

# Discourses surrounding sustainability and digitalization in Europe on Twitter over time

*Digitalization and sustainability transformations are contested change processes, accompanied by wide public discourse. But what concerns the public? Our analysis of the social media discourse on Twitter in the last decade reveals key discursive hubs such as smart cities and climate change, as well as blind spots such as sufficiency strategies. It also points to differences between societal and academic discourse, and where increased engagement of researchers and sustainability professionals would be needed to move forward.*

Mario Angst , Nadine Strauß 

## Discourses surrounding sustainability and digitalization in Europe on Twitter over time

GAIA 32/S1 (2023): 10–20

### Abstract

This study analyzes the discourses surrounding the interrelation between digitalization and sustainability in Europe on Twitter between 2010 and 2021. We identify 34,802 tweets related to the interrelation between digitalization and sustainability among 634,017 tweets discussing sustainability issues with explicit mentions of Europe. Based on a qualitative analysis of tweets, we identify the main domains discussed (and not discussed). We then sketch the development of the identified domains, as well as their relationship to each other over time, based on a quantitative analysis of their (co-)occurrences. We find that smart city and mobility were two of the most dominant and interrelated domains, particularly in the middle of the decade. In parallel, the domain of climate change has gained ever more attention since 2017 and has emerged as a discursive hub. We further develop hypotheses for how external factors and events (especially EU-level programs) likely led to increases in attention to some domains. Finally, we find that the Twitter discourse across domains mirrors common blind spots regarding sustainable digitalization discourses in its uncritical stance toward economic growth and its overreliance on efficiency in comparison to sufficiency concerns.

### Keywords

digitalization, discourse, social media, sustainability

Concerning their discursive power, sustainability and digitalization have become two of the most dominant “mega-trends” of our time (Lichtenthaler 2021, p. 64). Particularly since the COVID-19 crisis, public attention to the interrelation between sustainability transformations and digitalization has further intensified, following governments (e.g., *European Green Deal*), intergovernmental associations (e.g., *United Nation Sustainable Development Goals*), and companies promoting sustainability and digitalization as joint strategic goals (cf. Del Río Castro et al. 2021, Lichtenthaler 2021).

However, research in the past has focused primarily on the theoretical and conceptual relationship between digitalization and sustainability transformations (e.g., Lichtenthaler 2021, Seele and Lock 2017). Yet, both refer to socio-technical processes that are driven and accompanied by societal discourses. This is especially relevant to sustainability transformations, which are inherently negotiated in reference to the goal of sustainability as a normative concept. Beyond the normative and empirical discussion of how the potential for sustainability transformations co-evolves with digitalization, it becomes equally relevant to study the interrelation between the discourses of sustainability and digitalization in the public domain over time (e.g., Andersen et al. 2021, Galaz et al. 2021). Analyzing and understanding the public discourses concerning these two processes are of crucial importance to determine the feasibility and direction of transformations and to identify future pathways, main drivers, and dominating or (under)represented discourses.

Societal discourses take place in different forms and in different fora and environments. We take one of various possible approaches and analyze the social media discourse surrounding digitalization and sustainability on the platform Twitter, a widely researched social medium that has been identified to reflect public discourse on a variety of topics (e.g., politics: Ott 2017). In the context of this special issue, we also conducted our analysis with a specific focus on the European discourse: how has the social media discourse on the interplay between sustainability and digitalization in the European context been structured on Twitter between 2010 and 2021, and what were likely causes of changes to the structure of the discourse during this time?

Dr. Mario Angst | University of Zurich | Digital Society Initiative | Zurich | CH | [mario.angst@uzh.ch](mailto:mario.angst@uzh.ch)

Prof. Dr. Nadine Strauß | University of Zurich | Department of Communication and Media Research | Zurich | CH | [n.strauss@ikmz.uzh.ch](mailto:n.strauss@ikmz.uzh.ch)

© 2023 by the authors; licensee oekom. This Open Access article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).  
<https://doi.org/10.14512/gaia.32.S1.4>

Received May 2, 2022; revised version accepted January 17, 2023 (double-blind peer review).

## Conceptual framework

### Sustainability transformations and digitalization:

#### Two interrelated processes

We view sustainability as a normative concept of intra- and intergenerational justice and human flourishing within planetary boundaries (Steffen et al. 2015, Raworth 2017). As such, sustainability transformations are socio-technical change processes oriented toward sustainability (Schneidewind 2018), and they occur next to and interact with other socio-technical change processes, one of which is digitalization. Following Bockshecker et al. (2018, p. 8), we define digitalization as “the state of an organization or a society referring to its current digital development and usage of ICT innovations. Digitalization takes into account social as well as technical elements”.

Information and communication technologies (ICT) has become a constant in the new “digital age” (Schmidt and Cohen 2013). However, while digitalization offers new opportunities to save resources, engage citizens, and limit carbon emissions, scholars have also argued that the way new technologies are employed can be counter-productive to achieving sustainability goals. Some even advocate for a general “digital reset” to re-calibrate the relationship between digitalization and sustainability, arguing that fundamental developments in how digitalization plays out in societies in its orientation toward economic growth, its single-minded focus on efficiency, and its lack of participation must be reset to achieve societal transformation toward sustainability (D4S 2022). Regarding resource use, for example, digital infrastructures and technological devices (e.g., smartphones, batteries) have been shown to require large amounts of energy and natural resources, and their resource efficiency potentials are often dwarfed by rebound effects (Hilty and Aebischer 2015). Digitalization is also reshaping resource and power distributions within societies, with ambiguous results in terms of inter- and intragenerational justice, which is a core tenet of sustainability as a normative goal. Despite a plethora of research concerning sustainability transformations and digitalization separately, scholars have also pointed out that research investigating both societal change processes jointly is still limited (Del Río Castro et al. 2021, Lichtenthaler 2021) and often lacks interdisciplinary perspectives and approaches.

#### Sustainability and digitalization discourses

Sustainability transformations and digitalization are not only two interrelated socio-technical change processes, but they are also accompanied and driven by societal discourses. This is especially relevant to sustainability transformations, which are broadly discussed in reference to sustainability as a normative concept, but also interpreted, contested, and translated constantly by a variety of actors. This is also true for digitalization—which is not an entirely self-referential technical process—without reference to the normative stances of the actors involved in shaping it. Indeed, the course that digitalization follows in society is largely dependent on how it is framed and perceived in public discourses,

which is additionally affected by institutions and power structures (Marenco and Seidl 2021).

However, few empirical studies have investigated the interplay between digitalization and sustainability in the public discourse, and this is especially true for mediated societal discourses. Following Brenner and Hartl (2021, p. 4), the news media are central to “creating and reproducing discourses about how digitalization and sustainability interact”. Based on the agenda-setting theory (McCombs and Shaw 1972), it is argued that the news media, including social media, have the power to influence what concerns the public. In effect, much research in the field of communication studies has found that the issues being discussed prominently in the news media transfer to the minds of the public, thus ranking highly on the public agenda (e.g., Kioussis and McCombs 2004). In this study, we equally assert (but do not aim to prove empirically) that in combination with the relative attention given to various aspects of the discourse on digitalization and sustainability, their presentation on media platforms (e.g., Twitter) ultimately contributes to the conditions under which processes play out in societies, particularly in terms of regulation or policymaking (cf. Soroka 2002).

