C	1	186	186
Butavicius, M	4	133	33.25
Parsons, K	4	137	34.25
Harrison, B	4	133	33.25
Lin, T	4	120	30
Calic, D	3	114	38
Dommaraju, S	3	114	38
Mccormac, A	3	114	38
Pattinson, M	3	114	38
Alaca, F	1	107	107
Alsharnouby, M	1	107	107

Table 4. *Cont.*

Author	Articles	Citations	CPA
Chiasson, S	1	107	107
Zwaans, T	1	102	102
Caputo, Dd	1	100	100
Freeman, Jd	1	100	100
Johnson, Me	1	100	100
Pfleeger, Sl	1	100	100
Li, Y	4	88	22

3.4. Organizations

There are a few points of interest here. The co-authorship map (Figure 7) indicates that the University of Virginia has co-authored publications with organizations from various clusters (e.g., University of Michigan and University of Oklahoma), and Figure 8 tells us that Carnegie Mellon University shares a large number of references with other organizations, as this university is seen to anchor the network. Carnegie Mellon University is the institution with the most citations and number of articles published on this topic of study (Table 5). By specifically identifying key organizations in the phishing and human factors domain, these results can guide researchers seeking institutional collaboration or considering academic programs in this field.

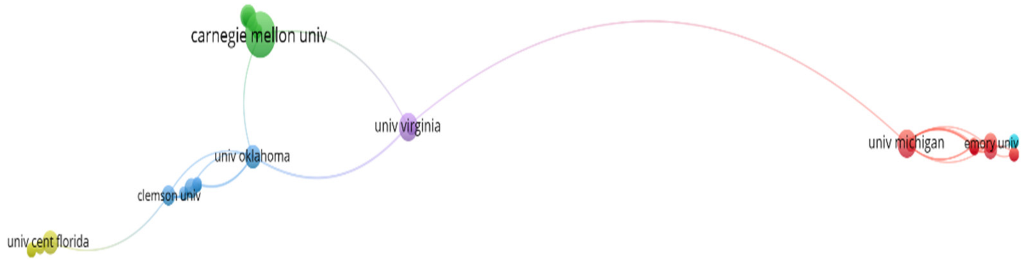

Figure 7. Co-authorship analysis of organizations.

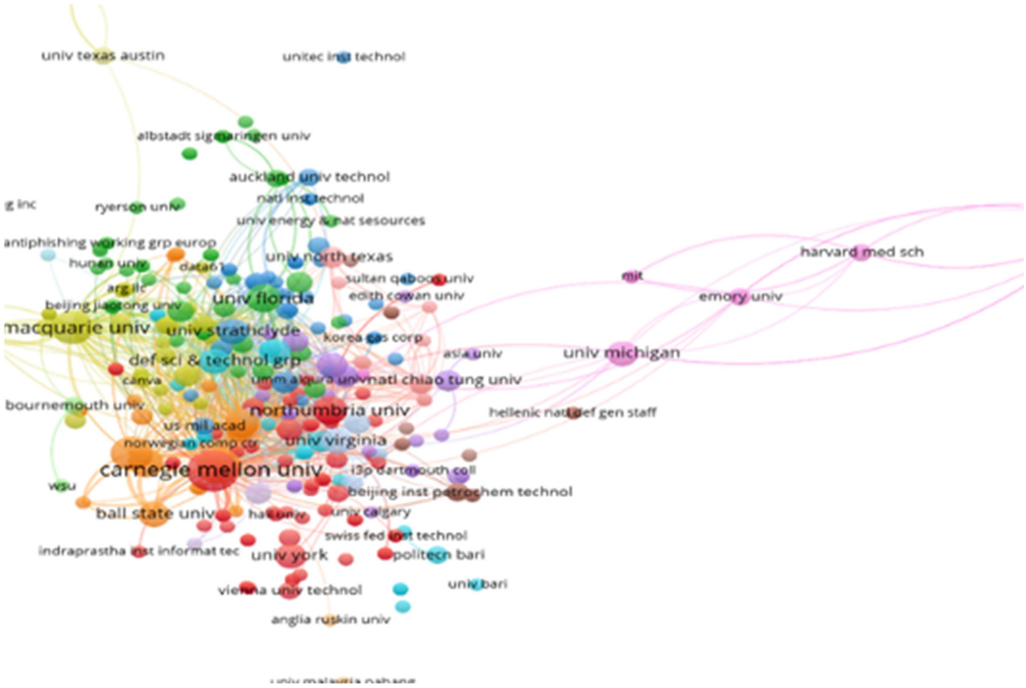

Figure 8. Bibliographic coupling analysis of organizations.

Table 5. Top 30 organizations.

Organization	Country	Articles	Citations	CPA
Carnegie Mellon University	USA	11	510	46.36
Suny Buffalo University	USA	7	435	62.14
University Texas Arlington	USA	8	414	51.75
Brock University	Canada	5	325	65.00
Ball State University	USA	4	323	80.75
Indraprastha Institute of Information Technology	India	1	264	264.00
Mississippi State University	USA	2	193	96.50
Microsoft Research	USA	1	186	186.00
University Adelaide	Australia	4	137	34.25
Defence Science & Technology Group	Australia	4	131	32.75
University Florida	USA	5	124	24.80
University San Francisco	USA	1	113	113.00
Northumbria University	U.K.	5	113	22.60
Carleton University	Canada	1	107	107.00
New York University	USA	2	107	53.50
Sogang University	South Korea	1	102	102.00
University Texas San Antonio	USA	5	102	20.40
Dartmouth College	USA	1	100	100.00
Vanderbilt University	USA	1	100	100.00
University Oklahoma	USA	3	83	27.67
Clemson University	USA	2	82	41.00
Suny Albany University	USA	1	80	80.00
University Massachusetts	USA	1	80	80.00
University South Carolina Upstate	USA	1	80	80.00
Vienna University of Technology	Austria	2	80	40.00
National Chiao Tung University	Taiwan	3	78	26.00
University Bath	U.K.	2	72	36.00
University Buffalo	USA	3	72	24.00
Nelson Mandela Metropolitan University	South Africa	3	65	21.67
University Of South Australia	Australia	1	61	61.00

