KOME – An International Journal of Pure Communication Inquiry Vol. 11. No. 1. (2023) • 2–27.
DOI: 10.17646/KOME.2023.1.1

How Do Social Media Machines Affect Self-Concept Research?

Systematic Literature Review of the Latest Trends

Katalin Fehér*, Attila Katona I.**

* University of Public Service, Associate Professor, e-mail: feher.katalin@uni-nke.hu
** University of Pannonia, Senior Research Fellow, e-mail: katona.attila@gtk.uni-pannon.hu

Advanced digital technologies broadly penetrate self-activities, such as algorithms, machine learning, or artificial intelligence. This trend is most evident on social media, where contents, attitudes and evaluative judgments meet on technology-driven platforms. Moreover, human networks also started communicating with social bots or conversational interfaces. All these challenges can trigger a redesign of self-concept via technology. Therefore, the paper investigates how social media machines affect self-concept-related academic research. First, pioneers of the field are presented. Second, the self-concept research in digital technology and social media is summarised. Topic networks illustrate critical research fields with the latest trends and future implications. Last but not least, we also investigate how emerging media phenomena affect academic trends in the case of social bots or fake news. The study aims to support the connected research in psychology, business, management, education, political science, medicine and media studies with an understanding of the latest trends. The additional goal is to highlight the potential of market-based research cooperation with academia supporting significant developments and funding.

Keywords: self-concept, social media machine, social media, information technology, systematic review, SMM, online identity, networked self

Introduction

Digital technology and social media are increasingly networking the individual. While there are available discussions about platforms and applications simply for years, a new face of social media has appeared with advanced digital technologies from machine learning (ML) to artificial intelligence (AI) (Alhajj, 2018). This movement results in constant user engagement with black-box machines (Rassameeroj and Wu, 2019), influencing and defining the selves intensively and in various ways (Feher, 2019). Thus, social media machines (SMM) (Newland, 2016) and the powers of data companies are also in the spotlight with privacy issues or bias problems in the increasingly invisible system operation (Tsesis, 2018). It is no coincidence that the Netflix-produced docudrama entitled "Social Dilemma" created a public and academic discourse with AI-driven future selves (Preston et al., 2021). In parallel, social bots and conversational media have started to communicate with human selves (Rheault and Musulan, 2021; Georgakopoulou et al., 2020) to influence them via technology-determined perceptions, beliefs and behaviour, especially on social media platforms (Kušen and Strembeck, 2019). These phenomena result in a specific machine-transmitted human experience which must be studied and understood, along with its consequences.

These changes confirm the necessity of the first summary of self-concept research in the context of social media machines (SMM). Since academic research has moved into a rapidly changing terrain with a relatively short history, only snapshot research is relevant with an exploratory goal. However, the time has come to investigate how academic research reflects social media technology regarding the self. Confirming this statement, a short historical overview of this field is also presented in this study as well as some projected future scenarios. According to this investigation, the related research is rapidly growing. Therefore, one can study which research topics and disciplines of self-concept have become fundamental in social media machines.

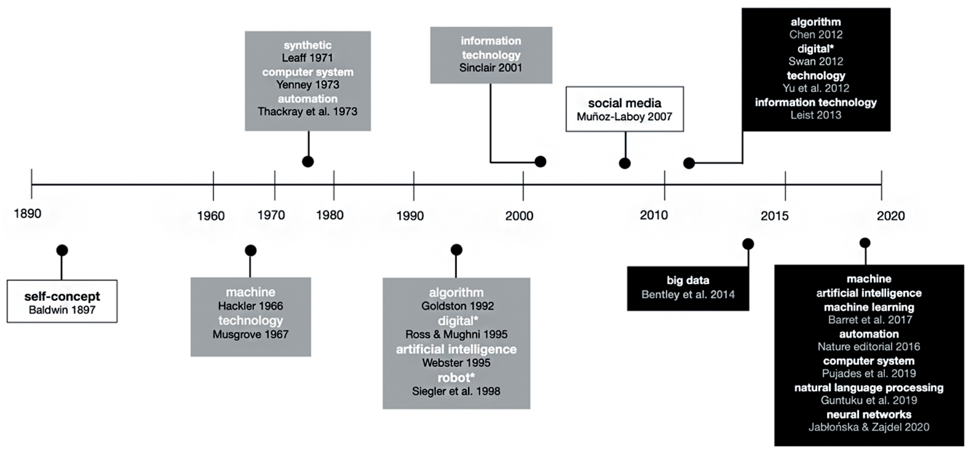
Consequently, the goal is to reveal and interpret the latest academic trends in self-concept research in relation to SMM through the first systematic review. It also aims to support research in psychology, medicine, education, business and their relevant interdisciplines with projected future directions.

Accordingly, the rest of the paper is organised as follows: Section *Background of the study* outlines theoretical considerations with a short historical overview and timeline, while Section *Methods* presents the research goal and methods in detail. Then, Section *Results* shows findings with implications, Sections *Discussion* and *Conclusions* offer a conclusion and projections for the future, and finally Section *Research limits* lists the inevitable barriers to research and assesses their negligible impact on the systematic review.

Background of the study

The number of scientific publications on self-concept has grown dynamically over the last fifty years. According to the leading academic databases such as Scopus and Web of Science (WoS), scientific interest in the topic has grown even more intensively in parallel with the first three decades of digital platforms and the first decade of the social media revolution. Previously, the leading subject areas were medicine, psychology and education. Although medicine and psychology have kept their dominant role in self-concept research, two displacements can be seen. While social science, nursing, business and management, art and humanities, or environmental studies focus more on self-concept than ever, education, sport and rehabilitation science have been even less interested in this field than previous trends. These results do not highlight the specific role of communication and media studies, computer science, or engineering. Nevertheless, they can be assumed with emerging social media developments. However, Scopus and WoS analytics point out that social sciences already contain these emerging disciplines in self-concept research. All these results confirm the changing trends in self-concept research and the relevance of the research goal.

Moving forward with these results, social media and technology-related publications were filtered resulting in noticeable milestones of self-concept research on a timeline (Figure 1).



1) White rectangle=self-concept & social media topics started to be published in academia
2) Grey rectangle=self-concept AND the type of the technology
3) Black rectangle=self-concept AND social media AND the type of the technology

Figure 1: Self-concept research on the timeline with technology and social media
Source: Compiled by the author.

The first academic publication on self-concept appeared more than a century ago, as Baldwin (1897) wrote on the topic of organisational behaviour. However, technology was not relevant to the early studies. Research in self-concept started to explore technology and machines in the 1960s, followed by automation and computer systems in the 1970s. The initial studies presented mostly psychometric and self-perception testing along with ergonomics and teaching machines (Hackler, 1966), (Musgrove, 1967; Thackray et al., 1973; Yenney, 1973). The adjective “synthetic” was also first mentioned at this stage as a synonym of “artificial” to study technology-related ego development (Leaff, 1971). These first milestones represent how digital technology began to be incorporated into the research of the self.

In the 1990s, digital systems, algorithms, AI, and robotics presented the next noticeable milestone in self-concept research (Ross and Mughni, 1995; Goldston et al., 1992;

Webster, 1995; Siegler et al., 1998). The initial publications focused on social perception, self-esteem and adaptive behaviour. Therefore, individuals, social aspects and emerging technologies were also relevant for academic research. The perception of self-motion and the use of digital displays were also studied at this stage, assuming the growth and deepening of human–machine interrelations. These directions led to the emergence of an umbrella term, "information technology", highlighting the first medical informatics for self-concepts (Sinclair, 2001). This initial perspective represented the options for the research of digital health services. In a broader sense, it was a signal of the changing trends in self-research with emerging (inter-)disciplines.

After this first era, social media developments introduced a new technological approach from network science to big data analysis (Tinati et al., 2014). The introductory journal paper was published in 2007, highlighting cultural, generational and ethnic issues in digital networks (Muñoz-Laboy et al., 2007). Through the exploration of this topic, social media entered the stage of academic publications on self-concept. The technological background was even less pronounced in this case, and remained so for some time. Since 2012, digital and information technology (Leist, 2013; Catherine et al., 2012; Swan, 2012), algorithms and big data were highlighted technologies in social media and self-concept research (Chen, 2012; Bentley et al., 2014), with discussion mostly focusing on human–machine interactions, personalisation and health communication. Over the past few years, advanced digital technologies have simultaneously been explored for self-concept research with computer systems, machine learning, artificial intelligence, automation, natural language processing, or neural networks (Pujades et al., 2019; Barrett et al., 2017; Editorial, 2016; Guntuku et al., 2019; Jabłońska and Zajdel, 2020). The initial publications discussed body image, mental health, online communities, social comparison, web-based self-management and self-reflection. These trends describe not only widespread applications of advanced digital technologies, but also even more diverse approaches to self-concept. Thus, this milestone highlights the relevance of the concept of "social media machines" to interpret future selves.

