

HUOM! Tämä on alkuperäisen artikkelin rinnakkaistallenne. Rinnakkaistallenne saattaa erota alkuperäisestä sivutukseltaan ja painoasultaan.

PLEASE NOTE! This is an electronic self-archived version of the original article. This reprint may differ from the original in pagination and typographic detail.

Viittaa alkuperäiseen lähteeseen osoitteessa:

Cite the final publication found at :

[Compass Conference: Transferable Skills for Research & Innovation, 2023, October 4 – 5, Helsinki, Finland.](#)

© 2024 The Author(s) & Haaga-Helia University of Applied Sciences

## **AGORA: Automated Generation of Test Oracles for REST APIs**

Alonso, Juan C.<sup>1\*</sup>; Segura, Sergio<sup>2</sup>; Ruiz-Cortés, Antonio<sup>2</sup>

*\*Corresponding author - presenter*

<sup>1</sup> *jvalenzuela@us.es; University of Seville, Spain*

<sup>2</sup> *University of Seville, Spain*

Test case generation tools for REST APIs have grown in number and complexity in recent years. However, their capabilities for automated input generation contrast with the simplicity of their test oracles (i.e., mechanisms for determining whether a test has passed or failed), limiting the types of failures they can detect. AGORA is an approach for the automated generation of test oracles for REST APIs through the detection of *invariants* (properties of the output that should always hold). AGORA aims to learn the expected behavior of an API by analyzing previous API requests and their corresponding responses. For this, we extended Daikon, a tool that applies machine learning techniques to dynamically detect likely invariants, including the definition of new types of invariants and the implementation of an *instrumenter* called Beet. Beet converts any OpenAPI specification and a collection of API requests to a format processable by Daikon.

AGORA achieved a total precision of 81.2% when tested on a dataset of 11 operations from 7 industrial APIs. AGORA revealed 11 bugs in APIs with millions of users: Amadeus, GitHub, Marvel, OMDb and YouTube. Our reports have guided developers in improving their APIs, including bug fixes and documentation updates in GitHub.

A preliminary version of this work obtained the first prize in the ESEC/FSE 2022 Student Research Competition (Alonso, Automated Generation of Test Oracles for RESTful APIs, 2022) and the second prize in the ACM SRC Grand Finals (Alonso, SRC Grand Finalists 2023,

2023). An extended version of this work has been accepted for publication in the ISSTA 2023 conference.

**Keywords:** REST APIs, test oracles, automated testing

## **Background of the study**

Web *Application Programming Interfaces* (APIs) allow heterogeneous software systems to interact over the network (Richardson, Amundsen, & Ruby, 2013). Companies such as Google, Microsoft and Apple provide web APIs to enable their integration into third-party systems. Most web APIs adhere to the REpresentational State Transfer (REST) architectural style, being known as REST APIs (Fielding, 2000). REST APIs are specified using the OpenAPI Specification (OAS) (OpenAPI Specification, 2023), which defines a standard notation for describing them in terms of operations, input parameters, and output formats.

Automated test case generation for REST APIs is a thriving research topic (Kim, Xin, Sinha, & Orso, 2022). Despite the promising results of existing proposals, they are all limited by the types of errors they can detect: uncontrolled server failures (labelled with a 500 status code) and disconformities with the API specification (e.g., the API response contains a field that is not present in the specification). Generally, there are dozens or hundreds possible test oracles for every API operation. For example, if we perform a search for songs in the Spotify API by setting a maximum number of results (*limit* parameter), the size of the response field containing an array of songs should be equal to or smaller than the value of *limit*. Currently, creating these test oracles is a costly and manual endeavour that requires domain knowledge.

## Aim of the study

We present AGORA, the first approach for the automated generation of test oracles for REST APIs. AGORA is based on the detection of *invariants*, properties of the output that should always hold, e.g., a JSON property from the response should always have the same value as one of the input parameters. For this, we developed an extension (so-called instrumenter) of Daikon (Ernst, et al., 2007), a popular tool that uses machine learning techniques for dynamic detection of likely invariants. Generated invariants can be used as potential test oracles, and they can also help to identify unexpected behaviour (bugs) in the software under test.