In Figure 7, the co-authorship analysis map was generated based on the co-authorship network of organizations, using the association strength normalization method. Each node (circle) represents an organization, and its size indicates the number of publications associated with that organization. The distance between nodes represents the strength of co-authorship relationships, meaning that organizations that have co-authored more papers appear closer together. The visualization was filtered by selecting organizations with at least a certain number of documents (in this case, a threshold of a minimum number of documents published was applied). The map's clusters are coloured to represent different research communities or collaboration networks.

For Figure 8, the bibliographic coupling analysis map was created to illustrate the extent to which organizations cite the same publications. VOSviewer computes this by measuring the number of references shared between organizations. The map uses association strength normalization to account for the varying number of publications per organization, ensuring that larger institutions like Carnegie Mellon University, with more publications and citations, do not dominate the visualization. As in the co-authorship analysis, the size of each node indicates the organization's prominence in terms of citations, while the thickness of the lines between nodes shows the strength of the bibliographic coupling. Clusters of organizations that frequently cite similar works are grouped together, highlighting major collaborative networks and identifying key institutions shaping the discourse on phishing and human factors.

3.5. Countries

Figure 9 presents a co-authorship map that highlights the most collaborative countries in this field. In this visualization, the size of each node corresponds to the number of publications, and the colour represents different clusters. Countries within the same cluster have collaborated more frequently on publications. The United States emerges as the most collaborative country, with a total of seven links, and it has the strongest collaboration with Canada based on link strength. The USA is also the most prolific country, accumulating a total of 2438 citations among 97 articles. Table 6 shows the top 30 countries in terms of total number of citations. By identifying these geographic patterns, this study contributes to the understanding of how regional and cultural factors may influence research priorities and collaboration opportunities.

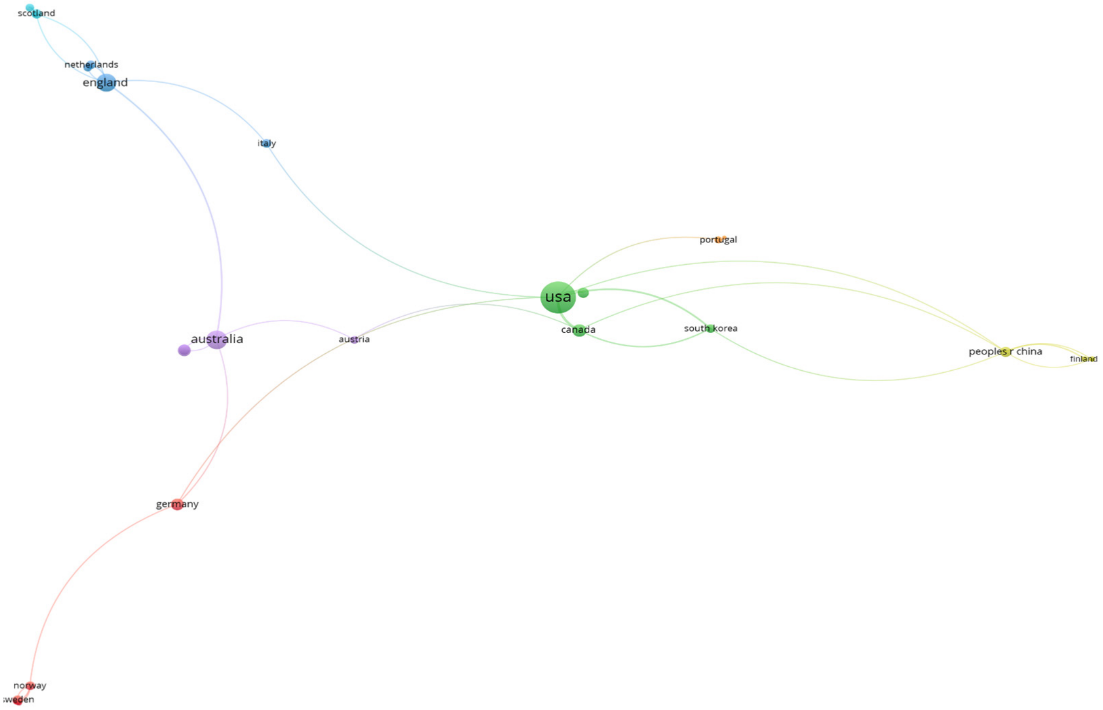

Figure 9. Co-authorship analysis of countries.

Table 6. Top 30 countries based on number of citations.