To sum up, the introductory fields of SMM, academic research in self-concept has begun to explore the technological operation behind social media for even more fields and disciplines. Psychology is key in this field regarding well-being and social needs (Thomas et al., 2021). Although specific topics are not yet on the timeline without relevant records, such as "robot" or "deep learning", we also expect these topics to come up in the near future. The summarised milestones point out that the changing digital transformation (Vial, 2019) has started to affect topics and research.

According to this timeline-based summary, the greatest change has been in the last few years with several emerging technologies, and the densest period is between 2016 and 2020. With the mean of this period, our research focused on the last three years to find the latest trends.

There are probably two reasons for these trends. On the one hand, more than fifty percent of the total population already has access to social media platforms (Hootsuite, 2021), facilitating self-related studies in numerous ways. On the other hand, social media technology has changed dramatically over the last few years. For example, the change is obvious with Facebook AI or Twitter Sentiment Analysis through machine learning.

They face noticeable challenges currently, such as using biased datasets (Houser, 2019) or privacy issues. With another example, the human–chatbot relationships are spreading, but little knowledge exists on how these connectivities influence the social context of the users and their emotional and social values. Even if a specific study has already reported positive and mixed impacts and suggested an initial model. These all affect academic research in self (Skjuve et al., 2021).

What is the outlook for self-concept research with these changes? First, if even more people join social media applications and the social media machines drive human perception and activities, social and psychological constructions will be significantly designed by technology. Second, this is especially critical for the phenomena of misinformation technology that can shape individuals' attitudes (Colliander, 2019) or allow fake accounts to result in inauthentic behaviour (Mazza et al., 2022). This field still appears in a small proportion in the case of self-concept research. However, self-expression or self-efficacy are already affected by them (Hilliard et al., 2015; Gesser-Edelsburg et al., 2018). Third, social bots, non-social interactions, or conversational media (Ferrara et al., 2016; Rheault and Musulan, 2021) also have a great potential to influence the technology-driven self. Fourth, quantified self or self-tracking allows the development of the profiles of the users (Puntoni et al., 2021; Druga et al., 2017; Neff and Nafus, 2016; Sadowski, 2019) for personalisation and predictive analysis. All these options augment the self, resulting in a broader spectrum for self-concept research.

Last but not least, what is the meaning of "self-concept" in this context? If cognitive structures started to be translated for AI, neural networks, machines, or deep learning, the suitable definition of self-concept is "cognitive structures with content, attitudes or evaluative judgments to make sense of the world" (Oyserman and Markus, 1998). However, this connected but older definition should be rethought using the research history summarised above, the possible future scenarios, and also, the interpreted results below.

Following all these considerations, a systematic literature review of self-concept research is presented in the context of social media machines with the latest trends.

Methods

The above-presented background (Section *Background of the study*) implies that self-concept research started to emerge in the last few years in SMM with diverse fields in SMM. Accordingly, the goal is to discover and interpret these diverse academic research topics over the last five years. Therefore, four research questions were formulated as follows.

- RQ1. What are the key topics of self-concept research?
- RQ2. What are the key research topics of self-concept research in the case of social media?
- RQ3. How is self-concept studied along with social media machines?
- RQ4. What are the latest trends of self-concept research in the context of social media machines?

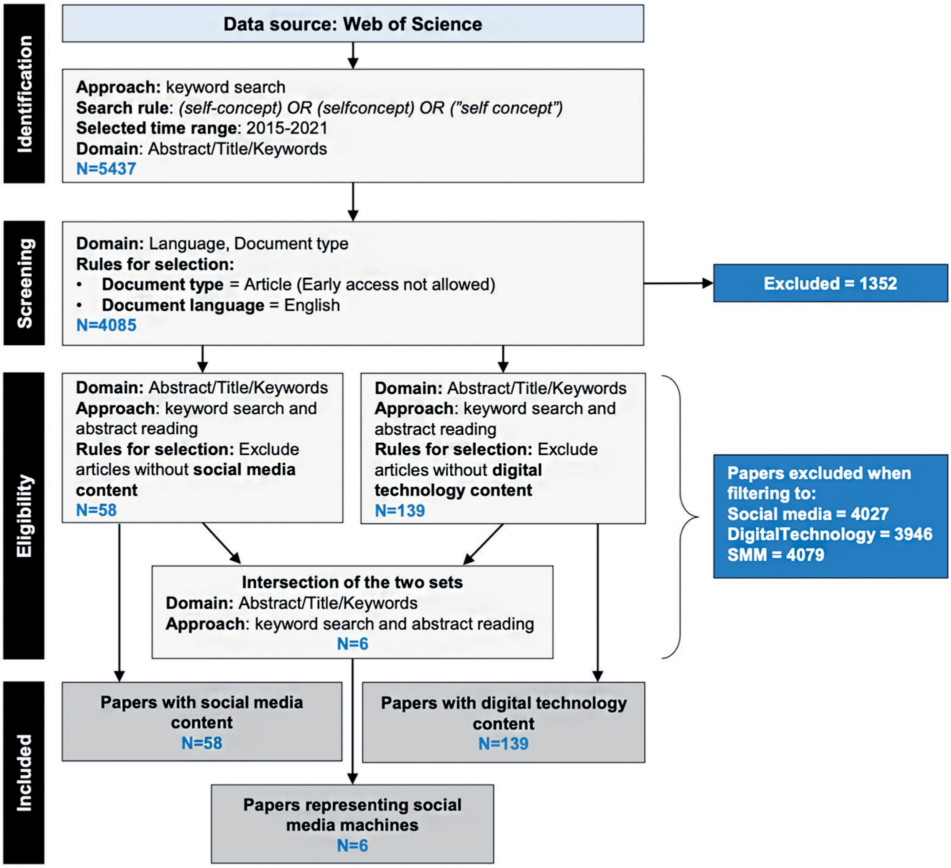


Figure 2: Data collection based on the PRISMA methodology
Source: Compiled by the author.

To answer these research questions, this study applied the PRISMA methodology. It supports a literature review with a systematic method using four categories as: 1. "identification"; 2. "screening"; 3. "eligibility"; and 4. "included". PRISMA was proposed by Moher and his colleagues (Moher et al., 2009) for the careful analysis of selected sources in one corpus.

PRISMA methodology

In this paper, several steps were applied from data collection to synthesising research findings. Figure 2 shows the flowchart as a structure of data collection and analysis.

1. *Identification.* As the first step of the identification, the three key academic databases were tested: Scopus, Dimensions and WoS. Selecting the database for this study, two primary considerations were taken into account. On the first hand, the research

aimed not to compare academic databases, as this would require a separate study. On the other hand, the selection was based on which database could provide the most relevant results. Testing all available options, three arguments were considered in favour of using the Web of Science (WoS) database. First, we had to test how many download options could be performed simultaneously, as the volume of hits and records for the same time range can change due to items added afterward. Our database tests indicated that WoS allows the download of the most significant data sets, and a one-time snapshot filter resulted in the most accurate data. Second, WoS datasets are known to support basic research (Stahlschmidt and Stephen, 2020), which is helpful in the case of research utilising secondary data for a literature review (Johnston, 2014). Third, WoS is a valuable tool for conducting a comprehensive and rigorous literature review, particularly in the case of citation searches in social sciences and humanities (Zahedi and Haustein, 2018).

According to Section *Background of the study,* it was obvious to investigate a five-year time interval. To conduct the search, a keyword search was applied considering the abstract, keywords and title of the papers with the search rule: (self-concept) or (selfconcept) or ("self concept"). This step resulted in 5,437 records. As a comparison, Scopus and Dimensions.ai were also tested with the same searching parameters but they gave nearly the same results as Web of Science.

2. *Screening.* To further screen the result set, the data was filtered to keep only refereed journal articles as "the gold standard" (?). Only English as the language of the papers was allowed during the screening. The world language results in the highest number of sources without translation and abbreviation anomalies. It must be noted that the specified language criteria only applied to the language of the abstract, title and keywords. Thus, papers were allowed where full text was written in another language. This decision supports the minimisation of data loss. To perform the screening step, the filtering options provided by the Web of Science online platform were applied.