While scholars have extensively studied the coverage of climate change (Hase et al. 2021) and, to a smaller degree, the representation of new, digitalization-adjacent technologies (e.g., nanotechnology: Metag and Marcinkowski 2013, digitalization in agriculture: Mohr and Höhler 2021) in the news media, the interrelation between the two discourses of digitalization and sustainability has received little scholarly attention thus far. Lenz (2021) offers the first qualitative account of three common narratives being used in the public discourse to describe the connection between sustainability and digitalization. On a general level, they distinguish digital technologies as problem solvers, digital tools as opportunities for participation and inclusion, and technological innovations as solutions to ecological disasters. In a similar vein, Brenner and Hartl (2021) qualitatively analyzed news media coverage of sustainability and digitalization in Austrian news from 1990 until 2019 and identified four frames, presenting the relationship between the two processes as a stand-alone challenge (frame 1), a result of the impact of digitalization on sustainability (frame 2), not leading to a sustainable solution (frame 3), and asserting digitalization as a positive catalyst for sustainability (frame 4).

These first analyses offer a qualitative overview of the narratives and frames being used in the discourse surrounding digitalization and sustainability, mainly based on print news. However, a more encompassing assay of how the discourse has evolved on a broader scale, on more timely platforms for public discourses (e.g., social media), and over time is currently lacking. In fact, previous research has suggested expanding the analysis of the discourse on digitalization and sustainability from the news media to social media (Brenner and Hartl 2021). Thus, we contribute to filling this gap by analyzing social media discourse, specifically on the microblogging service Twitter, with a specific focus on Europe.



## Methods

We empirically analyze the occurrence, relative weight, and, to some extent, framing and likely drivers of different aspects of the English-language discourse on the interplay between sustainability and digitalization in Europe over time on the social media platform Twitter between 2010 and 2021. To answer our research questions concerning the structure of the Twitter discourse over time and the external factors likely to explain them, we focus on three key aspects. First, we inductively identify distinct discourse domains. Second, we then analyze the occurrence and co-occurrence of these discourse domains over time and explore likely drivers of trends and patterns. Finally, we investigate the presence or absence of two key transversal discourse topics, which are core components of critical academic and policy discussions surrounding digitalization and sustainability (D4S 2022): framing of sustainability and digitalization regarding economic growth and the reliance on a narrative of resource use efficiency gains regarding digitalization, as well as considerations of sufficiency strategies.<sup>1</sup>

At the time of our analysis (spring 2022), Twitter was a privately owned social media platform, which at its core allowed users to formulate short statements restricted to 280 characters (so-called tweets, a form of microblogging), optionally accompanied by images, videos, or web links. Authors of tweets were able to share these with other platform users. In addition, users could interact in various ways with tweets, including resharing and liking them. While widely used, Twitter data are by no means a representative or unbiased source for analyzing societal discourse. There are clear limitations fundamentally linked to the fact that data are gathered on a platform not designed for research purposes and set up by a private company. In our case, an analysis of tweets is likely skewed toward aspects of the discourse specifically relevant to elite actors. This is because professional entities or representations (Sloan 2017) and higher socioeconomic classes (Yates and Lockley 2018) have been found to be likely over-represented among Twitter users, including academics, journalists, or politicians.

### Analysis pipeline

Our analysis proceeds in three steps. First, we obtain a dataset of tweets relating to the overall discourse on sustainability and digitalization between 2010<sup>2</sup> and 2021 in Europe by querying the Twitter application programming interface (API) with a set of keywords and then training a binary text classifier on the data obtained to arrive at a subset of relevant tweets. Second, we inductively identify a set of domains prevalent in the overall Twitter discourse and assign one or more discourse domains to tweets using pattern matches. Third, we proceed to classify tweets regarding the occurrence of transversal (cross-domain) discourse dimensions, which are (stances toward) economic growth and mentions of efficiency and sufficiency. To do so technically, we build on recent work using zero-shot learning approaches (Gambini et al. 2022).

### Step 1a: Querying the Twitter API

We query the Twitter Academic API (full archive search) with two queries to gather two datasets. First, we gather a starting set of English-language tweets using a relatively narrow search query based on terms used in Andersen (2021)—with a European focus added—and adapted to the requirements of the Twitter API endpoint (figure 1). This yields the dataset *sus\_digi\_eu*, comprising 15,592 unique tweets.

#### Query 1:

```
(digital OR digitalization OR digitalisation OR ict)
AND (sustainable OR sustainability OR SDG)
AND (europe OR european OR eu))
English language, no retweets
between 2010 and 2021
```

FIGURE 1: Twitter API query 1.

Second, we query the Twitter API with a second query (figure 2) designed to gather a much broader dataset of tweets relating to sustainability in Europe. This resulted in the dataset *sus\_eu*, comprising 634,017 unique tweets, which is a superset of *sus\_digi\_eu*.

#### Query 2:

```
(sustainable OR sustainability OR SDG)
AND (europe OR european OR eu)
English language, no retweets
between 2010 and 2021
```

FIGURE 2: Twitter API query 2.

### Step 1b: Classification of discourse-related tweets

Even a brief inspection of the *sus\_eu* dataset resulting from the keyword-based query to the Twitter API reveals it contains many tweets unrelated to the discourse on the interplay between digitalization and sustainability and, to a lesser degree, tweets that address issues outside the European context. Thus, we trained a binary text classifier using the natural language processing framework *spacy*, implemented in Python, to filter relevant tweets (Montani et al. 2022). The classifier makes use of the pretrained transformer model *distilroberta* (Liu et al. 2021).

To train the classifier, a team of three coders first manually annotated a gold standard evaluation set of 600 randomly sampled tweets (400 from *sus\_digi\_eu*, 200 from *sus\_eu*), based on an initial codebook. By comparing and resolving differences among coders, we then created a final codebook and a test set, which was held back from training. A further 586 randomly sampled tweets from *sus\_eu* were additionally added to this test set based

1 Computer code and access to tweet IDs to reproduce the analysis presented here is available at a public repository at <https://doi.org/10.5281/zenodo.7555375>.

The access to tweet IDs is in accordance with Twitter's developer policy regarding content redistribution at <https://developer.twitter.com/en/developer-terms/policy> (as of 26 April 2022).

2 2010 can be seen as the start of a phase for Twitter in which the microblogging platform added several crucial functions raising its popularity.

on the final codebook. Adding more samples from *sus\_eu* ensured a robust evaluation set that is more representative of the *sus\_eu* target set.