Country	Articles	Citations	CPA
USA	97	2438	25.13
Canada	11	491	44.64
England	25	377	15.08
Australia	29	340	11.72
India	7	331	47.29
South Korea	5	132	26.40
Austria	4	97	24.25
South Africa	9	85	9.44
Taiwan	5	79	15.80
Netherlands	5	72	14.40
Portugal	3	68	22.67
Germany	9	61	6.78
Saudi Arabia	9	47	5.22
Scotland	6	43	7.17
Sweden	6	40	6.67

Table 6. *Cont.*

Country	Articles	Citations	CPA
Peoples R China	7	32	4.57
Luxembourg	1	24	24.00
Spain	2	21	10.50
Switzerland	1	16	16.00
Kenya	2	16	8.00
New Zealand	4	16	4.00
U. Arab Emirates	4	16	4.00
Malaysia	7	14	2.00
Israel	2	11	5.50
Finland	2	10	5.00
Japan	6	10	1.67
Belgium	2	8	4.00
Italy	5	7	1.40
Greece	2	6	3.00
Jordan	2	6	3.00

3.6. Keywords

It is of value to identify emerging research fronts to pinpoint research endeavours within a specific scientific domain [22]. Figure 10 presents a map of keywords that co-occur. The original search criterion in our data query ("Phishing") was excluded from the keywords. Additionally, the keywords were checked for spelling differences and erroneous information. For example, differences in American vs. British English and wording of keywords were standardised. A thesaurus file was created and included in the analysis to take the differences into account. Based on the co-occurrence analysis of keywords in the field of phishing and human factor literature, several key themes and research directions emerge.

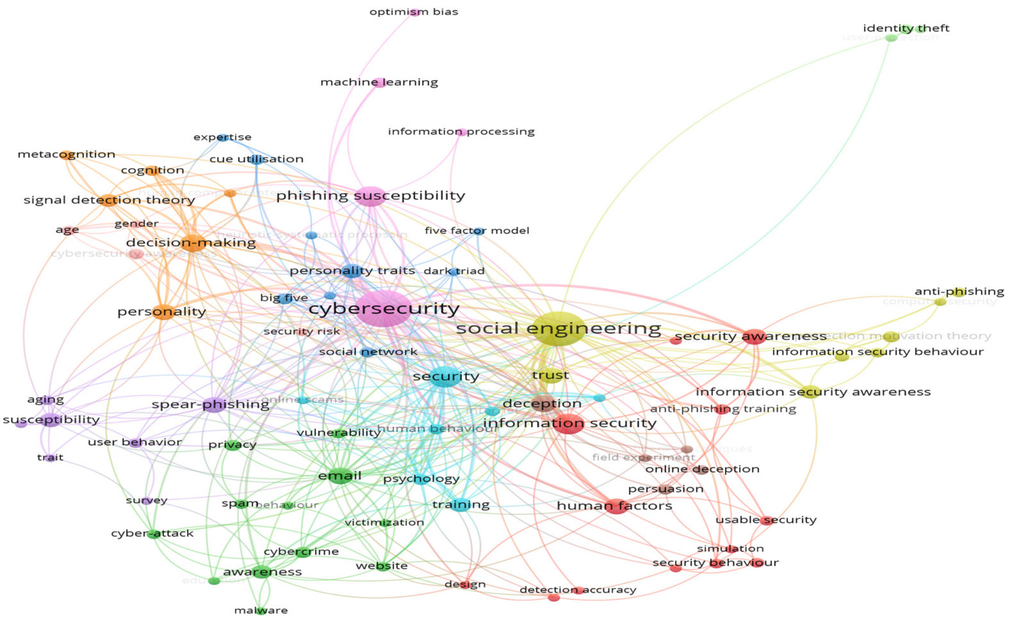

Figure 10. Co-occurrence analysis of keywords.

The largest cluster (red cluster) in Figure 10 focuses on the practical aspects of anti-phishing efforts, emphasizing training, user studies, and the development of usable security measures. This cluster highlights the importance of human factors in cybersecurity, particularly in the context of online banking. The light-blue cluster broadens the scope to include various forms of cyber threats and the role of user awareness and education in mitigating these risks. The blue and orange clusters delve into the psychological aspects of phishing susceptibility, exploring personality traits, cognitive processes, and decision-making models. This psychological focus is complemented by the yellow cluster, which examines theoretical frameworks such as the Protection Motivation Theory and the Theory of Planned Behaviour in the context of information security.

To generate this map, VOSviewer employed a co-occurrence analysis of keywords, using full counting as the method to consider each occurrence of the keywords equally, regardless of how many documents they appeared in. Keywords that met the minimum threshold of appearing in at least three documents were included. For normalization, the association strength method was applied, which ensures that the strength of the relationships between keywords is proportionally balanced based on their co-occurrence frequency.

The purple and pink clusters address demographic factors like age and gender in phishing susceptibility, while other clusters explore the broader context of online scams, deception techniques, and human behaviour on the internet. The inclusion of machine learning and information processing suggests an emerging trend towards incorporating advanced technological solutions in phishing detection and prevention.

The keyword analysis reveals a comprehensive and interdisciplinary approach to phishing research, encompassing technological solutions, psychological insights, and educational strategies. The field appears to be moving towards a more nuanced understanding of phishing susceptibility, considering individual differences, cognitive processes, and the broader context of online behaviour. This suggests that effective anti-phishing strategies will likely require a combination of technological innovations, targeted education, and interventions tailored to individual psychological and demographic factors. Each cluster and the corresponding keywords can be found in Table 7.

Table 7. Keyword clusters.