## Methodology

Figure 1 shows an overview of AGORA. At the core of the approach is Beet, a novel Daikon instrumenter. Beet receives three inputs: 1) the OAS specification of the API under test, 2) a set of API requests, and 3) the corresponding API responses. As a result, Beet returns an instrumentation of the API requests consisting of a declaration file (describing the format of the API operations inputs and outputs) and a data trace file (specifying the values assigned to each input parameter and response field in each API call). This instrumentation is processed by our customized version of Daikon, resulting in a set of likely invariants that can be used as test oracles.

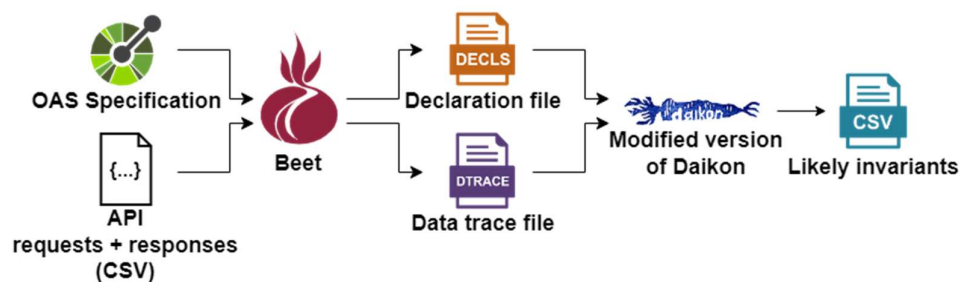


Figure 1: Workflow of AGORA

Beet has been implemented in Java and is open source. Our GitHub repository contains an exhaustive description of the instrumentation process with examples (Beet Repository, 2023).

To identify invariants that could be used as test oracles, we used a benchmark of 40 APIs (702 operations) systematically collected by the authors for a previous publication (Alonso, Martin-Lopez, Segura, Garcia, & Ruiz-Cortes, 2023).

We implemented 22 new invariants, suppressed 36 default Daikon invariants, and activated 9 invariants disabled by default in Daikon. This version of Daikon supports 105 distinct types of invariants, classified into five categories:

- **Arithmetic comparisons (48 invariants).** Specify numerical bounds (e.g., *size(return.artists[]) >= 1*) and relations between numerical fields.
- **Array properties (23 invariants).** Represent comparisons between arrays, such as subsets, supersets, or fields that are always member of an array (e.g., *return.hotel.hotelId in input.hotelIds[]*).
- **Specific formats (22 invariants).** Specify restrictions regarding the expected format (e.g., *return.href is Url*) or length of string fields.
- **Specific values (9 invariants).** Restrict the possible values of fields (e.g., *return.visibility one of {"public", "private"}*).
- **String comparisons (3 invariants).** Specify relations between string fields, such as equality (e.g., *input.name == return.name*) or substrings.

We refer the reader to AGORA GitHub repository (Beet Repository, 2023) for a more detailed description of each invariant.

## Results

For our experiments, we resorted to a set of 11 operations from 7 industrial APIs. For each operation, we automatically generated and executed API calls using the RESTest (Martin-Lopez, Segura, & Ruiz-Cortés, 2020) framework until obtaining 10K valid API calls per operation (110K calls in total). According to REST best practices (Richardson, Amundsen, & Ruby, 2013), we consider an API response as valid if it is labelled with a 2XX HTTP status code. We performed the following experiments to assess the effectiveness of AGORA for generating test oracles and detecting failures:

### **Experiment 1: Test oracle generation.**

For each API operation, we randomly divided the set of automatically generated requests (10K) into subsets of 50, 100, 500, 1K and 10K requests. Then, we ran AGORA using each subset as input and computed precision by manually classifying the inferred invariants as true or false positives. True positive invariants describe properties of the output that should always hold and therefore are valid test oracles. A false positive reflects a pattern that has been observed in all the API requests and responses provided as input but does not represent the expected behavior of the API.