In the final codebook, we chose to be relatively inclusive by treating the co-occurrence of explicit sustainability and digitalization mentions as necessary and sufficient for accepting a tweet. At the same time, we specified two explicit exclusion criteria: first, the tweet could not use “sustainable” explicitly in the sense of “long-term” or “enduring.” Without an explicit qualification, we considered the mere use of “sustainable” to be sufficient, as we judged this to give the actors participating in the discourse the benefit of the doubt regarding their use of the sustainability terminology, preventing us from imposing our own ontological stances about sustainability concepts. Second, a tweet could not refer explicitly and solely to an issue outside the European context and without European involvement. If no geographical location was mentioned explicitly, we accepted the tweet.

Based on the final codebook, we coded a training set of 3,841 tweets from the *sus\_eu* dataset to train the classifier. Classes in the large *sus\_eu* dataset were highly imbalanced, which led us to emphasize precision over recall in training the classifier, because in the presence of a large imbalance in our dataset, low precision would quickly result in a significantly high number of false positives. Currently, the classifier achieves a precision of 0.85 on the evaluation set and a recall of 0.69, leading to an F1 score of 0.76. Applying the classifier to the *sus\_eu* dataset yielded a dataset of 34,802 tweets. Given our relatively low recall on the test set, the true total number of tweets relevant to the discourse is likely higher, while we can be relatively more confident – given our emphasis on precision – that the tweets we identify are truly relevant.

### Step 2: Inductive identification and rule-based classification of domains in tweets

To identify inductively the discourse domains after initial filtering, we follow an approach inspired by what Carlsen and Ralund (2022) call computer-assisted text analysis, a variant of computational grounded theory. Following the workflow terminology of Carlsen and Ralund (2022), in a qualitative *discovery* step, we utilized our immersion into the corpus gained by annotating the first binary classifier to derive inductively a set of domains and to assign non-overlapping, single-, and multi-word search terms to them. We then extended these search terms by incorporating suggestions of similar single- and multi-word terms based on *sense2vec* (Trask et al. 2015). A common practice on Twitter is the use of so-called hashtags, which often combine multiple words into one string (e.g., “smart city” becomes “#smartcity”). We accounted for this by combining all our multi-word searches into single words, in addition to multi-word patterns.

In a *grounding* step, we then applied the search terms and explored them in context, updating them along the way and refining the domain set. We iterated in this way over five main rounds, arriving at a final set of 29 discourse domains and associated search terms. In a classification step, we then assigned

domain labels to tweets using a rule-based model based on our search terms.

### Step 3: Zero-shot classification of transversal discourse dimensions economic growth and efficiency

We evaluated tweets related to the discourse concerning whether they referred to two broader transversal discourse dimensions (i.e., spanning domains). We focused on the following two dimensions: reference to economic growth, which we also further qualitatively assessed concerning stances taken in tweets (a supportive or critical stance towards economic growth) and references to (resource) efficiency and sufficiency considerations. To classify these transversal dimensions, we relied on a zero-shot classifier trained on the *MultiNLI* dataset (Williams et al. 2018), and tweets were preprocessed for the zero-shot classifier following Gambini et al. (2022). We evaluated the performance of each classification against test sets of tweets annotated by a team of four annotators<sup>3</sup>.

### Contextualization of domain presence

To move beyond the description of domain occurrence and co-occurrence over time, we were interested in identifying potential causes of the spikes in certain domains during the period of analysis. To do so, we followed a qualitative explorative approach in conducting desk research. First, we used a search engine, Google, with the keywords related to the key domains within the period in which they were mostly present and with the qualifier “Europe”. Relevant hits were scrutinized and screened as to whether they dealt with the respective domains and Europe in the given timeframe. Second, because actors related to the European Union (EU) were behind a predominant number of accounts in the tweets analyzed, we also searched the website of the EU with the respective keywords to identify events, programs, or other external factors related to the domains. Our findings are presented as hypothetical explanations of the presence of dominant domains in our dataset over time.

## Results

### Domains over time: Dominant domains and developments

Table 1 (p. 14) lists the 29 domains we inductively identified and around which the discourse on Twitter revolved, including the labels assigned to them, which were used in visualizations in this article. Figure 3 (p. 15) shows the relative presence of domains in the discourse between 2010 and 2021. Some domains are stable components of the discourse over the decade in our analysis. Smart city concepts (SMC) and mobility (MOB) are present in the top 5 in every year, and discussions of them in terms of their relative weight in the discourse peaked mid-decade. The joint

>

<sup>3</sup> Economic growth classifier: Matthew's correlation coefficient (MCC) 0.67, balanced accuracy (BA) 0.85. Economic growth support: MCC 0.54, BA 0.79. Efficiency: MCC 0.58, BA 0.79.



**TABLE 1:** Domains identified in the English-language discourse on sustainability and digitalization in Europe on the social media platform Twitter between 2010 and 2021.

DOMAIN	LABEL
agriculture	AGR
air travel	AIR
biodiversity	BIO
blockchain technology	BLO
construction	BUI
circular economy	CIR
climate change	CLIM
recovery/COVID	COV
cultural heritage	CUL
data center	DAC
energy transition	ENE
EU programs	EUP
fashion	FAS
finance/economy/investments	FIN
fishing/ocean	FIS
forestry	FOR
health	HEA
smart home	HOM
green information and communication technologies	ICT
manufacturing	MAN
transport/mobility	MOB
future of work	OFF
pollution/waste	POL
raw material use	RES
smart city	SMC
small and medium enterprises	SME
smart village	SMV
tourism	TOU
water supply	WAT

presence of mobility and smart city concepts is likely because discussions of how to make cities “smart” often focus on sustainable forms of mobility. As such, the mobility domain intersects with the smart city concept domain to an extent. The predominant presence of smart city concepts and mobility between 2010 and 2021 could be explained by the European *Smart Cities Marketplace* program, the first EU publication, dating to 2010. Data centers (DAC) and sustainable ICT domains were much more prominent at the beginning of the decade, when the usage of cloud data storage and computing became widespread. Only in 2021 did data centers again gain prominence in the discourse, which might have been guided by the European data strategy, initiated in 2020.