Cluster	Topic	Keywords
1	Human Factors in Security and Usable Security	anti-phishing training, cognitive modelling, design, detection accuracy, experiment, human factors, information security, online banking, security awareness, security behaviour, simulation, usable security, user study
2	Cybersecurity Threats, Awareness, and Education	awareness, behaviour, cyber-attack, cybercrime, education, email, malware, privacy, spam, victimization, vulnerability, website
3	Personality Traits and Decision Making in Security Contexts	big five, cue utilization, dark triad, expertise, five factor model, heuristic-systematic processing model, personality traits, social network, systematic processing
4	Security Awareness and Behavioural Theories in Information Security	anti-phishing, computer security, information security awareness, information security behaviour, protection motivation theory, social engineering, theory of planned behaviour, trust
5	Age Factors and Phishing Susceptibility	aging, detection, spear-phishing, survey, susceptibility, trait, user behaviour

Table 7. *Cont.*

Cluster	Topic	Keywords
6	Psychological and Behavioural Aspects of Online Security	human behaviour, individual differences, internet, online scams, psychology, security, training
7	Cognitive Processes and Decision Making	cognition, decision-making, human-computer interaction, metacognition, personality, signal detection theory
8	Deception and Persuasion in Online Environments	deception, field experiment, influence techniques, online deception, persuasion
9	Information Processing in Cybersecurity and Phishing Susceptibility	cybersecurity, information processing, machine learning, optimism bias, phishing susceptibility
10	Demographics and Security Awareness	age, cybersecurity awareness, gender, security risk
11	Identity Theft and User Protection	identity theft, internet security, user protection

It is important to note that some synonyms or closely related terms appear in different clusters (e.g., “user behaviour” in cluster 5 and “human behaviour” in cluster 6). This occurrence is due to the way the clustering algorithm in VOSviewer groups keywords based on their co-occurrence patterns. Even though these terms are similar, they may co-occur with different sets of keywords in the literature, reflecting distinct contexts or research focuses. For instance, “user behaviour” might be associated with studies on user interactions with security interfaces, while “human behaviour” could relate to broader psychological aspects influencing phishing susceptibility. The presence of such overlaps indicates the interdisciplinary and interconnected nature of phishing research. This overlap provides valuable insights into the nuanced ways similar concepts are explored across different research themes.

The most frequent author keywords were used to understand the trend topics over the years for each cluster. The size of the x indicates the number of times each keyword has appeared for each year; the larger the size, the more frequent the term. An example of a trend chart—here for cluster 1—can be seen in Figure 11. This cluster focuses on user-centric aspects of security, with keywords like “security awareness”, “security behaviour”, and “user study” emphasizing the role of users in phishing prevention. Technical terms such as “detection accuracy” and “cognitive modelling” indicate an integration of psychological and computational approaches. Recently, “user study” and “usable security” have gained prominence, reflecting a shift towards understanding and enhancing user interaction with security systems. However, there has been a decline in the mention of “cognitive modelling” and “design”, suggesting a move towards more empirical, user-focused studies.

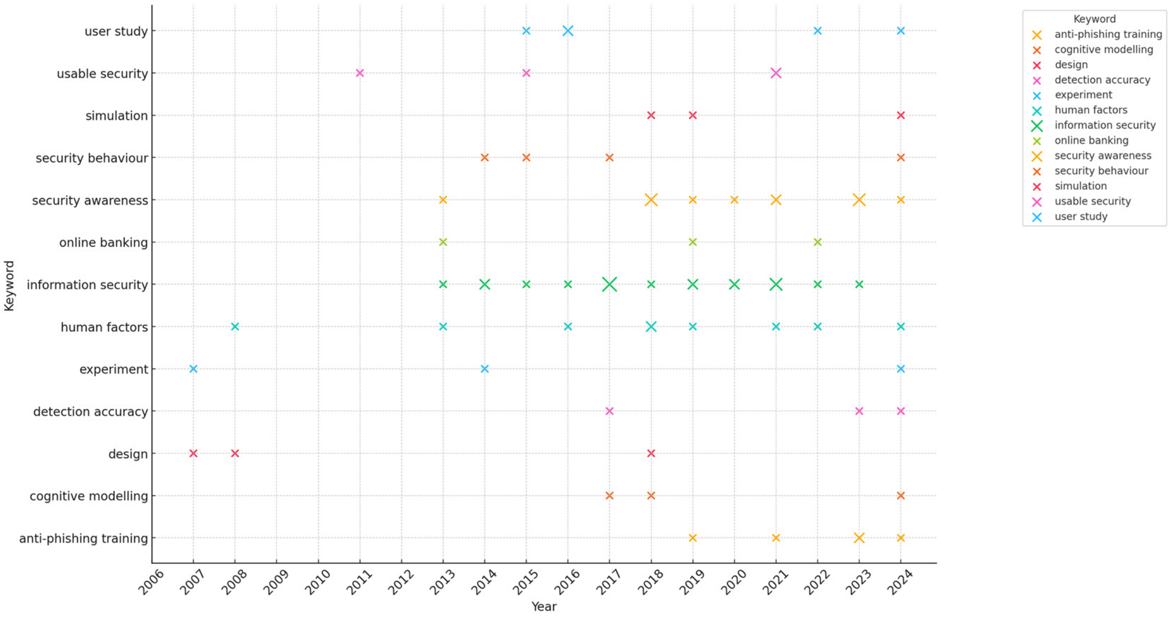

Figure 11. Trend topics based on author keywords for cluster 1.