We compared the effectiveness of AGORA against the default version of Daikon in identifying likely invariants for REST APIs. In both cases, Beet was used to transform API specifications and requests into inputs for Daikon. With this experiment, we aim to answer the following research questions:

***RQ1: Effectiveness of AGORA to generate test oracles.*** Table 1 shows the results for each API operation, set of API requests (50, 100, 500, 1K, 10K) and approach (AGORA vs default

Daikon). The columns labelled with “I” and “P” represent the number of likely invariants detected and the total precision, respectively.

When learning from the whole dataset, AGORA obtained a total precision of 81.2%, whereas the original configuration of Daikon achieved 51.4%. AGORA outperformed the default configuration of Daikon in all the API operations.

**RQ2: Impact of the size of the input dataset to the precision of AGORA.** The test suites of 50 requests seem to offer the best trade-off between effectiveness and cost of generating API requests. The total precision of AGORA improved from 73.2% with 50 API requests, to 81.2% with the complete dataset.

API - Operation	50 API calls				100 API calls				500 API calls				1K API calls				10K API calls			
	Daikon		AGORA		Daikon		AGORA		Daikon		AGORA		Daikon		AGORA		Daikon		AGORA	
	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)	I	P (%)
AmadeusHotel-hotelOffers	109	21.1	117	52.1	136	16.9	114	56.1	116	22.4	108	64.8	107	24.3	107	66.4	99	26.3	106	67.9
GitHub-createOrgRepo	82	95.1	198	98	82	95.1	198	98	80	96.2	198	98.5	80	96.2	198	98.5	80	96.2	198	98.5
GitHub-getOrgRepos	45	40	150	84.7	40	45	147	88.4	39	46.2	149	87.9	39	46.2	150	88	38	47.4	148	89.2
Marvel-comicById	178	29.8	115	47.8	194	28.9	127	46.5	178	33.7	119	52.1	167	35.9	106	58.5	140	45.7	96	65.6
OMDB-byIdOrTitle	7	57.1	16	93.8	7	57.1	16	93.8	7	57.1	16	93.8	8	50	17	88.2	7	57.1	16	93.8
OMDB-bySearch	4	100	5	100	7	57.1	7	71.4	5	80	6	83.3	5	80	6	83.3	6	83.3	7	85.7
Spotify-createPlaylist	22	100	41	100	22	100	41	100	22	100	41	100	22	100	41	100	22	100	41	100
Spotify-getAlbumTracks	46	45.7	68	85.3	45	46.7	67	86.6	42	50	66	87.9	42	50	66	87.9	41	53.7	66	89.4
Spotify-getArtistAlbums	53	43.4	55	81.8	49	49	52	88.5	35	68.6	50	92	32	75	50	92	31	83.9	52	92.3
Yelp-getBusinesses	60	28.3	30	40	55	30.9	33	36.4	46	37	25	48	45	37.8	23	52.2	41	39	22	50
YouTube-listVideos	228	31.6	194	57.2	227	32.2	199	56.3	218	35.8	191	62.3	225	36	200	61.5	201	41.3	196	65.3
<b>TOTAL</b>	<b>834</b>	<b>40.2</b>	<b>989</b>	<b>73.2</b>	<b>864</b>	<b>39.4</b>	<b>1001</b>	<b>73.5</b>	<b>788</b>	<b>44.5</b>	<b>969</b>	<b>77.8</b>	<b>772</b>	<b>45.9</b>	<b>964</b>	<b>78.8</b>	<b>706</b>	<b>51.4</b>	<b>948</b>	<b>81.2</b>

Table 1: Test oracle generation. I="Number of likely invariants", P="Precision (% valid test oracles)"

## Experiment 2: Failure detection.

To evaluate the effectiveness of the generated test oracles in detecting failures in the APIs under test, we systematically seeded *errors* in API responses using JSONMutator (JSONMutator, 2023), a tool that applies different mutation operators on JSON data, e.g., removing an array item.

For each API operation, we transformed the test oracles derived from the set of 50 test cases into executable assertions. Then, for each API operation, we randomly selected 1K API responses from the set of 10K test cases generated by RESTest.