In fact, energy (ENE) is another constant domain related to the discourse that remained in the top 5 until 2020. Akin to the other domains, energy has been a prominent theme discussed in Europe, spearheaded by the introduction of the *Energy Union Strategy* in 2015 (EC 2015), a key priority of the Juncker Commission (2014 to 2019). The most rapid increase in prominence in the latter half of the decade relates to the discussion of climate change (CLIM) since 2017/2018 and finance (FIN) since 2020. The high prevalence of climate change since 2015 is likely related to the accomplishment of the *Paris Agreement* in 2015, the more

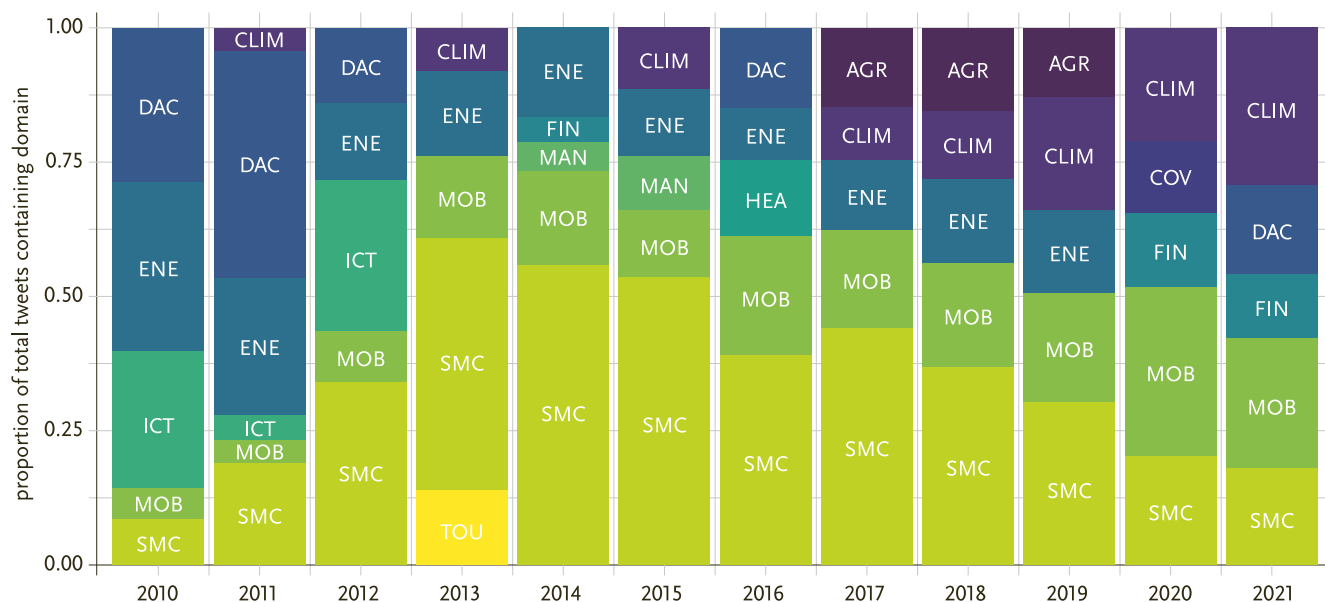
frequent releases of UN Intergovernmental Panel on Climate Change (IPCC) reports since 2018, and the more general increase in awareness of climate change in the wake of the *Fridays for Future* protests in Europe since 2018. The fact that the finance domain has been one of the main domains since 2019/2020 could be related to the EU-wide discussions of the EU taxonomy, a framework for sustainable investments in Europe and parallel developments in fintech (financial technology).

We can also analyze the stability and developments of domains by comparing domain occurrences in each year in relation to the previous year. Figure 4 (p. 16) shows such trend lines for every domain, emphasizing dominant domains in the yearly discourse using transparency and line width, as large swings in presence are much more likely among domains with few mentions. Further, this normalization regarding the relative weight of a domain in the overall discourse each year adjusts for the overall increase in tweets in general in our dataset over the analyzed period. Some striking results include the introduction of blockchain technologies (BLO) in 2018, which can partly be explained by the explosion of Bitcoin, Ethereum, and Litecoin transactions in the same year. Further, the large upswing in the health domain (HEA) in 2016 could be related to the publication of the seminal biannual report *Health at a Glance* by the Organisation for Economic Co-operation and Development (OECD) and the EU the same year, wherein the concept of sustainable health systems in Europe was mentioned for the first time.

Similarly, tourism (TOU) experienced an uptick in 2013 and again in 2018, where the latest increase could be explained by the *European Capitals of Smart Tourism* award, which was first handed out in 2018 and which was prominently tweeted about in our dataset. In 2017, we can identify the introduction of the agriculture domain (AGR), which is likely related to the strategic approach to EU agricultural research and innovation presented in Brussels in 2016, highlighting the potential of technology for sustainability in the farming sector and rural areas. In comparison, we observed a large increase in the presence of the EU-level program (EUP) domain in 2019 and 2020, which might be a result of the announcement and implementation of the EU digital agenda, the *European Green Deal*, and the Covid recovery plan, all of which emphasized exploiting the interplay between digitalization and sustainability.

### Transversal discourse dimensions

Results concerning the occurrence of our transversal discourse dimensions differ substantially between dimensions. Of all tweets labeled as dealing with economic growth (ca. 3%), 88% were labeled as taking a positive or supportive stance, while the remainders were inconclusive. There are so few critical stances toward economic growth that a meaningful analysis of its presence in the discourse was not possible (even though the few labeled examples generally illustrate that the classifier could identify them, albeit with low precision). For efficiency considerations, we find 4.6% of tweets classified as containing efficiency content, which is a sizable proportion of the discourse, especially



**FIGURE 3:** Top five domains per year (relative presence) in English-language tweets relating to the discourse on sustainability and digitalization in Europe between 2010 and 2021. In 2011, domains MOB and CLIM share place 5, because they had exactly the same number of tweets. See table 1 for domain label reference.

given that our classifier has relatively low recall on our test set and, as such, likely underestimates the true number of efficiency-related tweets. We also find about 0.05 % of tweets classified as sufficiency related, which is, however, primarily due to the presence of self-sufficiency considerations in the smart rural topic. Thus, as with critical stances toward economic growth, sufficiency in a broader sense is likely mostly absent from the broader Twitter discourse.<sup>4</sup>

#### Interrelation among domains: The discursive landscape

Some discourse domains are discussed in relation to each other more often than others. As such, the overall discourse on sustainability and digitalization in Europe can be seen as a network of interrelated domains, some of which cluster together to form sub-discourses that go beyond single domains. Analyzing domain interrelations in this way provides a more high-level overview of the evolution of the discursive landscape or discourse topology. We analyze domain interrelation by analyzing how often domains co-occur in the same tweets. Figure 5 (p. 17) illustrates such co-occurrences in the empirical example of the second- and fourth-most liked tweets in our datasets.

Figure 6 (pp. 18f.) shows a network visualization of domain co-occurrences over three phases. We assigned a phase each to the years between 2011 and 2014, between 2015 and 2019 (pre-pandemic), and between 2020 and 2021 (pandemic). We normalize counts of co-occurrences for every domain in the symmetric co-occurrence matrix, which accounts for the imbalance in domain occurrences and treats variations in every domain independently of its absolute occurrence. Figure 6 displays up to the ten most frequent co-occurrences (top ten co-occurrences) a do-

main has with other domains for every domain. This co-occurrence graph is an indication of a higher-level structure within the overall discourse, and we clustered domains within the graph using modularity<sup>5</sup> maximization (Brandes et al. 2008).