3. *Eligibility.* Our research questions refer to social media, digital technology and social media machines. Therefore, during the "Eligibility" phase, our aim is to generate three subsets of data to allow us to analyse these fields separately to answer the research questions. Filtering (keyword search) was thus applied first to find papers associated with the research fields mentioned above, as follows:

Within the screened dataset:

1. find the set of papers with respect to social media research
2. find the set of papers with technology-related content
3. identify the intersection of the two sets

First (step 1), a keyword search was applied across all the records and kept only those papers where the string "social media" was an existing substring of the abstract, title, or keywords. The filtering was also refined by manually reading these three data fields (abstract, title, keywords). In case of this subset, 4,027 papers were excluded from the original dataset and included only 58 articles associated with the field of social media.

As per step 2, to identify the subset of papers representing the digital technology research field, another keyword search (and refinement via reading) was performed. In the context of the process described, “digital technology content” was a designated category used to manually refine keywords in order to filter the intersection between social media and technology-related content. It was applied independently from the exclusion criteria, and its purpose was solely to improve the precision of the keyword search. In this case, the terms “technology”, “information technology”, “big data”, “algorithm”, “artificial intelligence”, “AI”, “machine learning”, “automation” and “neural network” were used as keywords, and 139 papers were found with technological content.

Finally, the intersection of the two sets was determined. Here keyword search was applied to include only papers containing both social media and digital technology-related keywords. The keyword search results were refined by reading the title, abstract and keywords of the papers. After the refinement, six papers were identified as the third subset, i.e. the papers regarding social media machines.

4. *Included.* In the analysis, we include three subsets of papers mentioned above. Due to their sizes, social media and digital technology-related subsets were analysed by quantitative analysis, while papers on social media machines had qualitative analysis performed.

Methods applied in the quantitative analysis

Text cleaning

In this paper, text analysis was conducted, and therefore, the cleaning and preprocessing of the textual data is the precondition of further analysis. During the text cleaning first, basic text cleaning was applied such as lowercase transformation, removal of punctuation marks, removal of special characters and tokenisation. English stopwords are also removed with the extension of a self-defined custom stopword list to avoid the strong impact of commonly used terms within the abstracts like “research”, “study”, “result”, “analysis”, etc. As a next step, POS (part of speech) tagger is used to eliminate words that correspond to a specific part of speech, such as verbs or adjectives. Finally, lemmatisation was also applied to reduce inflectional forms of words to the base form. The R software was used to perform the lowercase transformation, punctuation mark removal, tokenisation and stopword removal. To perform the aforementioned transformations, the “tm” (Feinerer et al., 2008), “tidytext” (Silge and Robinson, 2016), “dplyr” (Wickham et al., 2021) packages were used. POS tagging and lemmatisation were conducted in Python, using the “spacy” (Honnibal and Montani, 2017) and “nltk” (Bird et al., 2009) libraries.

Latent Dirichlet Allocation (LDA)

To reveal the latent structure of the abstracts, Latent Dirichlet Allocation (LDA) approach was applied (Papadimitriou et al., 1998; Chen et al., 2017; Maier et al., 2018; Jeong et al., 2019).

Before describing the process of LDA, its principal concepts, such as corpus, documents and terms need to be discussed. Corpus denotes the collected set of text data, specifically the concatenated string of title, keywords and abstract in the current case. A document represents a given element of the corpus, which is the concatenated string of titles, keywords and abstracts related to a given paper. Finally, terms refer to the words of the documents.

LDA considers two types of distributions during the process: 1. the distribution of documents over the hidden topics; and 2. the distribution of words within each topic. If the number of documents within the corpus is denoted by D, the number of desired topics is denoted by K, and V is the dictionary of the terms across the corpus, the LDA process can be described with the following steps (Jelodar et al., 2019):

1. For each topic $k(k \in \{1, ..., K\})$ choose a word distribution $\overrightarrow{\varphi}_t \sim Dir(\beta)$
2. For each document $d(d \in \{1, ..., D\})$ choose a topic distribution $\overrightarrow{\theta_d} \sim Dir(\alpha)$
3. For each word $w(w \in \{1, ..., N_d\})$ in each document d:
   i. Select a topic z_n from $Multinomial(\overrightarrow{\theta}_d)$
   ii. Select a word w_n from $Multinomial(\overrightarrow{\varphi}_{zn})$

Where N_d denotes the number of terms within d^{th} document, *Dir(α)* and *Dir(β)* are Dirichlet distributions with *α* and *β* distribution parameters, respectively. θ distributions with *α* and *β* distribution parameters, respectively. *θ* and *φ* are multinomial distributions from Dirichlet distributions. The hyperparameters are *T, α* and *β* which need to be selected by the researcher (Hou-Liu, 2018).

To select the optimal number of topics (K^*), three widely used metrics were applied:

1. Griffiths and Steyvers (2004) propose to select K^* where the harmonic mean of sampled log-likelihood values is maximal. The samples are retrieved by Gibbs-sampling in this approach.
2. Another approach proposed by Cao et al. (2009) aims to minimise the average cosine similarity of topic distributions.
3. The method developed by Arun et al. (2010) minimises the symmetric Kullback-Liebler divergence between *θ* and *φ*.

t-Distributed Stochastic Neighbor Embedding

t-Distributed Stochastic Neighbor Embedding, or t-SNE for short aims to preserve the local neighborhood structure from a high dimensional space. It is a widely-used technique to visualise multidimensional data (Cao and Wang, 2017). Let $\{x_i\}_{i=1}^{n}$ represent the high dimensional data points and $\{y_i\}_{i=1}^{n}$ the low dimensional points. First, it defines the probability of choosing a pair of points (p_{ij}) in the high dimensional space that can be described as the symmetrized conditional probabilities $p_{i|j}$ and $p_{j|i}$ (Kruiger et al., 2017; Cao and Wang, 2017):

$$p_{ij} = p_{ji} = \frac{p_{i|j} + p_{j|i}}{2n} \tag{1}$$

where $p_{i|j}$ is given by normelized Gaussian distribution as follows:

$$p_{i|j} = \frac{exp(-\frac{d_{ij}^2}{2\sigma_i^2})}{\sum_{k\neq j} exp(-\frac{d_{jk}^2}{2\sigma_i^2})}, p_{i|i} = 0 \tag{2}$$

Furthermore, the probability of choosing a point-pair in the low dimensional space (given by normalized Student's t-distribution) is determined as follows:

$$q_{ij} = q_{ji} = \frac{(1 + ||y_i - y_j||^2)^{-1}}{\sum_{k\neq l}(1 + ||y_k - y_l||^2)^{-1}} \tag{3}$$

Finally, the position of data points can be found in the low dimensional (output) space by minimizing the Kullback-Leibler divergence between q_{ij} and p_{ij} probabilities:

$$C_{KL} = \sum_{i\neq j} p_{ij} log \frac{p_{ij}}{q_{ij}} \tag{4}$$

In this paper, t-SNE was used to visualise the topic modelling result by transforming the multidimensional output of topic probabilities into a two-dimensional space.

Hierarchical clustering

In this paper, hierarchical cluster analysis was also used to characterise the structure of the extracted topics with respect to self-concept studies having social media context. Hierarchical clustering algorithms aim to search for nested clusters within the population following either an agglomerative or a divisive approach. The steps of the method can be described as follows (Govender and Sivakumar, 2020):

1. Each data point is considered a single cluster.
2. Cluster distances are computed.
3. A pair of clusters having minimum distance based on a given measure are combined and replaced by a single cluster. The distance matrix is recomputed afterward.
4. Steps 2 and 3 are iterated until every data point is contained by one cluster.

$$H(p,q) = \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{N_v}(\sqrt{p_i} - \sqrt{q_i})} \tag{5}$$

where N_v denotes the number of terms in vocabulary V, p and q represents the probability distributions of two extracted topics.

Results

The latest trends of self-concept research

The LDA method was applied to find the key topics of self-concept research. As the first step, simulation was conducted to identify a starting point for the of topics). The number of topics was increased iteratively, and the metrics described by subsection *Latent Dirichlet Allocation (LDA)* were calculated in each iteration. Figure 3 shows the simulation results.