Cluster 2 centres around broader security concerns such as "cyber-attack", "cybercrime", and "privacy", alongside foundational terms like "education" and "awareness". The increasing attention to "cybercrime" and "cyber-attack" highlights the growing complexity and scale of phishing threats. Meanwhile, the keyword "website" has appeared less frequently in recent years, possibly due to the evolution of web security standards and practices. Cluster 3 is heavily focused on psychological factors, including "personality traits", "dark triad", and "systematic processing", reflecting an interest in understanding individual differences in phishing susceptibility. Emerging topics like "social network" and "heuristic-systematic processing model" suggest a rising interest in how social dynamics and cognitive processes influence phishing behaviour. However, the "five factor model" seems to have received less attention recently, possibly overshadowed by more nuanced psychological models. Cluster 4 blends technical and social aspects, with terms like "computer security", "social engineering", and "trust". It also includes behavioural theories such as "protection motivation theory" and "theory of planned behaviour". The continuing focus on "social engineering" highlights its significance, particularly as it intersects with technical security measures. In contrast, "computer security" as a standalone term has become less frequent, likely due to the integration of more specific and advanced concepts in cybersecurity.

Cluster 5 is concerned with specific attack vectors like "spear-phishing" and demographic factors like "aging" and "trait", alongside general behavioural aspects such as "user behaviour". The increased use of keywords "spear-phishing" and "susceptibility" indicates a focus on targeted phishing attacks and understanding vulnerabilities. The term "survey" is less prominent, suggesting a shift towards more diverse methodologies beyond surveys. Cluster 6 focuses on the human side of security, with keywords like "psychology", "human behaviour", and "training". Broader terms like "security" and "internet" are also included. The increased attention to "training" highlights the importance of educational initiatives in enhancing security. However, the term "online scams" has seen fewer mentions in recent years, possibly due to a shift towards more specific forms of online threats. Cluster 7 explores cognitive and decision-making processes, including "cognition", "decision-making", and "signal detection theory", as well as human–technology interaction through "human-computer interaction". The rising attention to "signal detection theory" and "human-computer interaction" reflects the growing complexity of phishing detection and prevention strategies. In contrast, “metacognition” appears less frequently, possibly due to a shift towards more applied cognitive theories.

Cluster 8 centres on deception techniques, with keywords like “deception”, “persuasion”, and “influence techniques”. The inclusion of “field experiment” indicates an empirical approach to studying phishing. The growing attention to “online deception” and “persuasion” reflects the sophisticated tactics used in modern phishing campaigns. However, “field experiment” seems less frequent, suggesting a possible shift towards other experimental or observational methods. Cluster 9 emphasizes technological approaches to phishing prevention, including “cybersecurity”, “machine learning”, and “information processing” while also addressing psychological factors like “optimism bias”. The increase of “machine learning” and “phishing susceptibility” as keywords highlights the use of advanced technologies and psychological insights in combating phishing. Conversely, “information processing” appears less frequently, possibly due to its integration into more specific applications like machine learning. Cluster 10 explores demographic factors in security, with keywords like “age”, “gender”, and “cybersecurity awareness” alongside “security risk”, reflecting an interest in risk perception. The increased attention to “cybersecurity awareness”, especially in the context of demographic differences, underscores the importance of tailored educational initiatives. The declining frequency of “age” suggests a possible shift towards more complex demographic analyses. Finally, cluster 11 focuses on user protection, with terms like “user protection”, “internet security”, and “identity theft”. The declining frequency of “identity theft” may reflect a broader integration of identity-related concerns into general security practices.

By analysing keywords and their co-occurrence, emerging research themes were identified, such as the integration of machine learning in phishing detection and the exploration of individual psychological traits affecting susceptibility. This detailed examination of research trends contributes to the existing literature by pinpointing areas that require further investigation and by suggesting potential interdisciplinary collaborations.

4. Discussion

The bibliometric analysis of phishing and human factors research reveals several key issues worthy of further discussion. These findings provide specific insights into the unique challenges and developments within the field of phishing.

Firstly, the steady increase in publications since 2015, with a notable surge following major incidents like the 2014 Sony Pictures hack and the onset of the COVID-19 pandemic in 2020, underscores the growing recognition of phishing as a critical cybersecurity threat. These events were heavily influenced by sophisticated phishing campaigns that exploited human vulnerabilities, leading to significant data breaches and financial losses specific to phishing attacks. This trend aligns with the increasing sophistication of phishing attacks and the heightened vulnerability of individuals and organizations in an increasingly digital world. By highlighting this correlation, our study emphasizes the reactive nature of research to real-world events, suggesting that as phishing threats evolve, so does the academic focus on understanding and mitigating them.

Secondly, as indicated in Table 2, this study highlights the multidisciplinary nature of phishing research. Phishing uniquely combines elements of cybersecurity, psychology, and social engineering, necessitating an interdisciplinary approach. The importance of understanding demographic and psychological factors in phishing susceptibility is particularly significant because phishing attackers often exploit specific psychological triggers and social cues [28,29]. The enduring effectiveness of phishing, underscoring the continuous challenges in educating users and implementing cybersecurity measures, is also emphasised by some studies [32]. The interdisciplinary nature of the field is further evidenced by the diversity of journals publishing phishing research (Table 3). The prominence of journals focusing on human–computer interaction and human behaviour, alongside traditional cybersecurity outlets, reflects the field’s evolution towards a more holistic understanding of phishing. This highlights how phishing research specifically benefits from insights across multiple disciplines to address its complex nature. These findings enhance comprehension by revealing that phishing susceptibility is influenced by a complex interplay of psychological, cognitive, and social factors. This highlights the necessity of adopting multidisciplinary approaches to effectively address phishing threats, integrating insights from psychology, cognitive science, and human–computer interaction with technical cybersecurity measures specifically designed to counter phishing techniques.