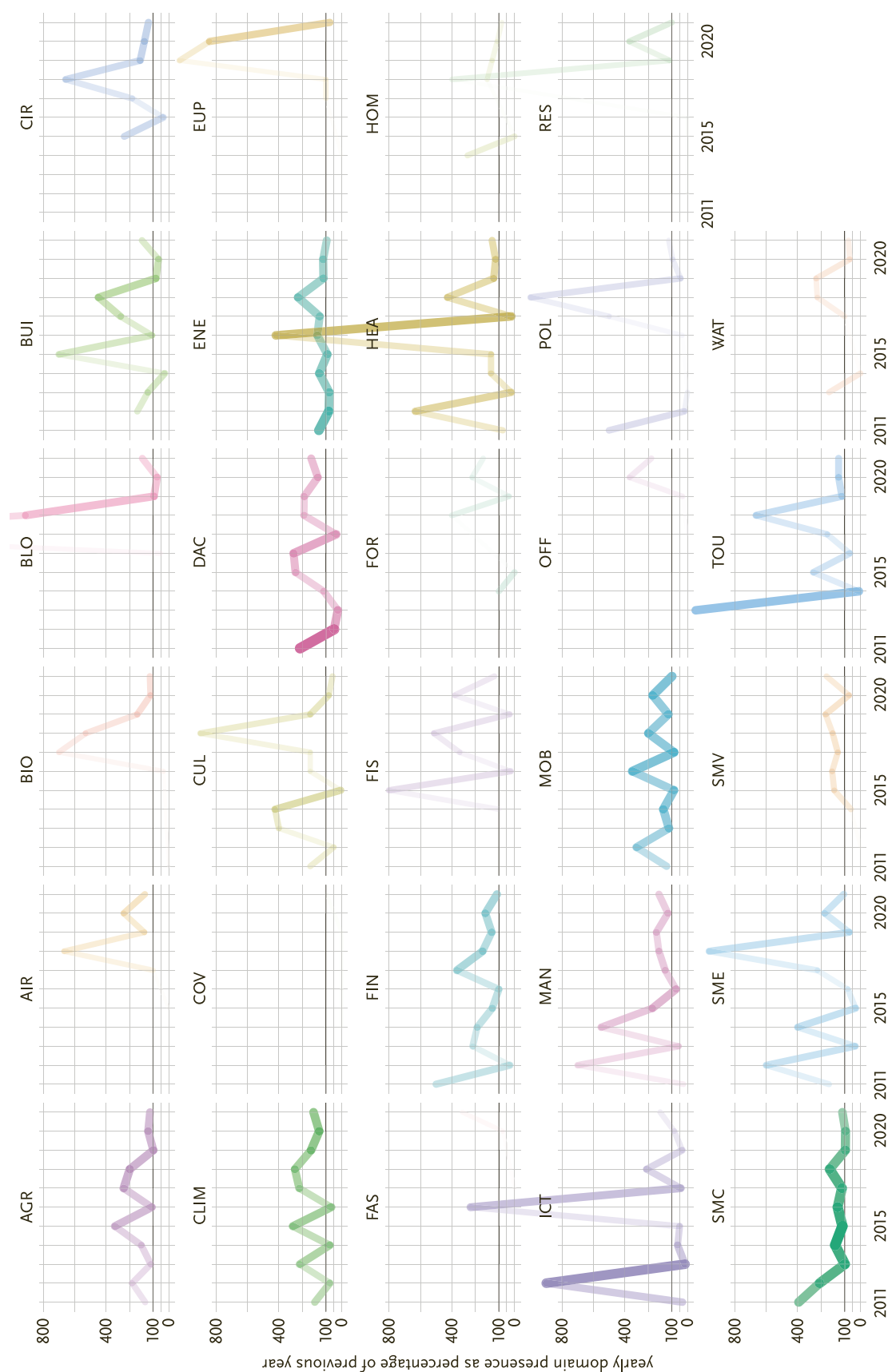
## Discussion

In this study, we analyzed the public discourse on the interrelationship between digitalization and sustainability on Twitter between 2010 and 2021. Our goal was to identify the main domains discussed in the discourse, to sketch the development and drivers of these domains and their interrelationship over time, and to investigate the discourse for the presence of key transversal elements highlighted in the fields of academia and policy. Some domains were specifically prominent in certain years but were not consistently discussed during the period under study. A likely explanation for such punctuations in attention were external factors, such as political issues or events, programs, or social movements (cf. Downs 1972) as highlighted in the contextualization of our results. Confirming previous research (Marenco and Seidl 2021), the results of the contextualization of the domains, as well as the dominance of EUPs as a discourse since

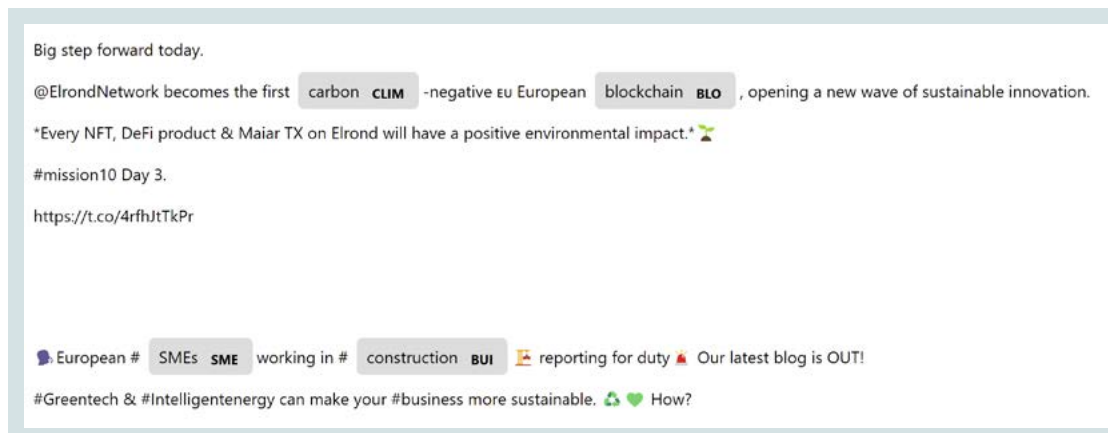
>

<sup>4</sup> For similar findings in the research community see Santarius and Wagner (2023, in this issue) and in the economy see Gotsch et al. (2023, in this issue).

<sup>5</sup> Modularity is a quality measure for clusterings: by maximising modularity a graph is subdivided in groups (clusters) that have a maximum number of within-group interrelations and a minimum number of external relations. For the calculation procedure see Brandes et al. (2008).



**FIGURE 4:** Trends in domain occurrence for each domain in English-language tweets relating to the discourse on sustainability and digitalization in Europe between 2010 and 2021. Line inflection points denote changes in yearly mentions of the domain as a percentage of mentions in the previous year. Transparency and width of lines relate to the share of all domain mentions a specific domain received in a given year (at the line inflection points). Reading example: The domain ICT made up a relatively large proportion of the overall discourse in 2012 and 2013 (wider and more solid line); it saw a large spike in attention in 2012 and 2016, compared to previous years (line inflection points). See table 1 for domain label reference.