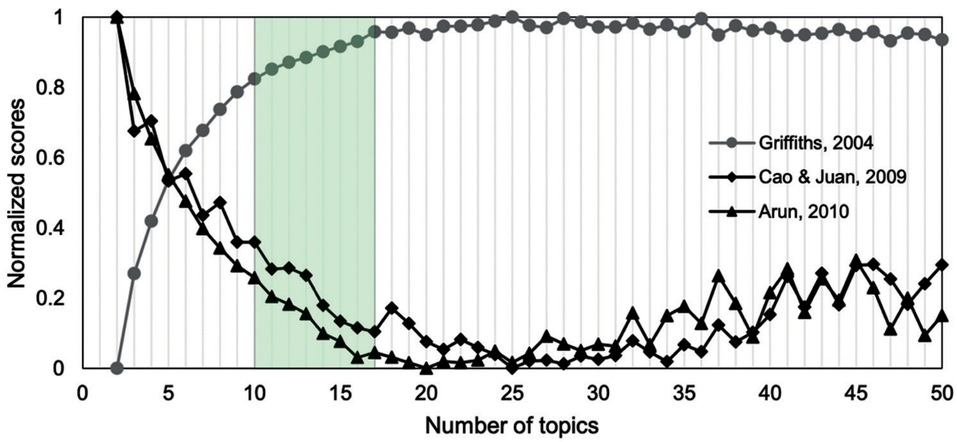


Figure 3: Optimal topic number selection (simulation)
Source: Compiled by the author.

As the results suggest, K^* should be selected within the 9–17 interval as a start. After a manual investigation of the topic detection results found that the most interpretable topic structure is provided when $K^* = 9$.

Answering RQ1, this study has discovered a total of nine extracted topics with issues for individual level from childhood to gender identity, and also, with social issues from education to moral questions. Although the method generated an optimal topic number and the extracted topics are mostly congruent, the resultant keywords present cross-cutting themes in a few cases, such as research in schooling and social or identity. Thus, we manually studied the papers behind these results to define nine congruent topic labels for the keyword selection. These topic labels confirm a broadly investigated self-concept research from the physical image to mental health, from social-moral identity to brand-related consumers, and from educational challenges to cognitive skills. The nine topic labels in Table 1 represent the latest trends in academic self-concept research.

Table 1: Extracted topics regarding self-concept

No.	Keywords	Topic label	No. of papers
1	patients, health, treatment, depression, life	mental health	368
2	children, adolescents, social, school, family	children and adolescents	440
3	social, cognitive, memory, task, feedback	cognition and recall	374
4	physical, body, activity, children, intervention	body image	342
5	learning, education, career, social, teachers	education	543
6	identity, leadership, moral, social, role	social identity	395
7	sexual, personality, clarity, identity, women	gender and sexuality	402
8	academic, achievement, school, mathematics, reading	academic self-concept	879
9	brand, consumers, social, consumer, consumption	consumption and brands	342

Source: Compiled by the author.

Interpreting the records, two main categories were found. First, social roles are presented for who is the self in a given context, be they a patient, a child, an adolescent, a teacher, an academic person, a leader, or a consumer. From this list, patients, children, or adolescents are traditional categories for self-concept research. "Academic" topic is obvious if research projects have been sampled with peers or students in higher education research. Additionally, a specific research field was also found manually, focusing on an academic career with motivation, achievements and self-efficacy. We can also see that "leader" and "consumer" are emerging research fields, mainly in the context of online social networks and digital platforms. This trend is expected to grow in the future. Second, the identity, gender, personality, sexuality, body, or health of individuals are also presented along with their goals, skills, achievements, interventions, treatments, or moral issues. As can be seen, none of these keywords highlight the elements of the self or specific emotional themes, such as self-consciousness, shyness, guilt, shame, or self-disclosure (Buss, 2001). Instead, self-concept is broadly researched in diverse fields, primarily in psychology, clinical psychology, medicine, educational studies, marketing and management. These results assume inter- and multidisciplinary approaches, which is significantly represented by the most cited paper from the 4,085 records with a social, cultural and cognitive mindfulness program (Schonert-Reichl et al., 2015).

Considering this broad landscape of self-concept research, our next goal was to identify the interconnected topic areas of the nine extracted topics to reveal their position to each other and interpret their more or less connected fields. To visualise the structure of the nine extracted topics, we applied the t-SNE dimension reduction approach as described in subsection *t-Distributed Stochastic Neighbor Embedding.* Figure 4 shows the results.

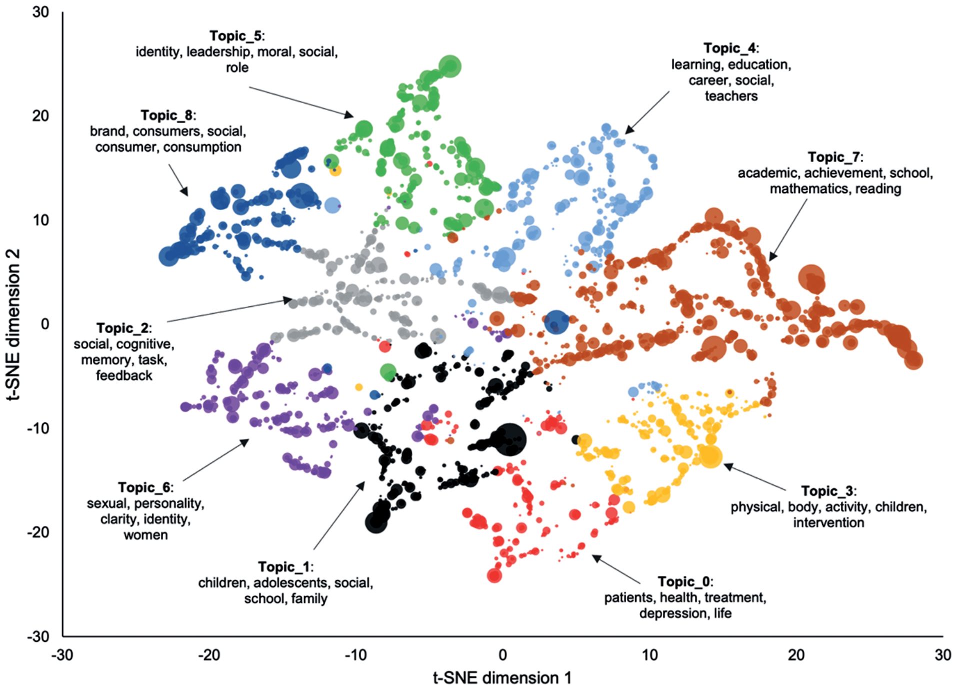


Figure 4: Distribution of papers in the two-dimensional space based on t-SNE
Source: Compiled by the author.

In Figure 4, the axes represent the reduced two output dimensions given by t-SNE. Each circle represents a paper, and the size of the circles reflects the total number of citing articles by each paper.

The most interconnected and cross-linked fields are “mental health”, “children and adolescents” and “gender and sexuality”. Investigating the papers behind these results, primarily self-perception issues and various treatments are discussed. For example, traumatic brain injury, emotion regulation, or sexual well-being are highlighted. The physically perceptible body and the human mind along with emotions are studied together in this way. Several interdisciplines explore these topics, primarily medical humanities or social psychology. In the figure, larger circles indicate topics with a higher number of citations. By far, the most cited article is about cognitive and social-emotional development through a mindfulness-based program (Schonert-Reichl et al., 2015) representing interdisciplinary research in applied social science and psychology with funding behind it. This article is the most referenced record from all extracted topics and also from the whole database with 213 citations. This result suggests that the authors are already well-known in their field with a high citation index. Also, they worked with an interdisciplinary and trending topic to cultivate wellbeing. The further articles are less cited, and they present diverse research topics from victimisation to pathological gambling.

The broadest interconnected fields are "cognition and recall", "academic self-concept" and "consumption and brands" with direct connections to "children and adolescents" or "gender and sexuality". The key topics in the papers are "belonging" and "performance", as the general dynamics of the self-concept. The manual scanning revealed diverse research fields from sexual minorities to luxury consumption. The related emotions or motivations are studied on an actual or symbolic level. Numerous interdisciplines are affected in this way, resulting in a focus on a measurable or quantified self. Even the most referenced topic label of the extracted topics with "academic self-concept" is in the direction of measurability, focusing on the issues of performance, intelligence and gender (Pekrun et al., 2017; Roth et al., 2015; Gaspard et al., 2015).