Thirdly, the focus on human factors represents a significant shift from purely technical approaches to cybersecurity, acknowledging the critical role of user behaviour and individual differences in vulnerability to phishing attacks. In phishing, unlike other cyber threats, the success of an attack often hinges directly on human interaction with deceptive content, making human factors particularly crucial. The high citation counts of these works indicate a growing consensus in the research community about the need for user-centric approaches to phishing prevention, which has been mirrored in practice by the introduction of user-centric approaches to service-desk cybersecurity operations [44]. The keyword analysis (Figures 10 and 11 and Table 7) reveals the prominence of terms related to user awareness, education, and psychological factors across multiple clusters, underscoring the shift towards human-centric approaches in phishing research. This shift enhances our understanding that technical defences alone are insufficient without considering the human element. Effective anti-phishing strategies must incorporate user education, awareness programs, and psychological insights to reduce vulnerability, acknowledging users as both a critical line of defence and a potential weakness in cybersecurity in the context of phishing.

Fourthly, collaboration patterns in the field, as evidenced in Figures 3–6, reveal an increasing trend towards national and international cooperation. The rise in collaborative efforts, particularly since 2015, indicates a growing recognition of the complex, multifaceted nature of phishing threats and the need for diverse expertise to address them effectively. Phishing attacks often have global implications, as attackers can target victims across borders using the internet, making international collaboration particularly pertinent in phishing research. The spike in international collaborations in 2020 may reflect the global nature of cybersecurity challenges highlighted by the COVID-19 pandemic. Analysing these publication patterns and collaboration networks (objective 2) is important because it identifies key contributors and collaboration hubs within the field. This knowledge facilitates targeted collaborations, encourages interdisciplinary research, and helps avoid duplication of efforts, ultimately accelerating advancements in phishing detection and prevention. These collaboration trends suggest a maturing research field and highlight opportunities for further cross-disciplinary and international research initiatives. These findings illustrate that combating phishing requires collaborative efforts that transcend national boundaries and disciplinary silos. By mapping these collaboration networks, the study reveals potential avenues for strengthening international partnerships and fostering interdisciplinary research, which are essential for developing comprehensive solutions to phishing threats that are globally coordinated.

Lastly, the central role of authors like Vishwanath, Kumaraguru, and Jakobsson in their respective clusters, as shown in Figures 4–6, indicates the formation of distinct research communities focusing on different aspects of phishing and human factors. For example, Vishwanath's work often focuses on cognitive processes in phishing susceptibility, while Kumaraguru emphasizes educational interventions, and Jakobsson explores technical aspects of phishing detection. This specialization is valuable for deepening our understanding of specific aspects of phishing, but it also highlights the need for increased cross-pollination between these research communities to develop more comprehensive anti-phishing strategies. By identifying these key researchers and their thematic focuses, the results map the intellectual structure of the field, enhancing understanding of how specialized contributions can be integrated. Encouraging cross-disciplinary collaboration among these communities can lead to more holistic and effective anti-phishing measures, leveraging diverse insights to address the multifaceted nature of phishing threats.

Overall, the findings enhance the understanding of phishing and the human factor by providing a comprehensive overview of the research landscape, identifying gaps, and suggesting directions for future research. By emphasizing the importance of interdisciplinary collaboration and the integration of human factors into cybersecurity strategies specifically targeting phishing, this study contributes to the development of more effective anti-phishing measures.

5. Conclusions

This article set out to investigate international research on phishing and human factors and provide new insights into this emerging field of study. The research clearly has its limitations in that it is based solely on secondary sources drawn from the Web of Science database, but the authors believe that new ground has been covered, and the four objectives set out in Section 1 have been addressed.

The findings of the study enhance the understanding of phishing and the human factor by revealing the dynamic and rapidly evolving nature of the field, highlighting significant progress in recognizing the importance of human-centric approaches in cybersecurity tailored to phishing threats. Through objective 2, publication patterns and collaboration networks were identified, which are instrumental in improving the field of phishing detection. By recognizing key researchers and institutions, targeted collaborations that combine technical expertise with psychological insights can be facilitated, leading to more effective anti-phishing strategies. Additionally, understanding these networks helps in promoting knowledge sharing, avoiding duplication of research efforts, and accelerating innovation in phishing detection technologies. The global nature of phishing threats calls for more international collaborations to understand cultural variations in phishing susceptibility and develop culturally sensitive prevention strategies. For instance, phishing emails may exploit cultural norms or use language-specific tactics, making international research crucial for developing effective countermeasures. While international collaborations spiked in 2020, the overall number of internationally collaborative articles was low compared to publications with authors all belonging to the same institution. In this context, Carnegie Mellon University stands out as the leading institution in terms of citations and the number of articles published, playing a central role in the research network. Geographically, as shown in Table 6, the USA emerges as the most collaborative country, having the most links with other countries and the highest total number of citations and articles. Table 5 illustrates this fact very well, as most of the top 30 organizations found in the table are located in the USA. Most publications are from universities, with only a few non-academic organizations like Microsoft Research Lab and Defence Science & Technology Group among the top 30 organizations in terms of citations and publication numbers. This suggests there is a need for more diverse perspectives, particularly from industry and governmental organizations. The integration of insights from various sectors can enrich the understanding of phishing and enhance the development of practical solutions.