**FIGURE 5:** Illustration of domain identification and domain co-occurrence in tweets.

2018, reinforce the impact of institutions, such as the EU, and thus, the political and regulatory power that structures can exert over public discourses on sustainability and digitalization.

In addition, the co-occurrence analysis of domains (figure 6) has allowed us to distinguish tentatively three phases in the Twitter discourse on sustainability and digitalization in Europe over time. The first phase of the discourse at the beginning of the decade centered much more strongly on the life-cycle impacts of digital technologies and was dominated by smart city concepts, as compared to the later stages.

The second phase, in the second half of the decade until the COVID-19 crisis, saw a second discursive hub emerge surrounding climate change. Two additional sub-discourses were identified, one connecting energy use both with digital infrastructure and the potential of smart grid technologies and a second concerning the potential for sustainability gains through digitalization in manufacturing and circular economy concepts.

During the COVID-19 crisis years, the third phase, climate change emerged as the dominant central discourse, being discussed based on a variety of domains. A distinct “smart rural” cluster also emerged during this time, including smart agriculture, forestry, and issues specific to digitalization in rural regions, such as broadband connectivity. The life-cycle impacts of digitalization were also being discussed again, as data centers were more frequently discussed in conjunction with energy use and climate impacts, as well as pandemic-induced changes in working patterns.

### Limitations

Our analyses of tweets from 2010 until 2021 are, of course, only a snapshot of the public discussions of sustainability and digitalization. Furthermore, beyond domain identification, we did not conduct an in-depth analysis of the presentation of the respective domains in terms of valence or arguments brought forward against or in favor of the current state of digitalization in the domain. Another limitation of our approach might be that we have missed possible additional domains that were not identified in the qualitative discovery step. Regarding our classifica-

tion step, we are confident that our rule-based approach has a relatively high precision in classifying domains, but it faces limitations in recall that could potentially be resolved using a statistical model for classification, though at the cost of increased complexity. Furthermore, we feel it is important to re-iterate that Twitter is a biased and flawed data source for understanding societal discourses in many aspects. As such, to generate a more encompassing map of the societal discourse, future work complementing ours would need to consider other mediums and fora in which discourse manifests, beyond elite discourses on specific social media platforms.

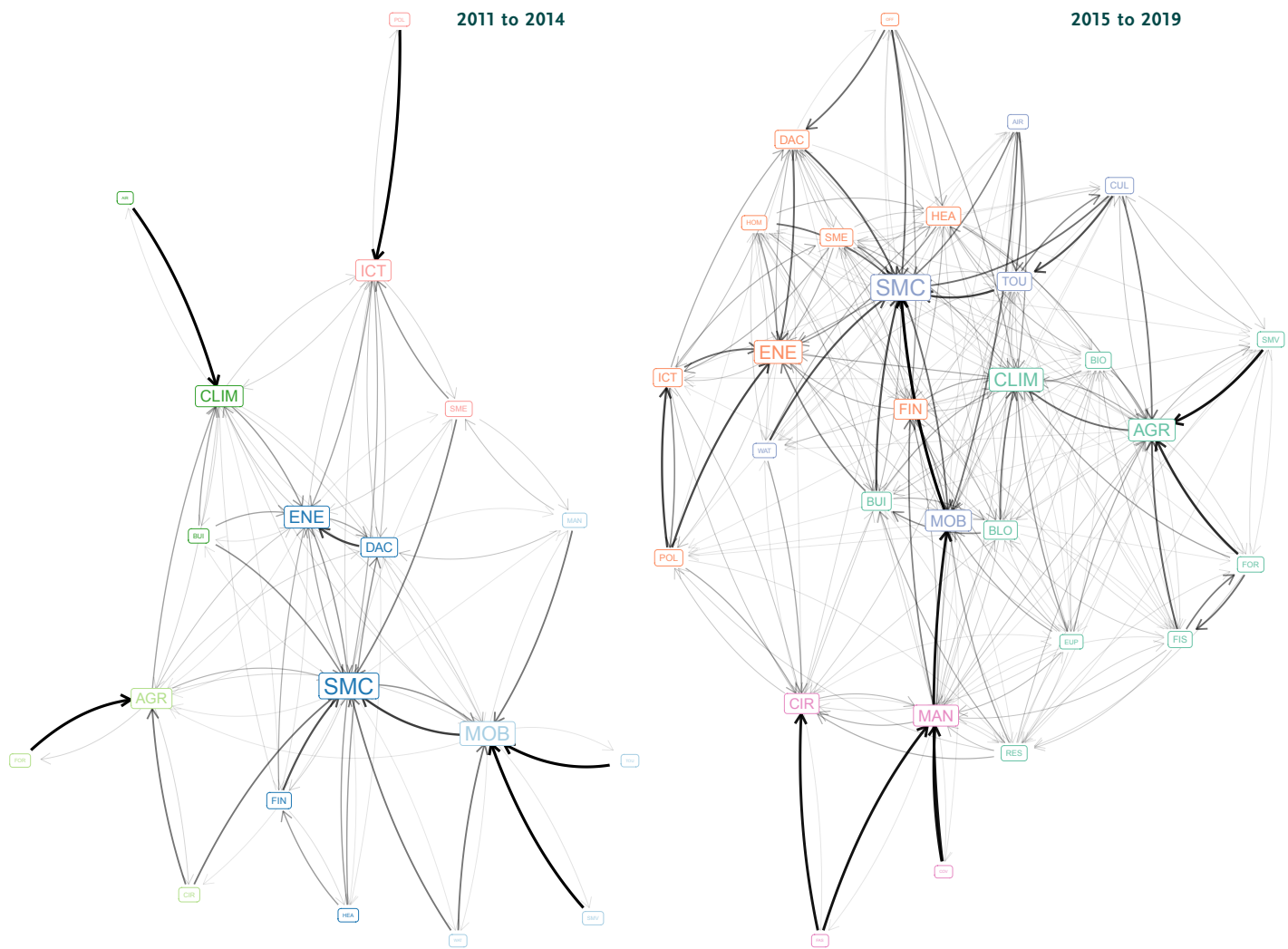
### Conclusion

This study aids in our understanding of how the discourse surrounding two of the most crucial socio-technical change processes of our time in Europe (and globally) – that is, sustainability transformations and digitalization – have evolved in the past decade. We were able to chart the development of a multi-faceted discourse using a multitude of domains over a decade, from concerns about the life-cycle impacts of ICT technology, to the rise in prominence of smart cities, to the establishment of climate change as a key discursive hub. Our results regarding domain clusters in the discourse point to potentials and the crucial importance of nexus approaches – which do not consider domains in isolation – in the research on and practices of digitalization and sustainability.