The most integrated field is "consumption and brands" with topics of organisational communication or consumer society. "Body image" for younger generations is also discussed partly in the consumer context reading the abstracts in detail. Specific fields are discussed in this field with a significant role of the "organisational psychology" interdiscipline. However, the most cited record presents adoption behaviour and consumer preference in mobile healthcare service systems (Dwivedi et al., 2016). The multidisciplinary article represents social and behavioural psychology, marketing, ICT and medicine, as well as epistemological and ontological paradigms for policymakers. The multidimensional approach and the research using technology point out the presumable future strategies of the authors to find the broadest possible audience with an innovative topic. Considering the result, further intensive expansion of business, management and marketing studies is expected in self-concept research, primarily recognised in the case of inter- and multidisciplines.

In conclusion, mostly congruent and partly interconnected topics are detected with broadly defined self-concept research. The traditional areas have kept their key topics with social or gender identity and younger generations and cognitive studies, consumer psychology and academic self-concept. Emotional topics are less presented in topic labels but more visible in line with diverse research fields. According to recent trends, self-concept research will be expanded mostly in the business and management sciences. However, psychology or medicine will still be the leading disciplines with their interdisciplines. As the most cited records, mindfulness, academic performance, or body image-related consumption can also emerge, especially in multidisciplines.

The latest trends of self-concept research in social media machines

After recognising current trends in self-concept studies, answer RQ2 and RQ3 were at the forefront, namely, to reveal the related and latest research trends in social media and SMM. Since a significant number of social media research on self-concept was found but just small data on the topic of social media machines, we interpreted the results in three steps. First, "social media" and "self-concept" research are summarised with 58 records.

Second, small data of SMM are revealed with six records. Third, "digital technology" and "self-concept" research are outlined together.

In the first step, specific extracted topics of "social media" and "self-concept" research were revealed by LDA for the 58 papers. Similarly to Figure 3, a simulation was conducted where we found that five topics can describe this subset well. Table 2 shows the top terms and labels related to the extracted topics.

Table 2: Extracted topics regarding self-concept and social media

No.	Keywords	Topic label	No. of papers
1	online, content, presentation, identity, people	online identity	10
2	online, satisfaction, relationship, individual, networking	networked self	10
3	sexual, adolescent, selfie, FOMO, behaviour	online behaviour	11
4	brand, consumer, fashion, brand love, love	brand engagement	13
5	political, role, activity, company, place	political issues	14

Source: Compiled by the author.

Answering RQ2, five topic labels were revealed with networked self and identity, online behaviour, as well as studies in brands and political communication (Table 2). The number of extracted topics was definitely narrowed from self-concept research in itself. This result may be evident, but numerous keywords just disappeared from the core keyword list, and new ones appeared. In detail, issues of individuals are not presented with topics of "children", "gender" or "patient", as well as social roles in education. Thus, most of the traditional key topics on self-concept research vanished from the extracted topics. Only three key fields of individuals remained with online behaviour and identity, and networked self. Scanning the records behind the data manually, these are studied mostly with issues in online congruent self, self-construction, or self-esteem. The topic of "body image" was also narrowed to sexuality and fashion. Primarily positive emotions are highlighted with these extracted topics, such as satisfaction, brand love, or brand engagement in relation to corporate brands and politicised behaviour. FOMO ("fear of missing out") is the only emotion without a positive meaning in this summarisation.

Interpreting the results briefly, the compact topic labels represent less traditional topics of self-concept research. The focus is on the online identity and networked self in the latest research trends. The connected emotions are more available than presented in the general self-concept research above. Further research fields are expected in consumption, while online political behaviour is also an emerging research field. The categories of "presentation", "networking", "selfie" and "FOMO" describe the social media-related fields, highlighting the challenges of the self on social platforms. These self-reflective aspects are also confirmed by the record with the highest citation number in online self-presentation and self-development (Yang and Brown, 2016).

Considering all these results, the directly connected disciplines are psychology, marketing and management, political science, organisational studies, network science, media and communication and their inter-disciplines.

Interpreting the results with topic hierarchy, we applied agglomerative hierarchical clustering as described by subsection *Hierarchical clustering.* Figure 5 shows the resultant cluster dendrogram.

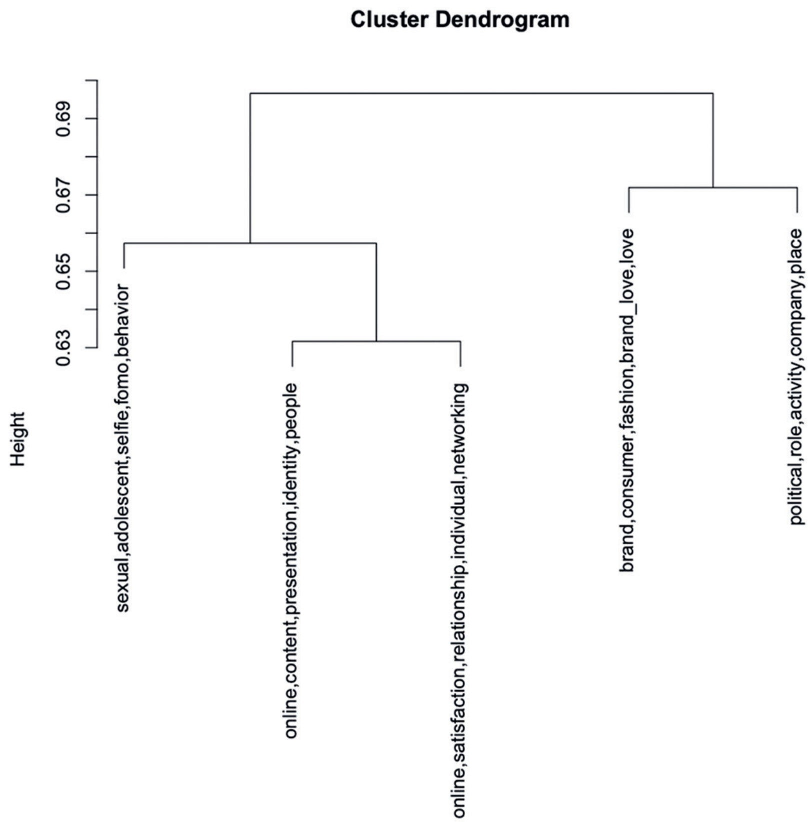


Figure 5: Hierarchy of self-concept research in the context of social media
Source: Compiled by the author.

The dendrogram visualises and clusters taxonomic relationships of the extracted topics. According to the results, marketing and political science are strongly correlated fields, presenting directly connected subtopics in the papers, such as self-expression, microtargeting to citizens, preferences and emotional engagement. The other part of the dendrogram on the left confirms the key dimensions summarised above for individuals with their representative and networked online self or identity. Thus, the dendrogram confirms the five topic labels of this study in social media and self-concept with their distribution.

Moving forward to RQ3 and its answer, only six papers represent social media machines, namely topics both with social media and technology in self-concept research. This result was unexpected but understandable. Only the last five years were studied, and significant change has just recently started for social media services. During this short time, academic research did not have enough time for multiple reflections in terms of research and publication. However, the number of related papers is expected to grow and technology-focused research is also assumed. The first papers of pioneers are available in Table 3.

Table 3: Self-concept research in social media machines: The pioneers

No.	Article	Digital technology	Social media	Disciplines	Funding	Citations
1	(Zarouali et al., 2020)	profiling algorithm	social media marketplace	political and behavioural science, new media studies	no	6
2	(Chen, 2019)	big data, machine learning	social network analysis	information and behavioural science, marketing	ministerial funding	1
3	(Thomas et al., 2019)	big data	social media content	psychology and cultural studies	no	1
4	(Tseng and Hsieh, 2019)	emoticon driven technology	mobile instant messaging	psychology, information and behavioural science	no	11
5	(Choi and Behm-Morawitz, 2018)	smart devices	networking sites	psychology and digital literacy	N/A	15
6	(Claffey and Brady, 2017)	artificial intelligence bots	firm-hosted virtual communities	psychology and marketing	no	30

Source: Compiled by the author.

The six academic papers represented are not valid for far-reaching conclusions. However, the first relevant articles reveal two perspectives with small data. First, research on individual or social self-concept has become available to consumers, political voters, and virtual or brand communities. This result confirms the expectation of emerging research in psychology and marketing or behavioural, information and political sciences.

Second, social media technology supports research in political microtargeting, personality profiling, smart devices and emotionally charged ads. This direction allows the discovery of the quantified selves with their behaviour and identity via platforms and applications. Additionally, the most cited work (Claffey and Brady, 2017) presents artificial intelligence bots for firm-hosted virtual communities. This result confirms the relevance of the research in conversational media with personal assistants and chatbots, as was assumed in the introduction.