One possible area concerns the growing integration of artificial intelligence in phishing detection and prevention strategies, combining insights from human behaviour with advanced technological solutions. In addition, the declining frequency of keywords related to specific demographic factors like age in recent years may indicate a need for more nuanced, intersectional approaches to understanding phishing vulnerability. For example, tailoring anti-phishing training to different age groups may improve effectiveness. The relative scarcity of keywords related to the long-term effectiveness of anti-phishing interventions points to another possible area for future research. Given the rapidly evolving nature of phishing threats, long-term studies tracking changes in phishing techniques and human responses over time could provide valuable insights. Future research could also focus on under-represented themes identified in the present study, such as personality, privacy, and cultural influences in cybersecurity. These areas could provide valuable insights into individual differences and their role in phishing susceptibility.

In conclusion, this bibliometric analysis revealed a dynamic and rapidly evolving field of research. Significant progress has been made in recognizing the importance of human factors in cybersecurity, but there remains ample opportunity for further integration of psychological insights with technological solutions and the pursuit of related avenues of research, as outlined above.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/info15100643/s1: Table S1: Metadata of 308 papers published up until 18 July 2024 and listed on the Web of Science database.

Author Contributions: Conceptualization, M.M. and B.M.; methodology, M.M., M.W. and B.M.; validation, M.M., M.W. and B.M.; formal analysis, M.M. and B.M.; data curation, M.M. and B.M.; writing—original draft preparation, M.M. and B.M.; writing—review and editing, M.M., M.W. and B.M.; visualization, M.M.; supervision, B.M. and M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available on the Web of Science database and in the Supplementary Materials File available with this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hadnagy, C.; Fincher, M. *Phishing Dark Waters: The Offensive and Defensive Sides of Malicious Emails*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
2. Canova, G.; Volkamer, M.; Bergmann, C.; Reinheimer, B. NoPhish app evaluation: Lab and retention study. In Proceedings of the NDSS Workshop on Usable Security, San Diego, CA, USA, 8 February 2015.
3. Ahmetoglu, H.; Das, R. A comprehensive review on detection of cyber-attacks: Data sets, methods, challenges, and future research directions. *Internet Things* **2022**, *20*, 100615. [CrossRef]
4. Caviglione, L.; Choraś, M.; Corona, I.; Janicki, A.; Mazurczyk, W.; Pawlicki, M.; Wasielewska, K. Tight arms race: Overview of current malware threats and trends in their detection. *IEEE Access* **2020**, *9*, 5371–5396. [CrossRef]
5. Goel, S.; Williams, K.; Dincelli, E. Got phished? Internet security and human vulnerability. *J. Assoc. Inf. Syst.* **2017**, *18*, 2. [CrossRef]
6. Anti-Phishing Working Group (APWG). Phishing Activity Trends Report, 4th Quarter 2023, APWG. 2023. Available online: https://docs.apwg.org/reports/apwg_trends_report_q1_2024.pdf?_gl=1*1c0xpns*_ga*MTgxODI2O-kzNC4xNzI3MTIxMzQ0*_ga_55RF0RHXSR*MTcyNzEyMTM0NC4xLjAuMTcyNzEyMTM0NC4wLjAuMA (accessed on 20 September 2024).
7. Frauenstein, E.D.; Flowerday, S. Susceptibility to phishing on social network sites: A personality information processing model. *Comput. Secur.* **2020**, *94*, 101862. [CrossRef]
8. Arduin, P.E. A cognitive approach to the decision to trust or distrust phishing emails. *Int. Trans. Oper. Res.* **2023**, *30*, 1263–1298. [CrossRef]
9. Krombholz, K.; Hobel, H.; Huber, M.; Weippl, E. Advanced social engineering attacks. *J. Inf. Secur. Appl.* **2015**, *22*, 113–122. [CrossRef]
10. Aleroud, A.; Zhou, L. Phishing environments, techniques, and countermeasures: A survey. *Comput. Secur.* **2017**, *68*, 160–196. [CrossRef]
11. Yeboah-Boateng, E.O.; Amanor, P.M. Phishing, SMiShing & Vishing: An assessment of threats against mobile devices. *J. Emerg. Trends. Comput. Inf. Sci.* **2014**, *5*, 297–307.
12. Alkhalil, Z.; Hewage, C.; Nawaf, L.; Khan, I. Phishing attacks: A recent comprehensive study and a new anatomy. *Front. Comput. Sci.* **2021**, *3*, 563060. [CrossRef]
13. Ackerley, M.; Morrison, B.W.; Ingrey, K.; Wiggins, M.W.; Bayl-Smith, P.; Morrison, N. Errors, irregularities, and misdirection: Cue utilization and cognitive reflection in the diagnosis of phishing emails. *Australas. J. Inf. Syst.* **2022**, *26*. [CrossRef]
14. Lawson, P.; Pearson, C.J.; Crowson, A.; Mayhorn, C.B. Email phishing and signal detection: How persuasion principles and personality influence response patterns and accuracy. *Appl. Ergon.* **2020**, *86*, 103084. [CrossRef] [PubMed]
15. Halevi, T.; Lewis, J.; Memon, N. A pilot study of cyber security and privacy related behavior and personality traits. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13–17 May 2013; pp. 737–744.
16. Curtis, S.R.; Rajivan, P.; Jones, D.N.; Gonzalez, C. Phishing attempts among the dark triad: Patterns of attack and vulnerability. *Comput. Hum. Behav.* **2018**, *87*, 174–182. [CrossRef]
17. Jensen, M.L.; Dinger, M.; Wright, R.T.; Thatcher, J.B. Training to mitigate phishing attacks using mindfulness techniques. *J. Manag. Inf. Syst.* **2017**, *34*, 597–626. [CrossRef]