What does appear in the discourse is as interesting as what does not appear in our analysis, which comes upon the eve of the recent *Digital Reset* report by Digitalization for Sustainability (D4S 2022). We find the discursive structure very much in line with some of the concerns about the direction of digitalization raised in the report, namely, an over-reliance on efficiency justifications to champion digitalization accompanied by little critical reflection on economic growth or mention of sufficiency strategies. We also find an almost total absence of discussions on the structural impacts of digitalization on sustainability (Hilty and

>





Aebischer 2015), such as those brought about by big tech companies' pressure on regulatory environments (Andersen 2021). In this way, as a complement to conceptual reflections and empirical assessments of sustainability and digitalization, our results offer a possibility for research in this area to reflect the differences between societal and academic discussions. In some recent, quickly developing domains (e.g., finance or blockchain technologies), sustainability research should interpret our results as a call for increased research on these domains and for researchers to assert themselves more forcefully in the societal discourses surrounding them. In our opinion, our results further imply a normative responsibility of researchers to improve awareness of underrepresented topics in the critical discourse beyond academia. In addition, there is a need for more discussions of structural changes to the conditions for sustainability due to the power of big tech, the importance of sufficiency strategies, and the actual value of efficiency arguments in the discourse surrounding sustainability and digitalization.

**Acknowledgement:** We would like to thank three anonymous reviewers for their helpful comments.

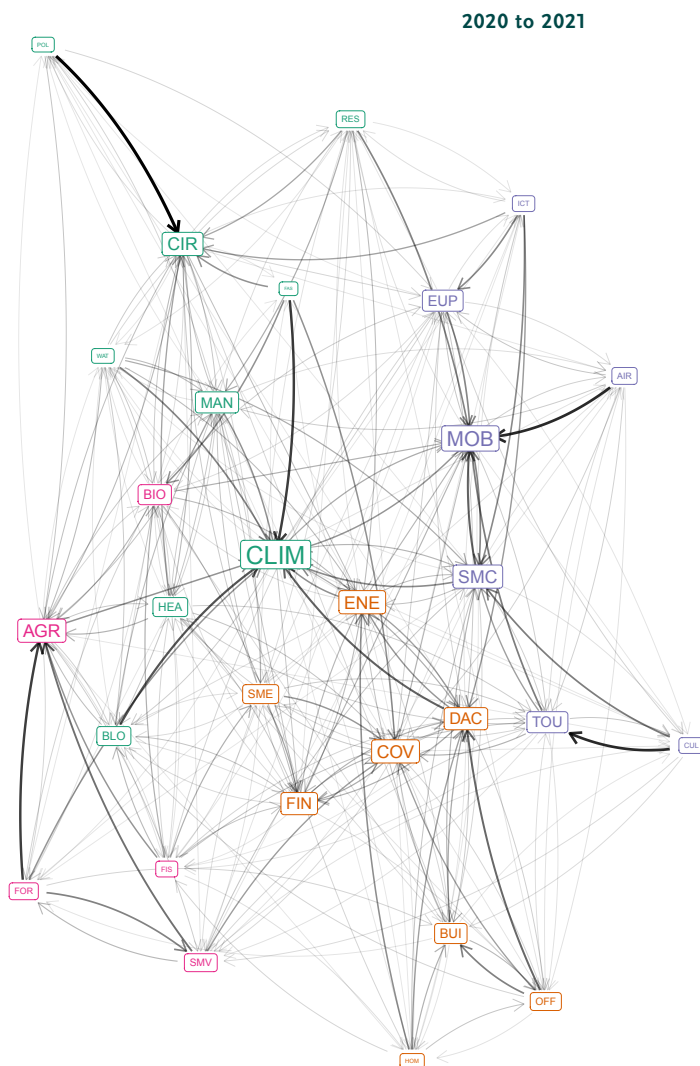
**Funding:** We thank the University of Zurich and the *Digital Society Initiative* for (partially) financing work on this article.

**Competing interests:** The authors declare no competing interests.

**Author contribution:** Both authors were involved in initial research design, data collection and analysis, manuscript drafting and writing the final manuscript.

## References

- Andersen, A. D. et al. 2021. On digitalization and sustainability transitions. *Environmental Innovation and Societal Transitions* 41: 96–98. <https://doi.org/10.1016/j.eist.2021.09.013>.
- Bockschecker, A., S. Hackstein, U. Baumöl. 2018. *Systematization of the term digital transformation and its phenomena from a socio-technical perspective: A literature review*. European Conference on Information Systems (ECIS) Research Paper 43. [https://aisel.aisnet.org/ecis2018\\_rp/43](https://aisel.aisnet.org/ecis2018_rp/43).
- Brandes, U. et al. 2008. On modularity clustering. *IEEE Transactions on Knowledge and Data Engineering* 20/2: 172–188. <https://doi.org/10.1109/TKDE.2007.190689>.



**FIGURE 6:** Interrelations of domains based on co-occurrences of domains in English-language tweets relating to the discourse on sustainability and digitalization in Europe across three phases between 2011 and 2021. The thickness of an arrow pointing from one domain  $i$  to another domain  $j$  is based on the number of co-occurrences  $i$  has with  $j$  as a percentage of  $i$ 's co-occurrences with all other domains. For every domain, the top ten co-occurrences are shown (if the domain has as many). Domains are sized by their overall number of co-occurrences. Domains were clustered by maximizing modularity (Brandes et al. 2008). Elements of the same cluster have the same color. See table 1 for domain label reference.

Brenner, B., B. Hartl. 2021. The perceived relationship between digitalization and ecological, economic, and social sustainability. *Journal of Cleaner Production* 315: 128128. <https://doi.org/10.1016/j.jclepro.2021.128128>.

Carlsen, H. B., S. Ralund. 2022. Computational grounded theory revisited: From computer-led to computer-assisted text analysis. *Big Data & Society* 9/1. <https://doi.org/10.1177/20539517221080146>.

D4S (Digitalization for Sustainability). 2022. *Digital reset. Redirecting technologies for the deep sustainability transformation*. Berlin: TU Berlin. <https://doi.org/10.14279/depositonce-16187.2>

Del Río Castro, G., M. C. González Fernández, A. Uruburu Colsa. 2021. Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *Journal of Cleaner Production* 280: 122204. <https://doi.org/10.1016/j.jclepro.2020.122204>.

Downs, A. 1972. Up and down with ecology: The issue-attention cycle. *Public Interest* 28: 38–50.

EC (European Commission). 2015. *A framework strategy for a resilient energy union with a forward-looking climate change policy*. COM/2015/080 final. Brussels: EC. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2015:80:FIN> (accessed January 24, 2023).

Galaz, V. et al. 2021. Artificial intelligence, systemic risks, and sustainability. *Technology in Society* 67: 101741. <https://doi.org/10.1016/j.techsoc.2021.101741>.

Gambini, M., T. Fagni, C. Senette, M. Tesconi. 2022. *Tweets2Stance: Users stance detection exploiting zero-shot learning algorithms on tweets*. arXiv:2204.10710 [cs.SI]. <http://arxiv.org/abs/2204.10710>.