Further details behind these records should also be considered with information about funding or citations. An average of eight citations per article was found, and only one of the records presents a granted research. Both results confirm a preliminary stage of self-concept research in SMM, albeit without serious support or attention. As the concept of SMM is being introduced in parallel, there was a decision to investigate the digital technology-connected research separately as well.

To extend the interpretation of small data and answer RQ3 with more detail, the social media topics (58) and the digital technology topics (139) are projected onto the original nine extracted topics of self-concept. Figure 6 shows the results.

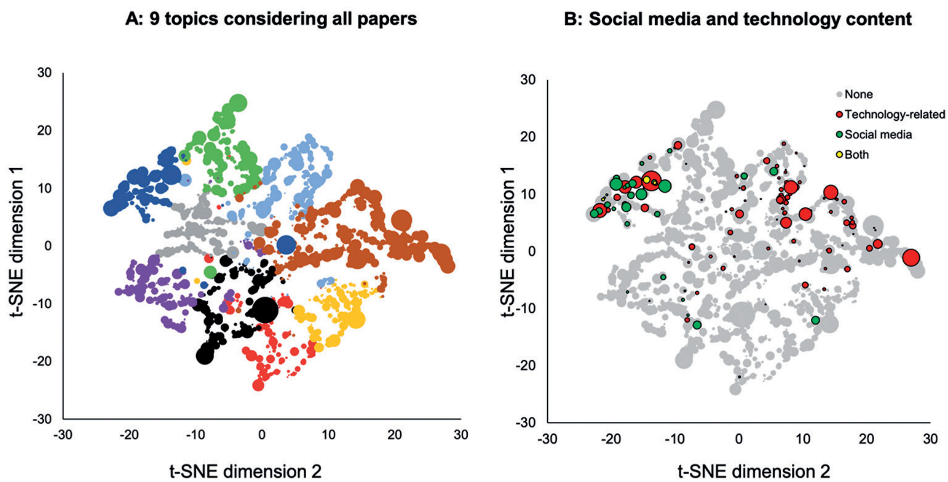


Figure 6: Embedding of technology and social media-related papers
Source: Compiled by the author.

Figure 6A shows the original pattern as a reference, while Figure 6B highlights the location of the referred subsets of papers. Red circles represent the technology-related papers (139), green circles denote the papers with social media content (58), and yellow circles show the intersection of the two sets (6).

According to the results, self-concept research with a social media focus is well connected to consumption, brand love, or organisational communication, mostly in business and management studies. The topics of body image, social identity, children and adolescents are also clearly affected fields. This result confirms the five topic labels of social media research with behaviour-based online identity and networked self. This result is also consistent with the review article of Hollenbaugh about research in self-presentation (Hollenbaugh, 2021).

However, the number of articles focusing on digital technology is more than twice compared to records of social media research. Therefore, digital technology is a more spreading field for the extracted topics, confirming "machine design" for social media and the "user" category for the self. In digital technology, the most connected topics are also consumption, brand love or organisational communication and social media research. Therefore, these fields are expected to be the most investigated in the context of social media machines.

Additionally, technology-related research in the self-concept is also demanding in education and learning performance in different studies. This could even be due to methodological changes supported by technology, but after scanning the 139 articles manually, this assumption is not confirmed. Instead, the records in education and learning performance mostly discuss topics in human–machine relations, recalling the timeline details in section *Background of the study* with ergonomics, teaching machines and digital displays. Nevertheless, an exploration of SMM-based methodology developments is expected.

Considering these results and the original background (section *Background of the study*) of this paper, a research responsibility is detected in self-concept research for trustworthy technologies. This implication is particularly valid if research of misinformation and fake media finds its place in the investigation of self. Only one record is available about fake media or misinformation in the whole database of 4,085 records (Colliander, 2019). Even if the verification of multimedia content has become a crucial issue, mostly in the case of user-generated content (Varshney and Vishwakarma, 2022). However, it could be a notification of an unexplored but critical research field. The research responsibility is more valid if "personalisation" also appears in the database with influences on human behaviour. These topics certainly play a key role in the research of data companies. However, this result has found an almost missing area for academic research.

Assuming the non-human participation in self-related interactions, the topics of social bots, human–bot communication, AI bots, chatbots, personal assistants, or other conversational media were manually filtered out. A few records already represent this research field (Nathanson, 2017; Claffey and Brady, 2017; Anshar and Williams, 2016). Despite the widespread adoption of these technologies, only pioneers have found a connection to self-concept research so far. It is important to highlight that these technologies started to imitate the self, and synthetic selves have started to influence human existence. This unexplored but fundamental field defines the human self-concept even more intensely. Therefore, the relevant research should be expanded reasonably quickly.

Accordingly, two implications have become available in line with the results of digital technology. First, traditional and also emerging self-concept topics should be explored with digital technologies and their consequences. Second, if scientific research and data companies collaborate in research of self-concept influence, it is a responsible way to develop trustworthy SMM.

Discussion

Although the applied database does not comprehensively cover the self-concept research, the latest trends with the background give us a comprehensive summary. Thus, the self was summarised as a concept with the most represented topics and the most cited records (Table 4).

Table 4: The latest trends of self-concept research in social media machines

Category	The self as a concept	Key topics	Most cited records
Digital Technology (139)	user and consumer	brand engagement, organizational communication, education, learning performance, human–bot interaction	Dwivedi et al. (2016) (127)
Social Media (58)	online identity and networked self	online behaviour, brand engagement, organizational communication, political influence	Yang and Brown (2016) (52)
Social Media Machine (6)	user, citizen, consumer	personalization, profiling, micro-tagging, emotionally charged ads, AI bots	Claffey and Brady (2017) (6)

Source: Compiled by the author.

This summary clearly reveals the current concepts of the selves in relation to SMM with user-based identity, networked self, consumer, or citizen. Interpreting the detailed results with interconnected research fields, organisational communication, political microtargeting and learning performance are trending. The most cited records highlight only the consumption and marketing research with keywords of customer engagement, consumer adoption, planned behaviour, user acceptance, or identity development.

Considering all these results, the highest potential of the latest trends is also available for research in business and politics with interdisciplines of the leading field of psychology. Applied sciences suggest exploring this potential with business collaborations and consulting. The previously mentioned research responsibility in influenced self-concept and human behaviour should also be highlighted if the self is multidimensional for SMM with online identity or networked self as a user, customer and citizen combined. Medicine has a low representation, but this position could be changed with digitalised health services.

Conclusions

In terms of academic contribution, first, this study is different from other systematic reviews of self-concept, focusing only on the latest trends in relation to social media machines. Thus, this paper has explored the key research topics and highlighted the outlook for the future, and has outlined the potential of non-academic joint research ventures for trustworthy social media technology. An unexpected result was that medicine, nursing and health care had a low representation in the results. However, their technologies are widespread, and medicine still has a key role in self-concept research in general. The emerging role of consumption and brand research is more obvious if the self is available and measurable in SMM. Further emerging trends are expected in both areas, especially if conversational or fake media are more connected to academic research.

In line with this, a great potential is for proactive research to explore the effectiveness and negative side effects of SMM. However, the reflective research approach also has the option to avoid the negative effects of digital technology and interpret the changes. Hence, this study went beyond the previous literature by supporting the researchers and practitioners to explore the responsibility of this domain.

Last but not least, this study implies the question: Are we, academic researchers, ready for social media machines with technology-driven selves? From the perspective of this systematic literature review, we are not yet ready. However, the first signs of future trends already indicate an expected change point. In the meantime, researchers have a clear responsibility to seek to understand how self-concept is influenced by social media machines.

Research limits

The present research has basic limitations, as conference papers were excluded from the studied databases. The latter would be relevant, as their results appear faster compared to the publication process of journal articles. However, this study focused only on high-ranked academic outputs with carefully considered and well-designed articles for a more focused and less diverse output. Despite these limits, this study presents the latest research topic trends in the SMM-driven self-concept with future opportunities for research and development.

Funds

Sincere gratitude to 1. the Fulbright Research Grant for the opportunity to work on this project, grant number: 1221101; 2. the Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences, grant number: BO/00045/19/9; 3. New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, Bolyai+, grant number: UNKP-22-5-NKE-87; and 4. the TKP2020-NKA-10 project financed under the 2020-4.1.1-TKP2020 Thematic Excellence Programme from the National Research, Development and Innovation Fund of Hungary.