18. Arachchilage, N.A.; Hameed, M.A. Teaching Johnny to thwart phishing attacks: Incorporating the role of Self-efficacy into a Game application. In *Machine Learning for Computer and Cyber Security*; CRC Press: Boca Raton, FL, USA, 2019; pp. 337–350.
19. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [CrossRef]
20. Pejić-Bach, M.; Jajić, I.; Kamenjarska, T. A Bibliometric Analysis of Phishing in the Big Data Era: High Focus on Algorithms and Low Focus on People. *Procedia Comput. Sci.* **2023**, *219*, 91–98. [CrossRef]
21. Mutluturk, M.; Metin, B. Mapping The Phishing Attacks Research Landscape: A Bibliometric Analysis And Taxonomy. *J. Theor. Appl. Inf. Technol.* **2023**, *101*, 6758–6780.
22. Munshi, A.; Singla, A.R.; Trivedi, K.J.; Jegede, O.O.; Abodunde, O.O.; Sonkar, S.K.; Kumar, S.; Mahala, A.; Tripathi, M.; Ramkumar, S.; et al. Scientometric-based Knowledge Map of Food Science and Technology Research in India. *J. Sci. Res.* **2022**, *11*, 409–418. [CrossRef]
23. Baker, H.K.; Kumar, S.; Pandey, N. Forty years of the journal of futures markets: A bibliometric overview. *J. Futures Mark.* **2021**, *41*, 1027–1054. [CrossRef]
24. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. Science mapping software tools: Review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1382–1402. [CrossRef]
25. Pranckutė, R. Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications* **2021**, *9*, 12. [CrossRef]
26. Van Eck, N.J.; Waltman, L. VOSviewer manual. *Leiden Univeristeit Leiden* **2013**, *1*, 1–53.
27. Trend, M. *A Constant State of Flux: Trend Micro 2020 Annual Cybersecurity Report*; Trend Micro: Shibuya, Tokyo, 2021.
28. Sheng, S.; Holbrook, M.; Kumaraguru, P.; Cranor, L.F.; Downs, J. Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta Georgia, GA, USA, 10–15 April 2010; pp. 373–382.
29. Parsons, K.; Calic, D.; Pattinson, M.; Butavicius, M.; McCormac, A.; Zwaans, T. The human aspects of information security questionnaire (HAIS-Q): Two further validation studies. *Comput. Secur.* **2017**, *66*, 40–51. [CrossRef]
30. Vishwanath, A.; Herath, T.; Chen, R.; Wang, J.; Rao, H.R. Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decis. Support Syst.* **2011**, *51*, 576–586. [CrossRef]
31. Herley, C. So long, and no thanks for the externalities: The rational rejection of security advice by users. In Proceedings of the 2009 Workshop on New Security Paradigms Workshop, Oxford, UK, 8–11 September 2009; pp. 133–144.
32. Alsharnouby, M.; Alaca, F.; Chiasson, S. Why phishing still works: User strategies for combating phishing attacks. *Int. J. Hum. Comput. Stud.* **2015**, *82*, 69–82. [CrossRef]
33. Williams, E.J.; Hinds, J.; Joinson, A.N. Exploring susceptibility to phishing in the workplace. *Int. J. Hum. Comput. Stud.* **2018**, *120*, 1–13. [CrossRef]
34. Caputo, D.D.; Pfleeger, S.L.; Freeman, J.D.; Johnson, M.E. Going spear phishing: Exploring embedded training and awareness. *IEEE Secur. Priv.* **2013**, *12*, 28–38. [CrossRef]
35. Vishwanath, A.; Harrison, B.; Ng, Y.J. Suspicion, cognition, and automaticity model of phishing susceptibility. *Commun. Res.* **2018**, *45*, 1146–1166. [CrossRef]
36. Wang, J.; Herath, T.; Chen, R.; Vishwanath, A.; Rao, H.R. Research article phishing susceptibility: An investigation into the processing of a targeted spear phishing email. *IEEE Trans. Prof. Commun.* **2012**, *55*, 345–362. [CrossRef]
37. Van Eck, N.J.; Waltman, L. VOSviewer manual. *Man. VOSviewer Version* **2023**, *1*, 1–55. Available online: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf (accessed on 1 June 2024).
38. Van Eck, N.J.; Waltman, L. CitNetExplorer: A new software tool for analyzing and visualizing citation networks. *J. Informetr.* **2014**, *8*, 802–823. [CrossRef]
39. Van Eck, N.J.; Waltman, L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics* **2017**, *111*, 1053–1070. [CrossRef] [PubMed]
40. Mas-Tur, A.; Roig-Tierno, N.; Sarin, S.; Haon, C.; Sego, T.; Belkhouja, M.; Porter, A.; Merigó, J.M. Co-citation, bibliographic coupling and leading authors, institutions and countries in the 50 years of Technological Forecasting and Social Change. *Technol. Forecast. Soc. Change* **2021**, *165*, 120487. [CrossRef]
41. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]
42. Cancino, C.; Merigó, J.M.; Coronado, F.; Dessouky, Y.; Dessouky, M. Forty years of Computers & Industrial Engineering: A bibliometric analysis. *Comput. Ind. Eng.* **2017**, *11*, 614–629.
43. Van Eck, N.J.; Waltman, L. *Manual for VOSviewer*, version 1.6.10; CWTS Univ: Leiden, The Netherlands, 2019.
44. Rezaeian, M.; Wynn, M. Cybersecurity and the Evolution of the Customer-Centric Service Desk. *Int. J. Adv. Intell. Syst.* **2019**, *12*, 147–157. Available online: https://eprints.glos.ac.uk/8007/ (accessed on 11 August 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.