Gotsch, M., N. Martin, E. Eberling, S. Shirinzadeh, D. Osiek. 2023. The contribution of data science applications to a green economy. *GAIA* 32/S1: 33–39. <https://doi.org/10.14512/gaia.32.S1.6>.

Hase, V., D. Mahl, M. S. Schäfer, T. R. Keller. 2021. Climate change in news media across the globe: An automated analysis of issue attention and themes in climate change coverage in 10 countries (2006–2018). *Global Environmental Change* 70: 102353. <https://doi.org/10.1016/j.gloenvcha.2021.102353>.

Hilty, L. M., B. Aebischer. 2015. ICT for sustainability: An emerging research field. In: *ICT Innovations for Sustainability: Advances in Intelligent Systems and Computing*. Edited by L. M. Hilty, B. Aebischer. Cham: Springer International. 3–36. [https://doi.org/10.1007/978-3-319-09228-7\\_1](https://doi.org/10.1007/978-3-319-09228-7_1).

Kiousis, S., M. McCombs. 2004. Agenda-setting effects and attitude strength: Political figures during the 1996 presidential election. *Communication Research* 31: 36–57. <https://doi.org/10.1177/0093650203260205>.

Lenz, S. 2021. Is digitalization a problem solver or a fire accelerator? Situating digital technologies in sustainability discourses. *Social Science Information* 60/2: 188–208. <https://doi.org/10.1177/05390184211012179>.

- Lichtenthaler, U. 2021. Digitainability: The combined effects of the mega-trends digitalization and sustainability. *Journal of Innovation Management* 9/2: 64–80. [https://doi.org/10.24840/2183-0606\\_009.002\\_0006](https://doi.org/10.24840/2183-0606_009.002_0006).
- Liu, Y. et al. 2019. RoBERTa: A robustly optimized BERT pretraining approach. arXiv:1907.11692. <https://doi.org/10.48550/arXiv.1907.11692>.
- Marenco, M., T. Seidl. 2021. The discursive construction of digitalization: A comparative analysis of national discourses on the digital future of work. *European Political Science Review* 13/3: 391–409. <https://doi.org/10.1017/S175577392100014X>.
- McCombs, M. E., D. L. Shaw. 1972. The agenda-setting function of mass media. *Public Opinion Quarterly* 36/2: 176–187. <https://doi.org/10.1086/267990>.
- Metag, J., F. Marcinkowski. 2014. Technophobia towards emerging technologies? A comparative analysis of the media coverage of nanotechnology in Austria, Switzerland and Germany. *Journalism* 15/4: 463–481. <https://doi.org/10.1177/1464884913491045>.
- Mohr, S., J. Höhler. 2021. Media coverage of digitalization in agriculture: An analysis of media content. SSRN. <https://ssrn.com/abstract=3971185> (accessed April 30, 2022).
- Montani, I. et al. 2022. explosion/spaCy: v3.2.2: Improved NER and Parser Speeds, Bug Fixes and More.
- Ott, B. L. 2017. The age of Twitter: Donald J. Trump and the politics of debasement. *Critical Studies in Media Communication* 34/1: 59–68. <https://doi.org/10.1080/15295036.2016.1266686>.
- Raworth, K. 2017. Why it's time for doughnut economics. *IPPR Progressive Review* 24/3: 216–222. <https://doi.org/10.1111/newe.12058>.
- Santarius, T., J. Wagner. 2023. Digitalization and sustainability: A systematic literature analysis of ICT for Sustainability research. *GAIA* 32/S1: 21–32. <https://doi.org/10.14512/gaia.32.S1.5>.
- Schmidt, E., J. Cohen. 2013. *The new digital age: Reshaping the future of people, nations and business*. New York: Knopf.
- Schneidewind, U. 2018. *Die Große Transformation. Eine Einführung in die Kunst gesellschaftlichen Wandels*. Frankfurt am Main: S. Fischer.
- Seele, P., I. Lock. 2017. The game-changing potential of digitalization for sustainability: Possibilities, perils, and pathways. *Sustainability Science* 12/2: 183–185. <https://doi.org/10.1007/s11625-017-0426-4>.
- Sloan, L. 2017. Who tweets in the United Kingdom? Profiling the Twitter population using the British Social Attitudes Survey 2015. *Social Media + Society* 3/1: 2056305117698981. <https://doi.org/10.1177/2056305117698981>.
- Soroka, S. N. 2002. Issue attributes and agenda-setting by media, the public, and policymakers in Canada. *International Journal of Public Opinion Research* 14/3: 264–285. <https://doi.org/10.1093/ijpor/14.3.264>.
- Steffen, W. et al. 2015. Planetary boundaries: Guiding human development on a changing planet. *Science* 347/6223: 736. <https://doi.org/10.1126/science.1259855>.
- Trask, A., P. Michalak, J. Liu. 2015. *sense2vec: A fast and accurate method for word sense disambiguation in neural word embeddings*. arXiv:1511.06388. <https://doi.org/10.48550/arXiv.1511.06388>.
- Williams, A., N. Nangia, S. Bowman. 2018. A broad-coverage challenge corpus for sentence understanding through inference. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies 1 (Long Papers)*. Edited by M. Walker, H. Ji, A. Stent. New Orleans: NAACL. 1112–1122. <https://doi.org/10.18653/v1/n18-1101>.
- Yates, S., E. Lockley. 2018. Social media and social class. *American Behavioral Scientist* 62/9: 1291–1316. <https://doi.org/10.1177/0002764218773821>.



#### Mario Angst

PhD political science. Currently postdoc and project lead at the *Digital Society Initiative (DSI)* of the University of Zurich, CH. Research interests: governance networks, sustainable digitalization, urban sustainability, policy in the Anthropocene, computational social science.



#### Nadine Strauß

PhD communication science. Currently assistant professor of strategic communication and media management at the Department of Communication and Media Research at the University of Zurich, CH. Research interests: sustainability communication, corporate communication, financial communication, and financial journalism.

## Nachhaltigkeit

# A-Z



## B like blueprint

Governments worldwide hope that digital technologies can provide key solutions. Yet this report shows that digitalisation, in its current and mainstream form, is rather aggravating than solving crises at hand. What is needed instead is a deep sustainability transformation that fundamentally reorganises the economy and all its sectors.

S. Lange, T. Santarius, L. Dencik, T. Diez, H. Ferreboeuf, S. Hankey, A. Hilbeck, L. M. Hilty, M. Höjer, D. Kleine, J. Pohl, L. Reisch, M. Ryghaug, P. Staab, T. Schwanen

#### Digital Reset

Redirecting Technologies for the Deep Sustainability Transformation

104 Seiten, Broschur, komplett vierfarbig mit zahlreichen Illustrationen, 22 Euro  
ISBN 978-3-98726-022-3

Bestellbar im Buchhandel und unter [www.oekom.de](http://www.oekom.de).  
Auch als E-Book erhältlich.

Die guten Seiten der Zukunft