References

Alhajj, R. (2018). Advanced Technology and Social Media Influence on Research, Industry and Community. In Amine, A., Mouhoub, M., Ait Mohamed, O. and Djebbar, B. (eds.): *Computational Intelligence and Its Applications*. CIIA 2018. IFIP Advances in Information and Communication Technology, vol. 522. Cham: Springer. 1–9. Online: https://doi.org/10.1007/978-3-319-89743-1_1

Anshar, M. and Williams, M.-A. (2016). Evolving Synthetic Pain into an Adaptive Self-Awareness Framework for Robots. *Biologically Inspired Cognitive Architectures,* 16, 8–18. Online: https://doi.org/10.1016/j.bica.2015.10.002

Arun, R., Suresh, V., Madhavan, C. V. and Murthy, M. N. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. In Zaki, M. J., Yu, J. X., Ravindran, B. and Pudi, V. (eds.): *Advances in Knowledge Discovery and Data Mining*. Berlin–Heidelberg: Springer. 391–402. Online: https://doi.org/10.1007/978-3-642-13657-3_43

Baldwin, J. M. (1897). The Psychology of Social Organization. *Psychological Review,* 4(5), 482–515. Online: https://doi.org/10.1037/h0072331

Barrett, P. M., Steinhubl, S. R., Muse, E. D. and Topol, E. J. (2017). Digitising the Mind. *The Lancet,* 389(10082), 1877. Online: https://doi.org/10.1016/S0140-6736(17)31218-7

Bentley, R. A., O'Brien, M. J. and Brock, W. A. (2014). More on Maps, Terrains, and Behaviors. *Behavioral and Brain Sciences,* 37(1), 105–119. Online: https://doi.org/10.1017/S0140525X1300277X

Bird, S., Klein, E. and Loper, E. (2009). *Natural Language Processing with Python. Analyzing Text with the Natural Language Toolkit*. O'Reilly Media.

Buss, A. (2001). *Psychological Dimensions of the Self*. Thousand Oaks: Sage Publications.

Cao, J., Xia, T., Li, J., Zhang, Y. and Tang, S. (2009). A Density-Based Method for Adaptive LDA Model Selection. *Neurocomputing,* 72(7–9), 1775–1781. Online: https://doi.org/10.1016/j.neucom.2008.06.011

Cao, Y. and Wang, L. (2017). Automatic Selection of t-SNE Perplexity. *JMLR: Workshop and Conference Proceedings,* 1, 1–7.

Catherine, H. Y., Parsons, J., Mamdani, M., Lebovic, G., Shah, B. R., Bhattacharyya, O., Laupacis, A. and Straus, S. E. (2012). Designing and Evaluating a Web-Based Self-Management Site for Patients with Type 2 Diabetes – Systematic Website Development and Study Protocol. *BMC Medical Informatics and Decision Making,* 12(1), 1–10. Online: https://doi.org/10.1186/1472-6947-12-57

Chen, A. T. (2012). Exploring Online Support Spaces: Using Cluster Analysis to Examine Breast Cancer, Diabetes and Fibromyalgia Support Groups. *Patient Education and Counseling,* 87(2), 250–257. Online: https://doi.org/10.1016/j.pec.2011.08.017

Chen, P., Zhang, N. L., Liu, T., Poon, L. K., Chen, Z. and Khawar, F. (2017). Latent Tree Models for Hierarchical Topic Detection. *Artificial Intelligence,* 250, 105–124. Online: https://doi.org/10.1016/j.artint.2017.06.004

Chen, Y. (2019). The Sustainable Development of Social Media Contents: An Analysis of Concrete and Abstract Information on Cultural and Creative Institutions with "Artist" and "Ordinary People" Positioning. *Sustainability,* 11(15). Online: https://doi.org/10.3390/su11154131

Choi, G. Y. and Behm-Morawitz, E. (2018). Teach Me about Yourself(ie): Exploring Selfie-Takers' Technology Usage and Digital Literacy Skills. *Psychology of Popular Media Culture,* 7(3), 345–360. Online: https://doi.org/10.1037/ppm0000130

Claffey, E. and Brady, M. (2017). Examining Consumers' Motivations to Engage in Firm-Hosted Virtual Communities. *Psychology & Marketing,* 34(4), 356–375. Online: https://doi.org/10.1002/mar.20994

Colliander, J. (2019). “This Is Fake News”: Investigating the Role of Conformity to Other Users’ Views When Commenting on and Spreading Disinformation in Social Media. *Computers in Human Behavior,* 97, 202–215. Online: https://doi.org/10.1016/j.chb.2019.03.032

Druga, S., Williams, R., Breazeal, C. and Resnick, M. (2017). “Hey Google Is It OK If I Eat You?” Initial Explorations in Child-Agent Interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children.* 595–600. Online: https://doi.org/10.1145/3078072.3084330

Dwivedi, Y. K., Shareef, M. A., Simintiras, A. C., Lal, B. and Weerakkody, V. (2016). A Generalised Adoption Model for Services: A Cross-Country Comparison of Mobile Health (M-Health). *Government Information Quarterly,* 33(1), 174–187. Online: https://doi.org/10.1016/j.giq.2015.06.003

Hilliard, M. E., Sparling, K. M., Hitchcock, J., Oser, T. K. and Hood, K. K. (2015). The Emerging Diabetes Online Community. *Current Diabetes Reviews,* 11(4), 261–272. Online: https://doi.org/10.2174/1573399811666150421123448

Editorial (2016). Post-publication Criticism Is Crucial, but Should Be Constructive: In an Era of Online Discussion, Debate Must Remain Nuanced and Courteous. *Nature,* 540, 7–8. Online: https://doi.org/10.1038/540007b

Feher, K. (2019). Digital Identity and the Online Self: Footprint Strategies – An Exploratory and Comparative Research Study. *Journal of Information Science,* 47(2), 192–205. Online: https://doi.org/10.1177/0165551519879702

Feinerer, I., Hornik, K. and Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software,* 25(5), 1–54. Online: https://doi.org/10.18637/jss.v025.i05

Ferrara, E., Varol, O., Davis, C., Menczer, F. and Flammini, A. (2016). The Rise of Social Bots. *Communications of the ACM,* 59(7), 96–104. Online: https://doi.org/10.1145/2818717

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U. and Nagengast, B. (2015). More Value Through Greater Differentiation: Gender Differences in Value Beliefs about Math. *Journal of Educational Psychology,* 107(3), 663–677. Online: https://doi.org/10.1037/edu0000003

Georgakopoulou, A., Iversen, S. and Stage, C. (2020). Analysing Quantified Stories on Social Media. In *Quantified Storytelling. A Narrative Analysis of Metrics on Social Media.* Cham: Palgrave Macmillan. 1–30. Online: https://doi.org/10.1007/978-3-030-48074-5_1

Gesser-Edelsburg, A., Diamant, A., Hijazi, R. and Mesch, G. S. (2018). Correcting Misinformation by Health Organizations during Measles Outbreaks: A Controlled Experiment. *PloS One,* 13(12). Online: https://doi.org/10.1371/journal.pone.0209505

Goldston, R. B., Gara, M. A. and Woolfolk, R. L. (1992). Emotion Differentiation: A Correlate of Symptom Severity in Major Depression. *Journal of Nervous and Mental Disease,* 180(11), 712–718. Online: https://doi.org/10.1097/00005053-199211000-00005

Govender, P. and Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Pollution Research,* 11(1), 40–56. Online: https://doi.org/10.1016/j.apr.2019.09.009

Griffiths, T. L. and Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences,* 101(Suppl 1), 5228–5235. Online: https://doi.org/10.1073/pnas.0307752101

Guntuku, S. C., Schneider, R., Pelullo, A., Young, J., Wong, V., Ungar, L., Polsky, D., Volpp, K. G. and Merchant, R. (2019). Studying Expressions of Loneliness in Individuals Using Twitter: An Observational Study. *BMJ Open,* 9(11). Online: https://doi.org/10.1136/bmjopen-2019-030355

Hackler, J. C. (1966). Boys, Blisters, and Behavior: The Impact of a Work Program in an Urban Central Area. *Journal of Research in Crime and Delinquency,* 3(2), 155–164. Online: https://doi.org/10.1177/002242786600300209

Hollenbaugh, E. E. (2021). Self-Presentation in Social Media: Review and Research Opportunities. *Review of Communication Research,* 9, 80–98.

Honnibal, M. and Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

Hootsuite (2021). *Global Digital Overviews.* Online: https://datareportal.com/reports

Hou-Liu, J. (2018). *Benchmarking and Improving Recovery of Number of Topics in Latent Dirichlet Allocation Models.*

Houser, K. A. (2019). Can AI Solve the Diversity Problem in the Tech Industry: Mitigating Noise and Bias in Employment Decision-Making. *Stanford Technology Law Review,* 22, 290–354.

Jabłońska, M. R. and Zajdel, R. (2020). Artificial Neural Networks for Predicting Social Comparison Effects among Female Instagram Users. *PloS One,* 15(2). Online: https://doi.org/10.1371/journal.pone.0229354

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y. and Zhao, L. (2019). Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications,* 78(11), 15169–15211.

Jeong, B., Yoon, J. and Lee, J.-M. (2019). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management,* 48, 280–290.

Johnston, M. P. (2014). Secondary data analysis: A method of which the time has come. *Qualitative and Quantitative Methods in Libraries,* 3(3), 619–626.

Kruiger, J. F., Rauber, P. E., Martins, R. M., Kerren, A., Kobourov, S. and Telea, A. C. (2017). Graph layouts by t-SNE. *Computer Graphics Forum,* 36, 283–294.

Kušen, E. and Strembeck, M. (2019). Something draws near, I can feel it: An analysis of human and bot emotion-exchange motifs on Twitter. *Online Social Networks and Media,* 10, 1–17.

Leaff, L. A. (1971). Affect Versus Feeling: On the Concept of a Structuralized Ego Response. *Journal of the American Psychoanalytic Association,* 19(4), 780–786. Online: https://doi.org/10.1177/000306517101900410

Leist, A. K. (2013). Social media use of older adults: a mini-review. *Gerontology,* 59(4), 378–384.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T. et al. (2018). Applying lda topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures,* 12(2–3), 93–118.

Mazza, M., Cola, G. and Tesconi, M. (2022). Ready-to-(ab) use: From fake account trafficking to coordinated inauthentic behavior on twitter. *Online Social Networks and Media,* 31.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Group, P. et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *PLoS Med,* 6(7).

Muñoz-Laboy, M., Weinstein, H. and Parker, R. (2007). The Hip-Hop Club Scene: Gender, Grinding and Sex. *Culture, Health & Sexuality,* 9(6), 615– 628. Online: https://doi.org/10.1080/13691050701528590

Musgrove, F. (1967). Social Class and Levels of Aspiration in a Technological University. *The Sociological Review,* 15(3), 311–322. Online: https://doi.org/10.1111/j.1467-954X.1967.tb01284.x

Nathanson, E. (2017). Native voice, self-concept and the moral case for personalized voice technology. *Disability and Rehabilitation,* 39(1), 73–81.

Neff, G. and Nafus, D. (2016). *Self-Tracking.* MIT Press. Newland, J. (2016). The Social Media Machine. *Nurse Practitioner,* 41(10), 11–11. Online: https://doi.org/10.1097/01.NPR.0000499555.41195.c1

Oyserman, D. and Markus, H. R. (1998). Self as Social Representation. In Flick, U. (ed.): *The Psychology of the Social.* Cambridge: Cambridge University Press. 107–125.

Papadimitriou, C. H., Tamaki, H., Raghavan, P. and Vempala, S. (1998). Latent semantic indexing: A probabilistic analysis. In Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, PODS '98, 159–168, New York, NY, USA. Association for Computing Machinery.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K. and Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development,* 88(5), 1653–1670.

Preston, P., Mansell, R., Murdock, G., Siapera, E. and Na, Y. (2021). Special Section. *The Social Dilemma,* by Jeff Orlowski, Netflix, 2020. *The Political Economy of Communication,* 8(2), 75–103.

Pujades, S., Mohler, B., Thaler, A., Tesch, J., Mahmood, N., Hesse, N., Bülthoff, H. H., and Black, M. J. (2019). The virtual caliper: rapid creation of metrically accurate avatars from 3D measurements. *IEEE Transactions on Visualization and Computer Graphics,* 25(5), 1887–1897.

Puntoni, S., Reczek, R. W., Giesler, M. and Botti, S. (2021). Consumers and artificial intelligence: an experiential perspective. *Journal of Marketing,* 85(1), 131–151.

Rassameeroj, I. and Wu, S. F. (2019). Reverse Engineering of Content Delivery Algorithms for Social Media Systems. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS),* IEEE. 196–203. Online: https://doi.org/10.1109/SNAMS.2019.8931859

Rheault, L. and Musulan, A. (2021). Efficient Detection of Online Communities and Social Bot Activity during Electoral Campaigns. *Journal of Information Technology & Politics,* 18(3), 324–337. Online: https://doi.org/10.1080/19331681.2021.1879705

Ross, L. E. and Mughni, W. N. (1995). Effect of alcohol on the threshold for detecting angular acceleration. *Aviation, Space, and Environmental Medicine,* 66(7), 635–640.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., and Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence,* 53, 118–137.

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society,* 6(1).

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F. and Diamond, A. (2015). Enhancing cognitive and social-emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental Psychology,* 51(1), 52–66. Online: https://doi.org/10.1037/a0038454

Siegler, I., Israël, I. and Berthoz, A. (1998). Shift of the Beating Field of Vestibular Nystagmus: An Orientation Strategy? *Neuroscience Letters,* 254(2), 93–96. Online: https://doi.org/10.1016/S0304-3940(98)00671-5

Silge, J. and Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *JOSS,* 1(3).

Sinclair, M. (2001). Information technology skills of midwives in Northern Ireland. A survey report. *The Practising Midwife,* 4(1), 23–28.

Skjuve, M., Følstad, A., Fostervold, K. I. and Brandtzaeg, P. B. (2021). My chatbot companion – A study of human-chatbot relationships. *International Journal of Human-Computer Studies,* 149.

Stahlschmidt, S. and Stephen, D. (2020). Comparison of Web of Science, Scopus and Dimensions databases. *KB Forschungspoolprojekt,* 1–37.

Swan, M. (2012). Health 2050: The realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. *Journal of Personalized Medicine,* 2(3), 93–118.

Thackray, R. I., Jones, K. N. and Touchstone, R. M. (1973). Self-estimates of Distractibility as Related to Performance Decrement on a Task Requiring Sustained Attention. *Ergonomics,* 16(2), 141–152.

Thomas, J., Al-Shehhi, A., Al-Ameri, M. and Grey, I. (2019). We tweet Arabic; I tweet English: self-concept, language and social media. *Heliyon,* 5(7).

Thomas, V., Balzer Carr, B., Azmitia, M. and Whittaker, S. (2021). Alone and online: Understanding the relationships between social media, solitude, and psychological adjustment. *Psychology of Popular Media,* 10(2), 201–211. Online: https://doi.org/10.1037/ppm0000287

Tinati, R., Halford, S., Carr, L. and Pope, C. (2014). Big data: methodological challenges and approaches for sociological analysis. *Sociology,* 48(4), 663–681.

Tseng, T. H. and Hsieh, S. H. (2019). Determinants of emoticon usage in mobile instant messaging: a construal level theory perspective. *Behaviour & Information Technology,* 38(3), 289–301.

Tsesis, A. (2018). Marketplace of Ideas, Privacy, and the Digital Audience. *Notre Dame Law Review,* 94(4), 1585–1629.

Varshney, D. and Vishwakarma, D. K. (2022). A unified approach of detecting misleading images via tracing its instances on web and analyzing its past context for the verification of multimedia content. *International Journal of Multimedia Information Retrieval,* 11(3), 445–459.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems,* 28(2), 118– 144.

Webster, C. (1995). Computer modeling of adaptive depression. *Behavioral Science,* 40(4), 314–330.

Wickham, H., François, R., Henry, L., and Müller, K. (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7.

Yang, C.-c. and Brown, B. B. (2016). Online Self-Presentation on Facebook and Self Development During the College Transition. *Journal of Youth and Adolescence,* 45(2), 402–416. Online: https://doi.org/10.1007/s10964-015-0385-y

Yenney, E. H. (1973). *Correlation between Two Standardized Tests of Depression.* Psychological reports.

Zahedi, Z. and Haustein, S. (2018). On the Relationships between Bibliographic Characteristics of Scientific Documents and Citation and Mendeley Readership Counts: A Large-Scale Analysis of Web of Science Publications. *Journal of Informetrics,* 12(1), 191–202. Online: https://doi.org/10.1016/j.joi.2017.12.005

Zarouali, B., Dobber, T., De Pauw, G., and de Vreese, C. (2020). Using a personality-profiling algorithm to investigate political microtargeting: Assessing the persuasion effects of personality-tailored ads on social media. *Communication Research,* 49(8), 1066–1091. Online: https://doi.org/10.1177/0093650220961965