We used JSONMutator to introduce a single error on each API response. Then, we ran the assertions and marked the failure as detected if at least one of the test assertions (i.e., test oracles) was violated. We repeated this process 10 times per operation to minimize the effect of randomness computing the average percentage of failures detected.

**RQ3: Effectiveness of the generated test oracles for detecting failures.** Table 2 shows the number of test assertions (i.e., test oracles) and the Failure Detection Ratio per API operation. Overall, test oracles generated by AGORA identified 57.2% of the failures.

API - Operation	Assertions	FDR (%)
AmadeusHotel-getMultiHotelOffers	61	60
GitHub-createOrganizationRepository	194	92.3
GitHub-getOrganizationRepositories	127	62.9
Marvel-getComicById	55	37.7
OMDB-byIdOrTitle	15	36.5
OMDB-bySearch	5	20.6
Spotify-createPlaylist	41	84.8
Spotify-getAlbumTracks	58	70.6
Spotify-getArtistAlbums	45	76.5
Yelp-getBusinesses	12	23.5
YouTube-listVideos	111	64.2
<b>TOTAL</b>	<b>724</b>	<b>57.2</b>

Table 2: Failure Detection Ratio per API operation.

## Detected Faults

AGORA detected 11 domain specific bugs in 7 operations from 5 APIs with millions of users worldwide. For example, one of the reported invariants in the Amadeus Hotel API led to the identification of 55 hotel room offers with zero beds (*return.room.typeEstimated.beds* $\geq 0$ ). When performing a search in YouTube using the *regionCode* input parameter, the returned



videos must be available in the provided region. However, a violation of the invariant `input.regionCode in return.contentDetails.regionRestriction.allowed[]`, led us to detect 81 cases in which the returned videos were not available in the provided region. These errors have been confirmed by the developers.

## Conclusions

We present the following original research and engineering contributions:

- AGORA, a black-box approach for the automated generation of test oracles for REST APIs.
- Beet, a Daikon instrumenter for REST APIs readily integrable into existing test case generation tools for REST APIs.
- A version of Daikon supporting the detection of 105 distinct types of invariants in REST APIs.
- An empirical evaluation in terms of precision and failure detection in 11 operations from 7 industrial APIs, including reports of 11 real-world bugs.

Future lines of work include:

- Automated generation of assertions from the reported invariants. This extension will be performed during a research stay with Professor Michael Ernst, creator of Daikon.
- Deploying AGORA to ease its integration into other applications. We have been contacted by Schemathesis (Schemathesis, 2023), a company that offers testing as a service of REST APIs, to implement AGORA into their framework.

## References

- Alonso, J. (2022). Automated Generation of Test Oracles for RESTful APIs. *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, (pp. 1808–1810). Singapore.
- Alonso, J. (2023). *SRC Grand Finalists 2023*. Retrieved from <https://src.acm.org/grand-finalists/2023>
- Alonso, J., Martin-Lopez, A., Segura, S., Garcia, J., & Ruiz-Cortes, A. (2023). ARTE: Automated Generation of Realistic Test Inputs for Web APIs. *IEEE Transactions on Software Engineering*, 348 - 363.
- Beet Repository*. (2023). Retrieved from <https://github.com/isa-group/Beet>
- Ernst, M., Perkins, J., Guo, P., McCamant, S., Pacheco, C., Tschantz, M., & Xiao, C. (2007). The Daikon system for dynamic detection of likely invariants. *Science of Computer Programming*, 35-45.
- Fielding, R. T. (2000). *Architectural Styles and the Design of Network-based Software Architectures*. University of California, Irvine.
- JSONMutator*. (2023). Retrieved from <https://github.com/isa-group/JSONmutator>
- Kim, M., Xin, Q., Sinha, S., & Orso, A. (2022). Automated Test Generation for REST APIs: No Time to Rest Yet. *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*, (pp. 289–301). Virtual, South Korea.

Martin-Lopez, A., Segura, S., & Ruiz-Cortés, A. (2020). RESTest: Black-Box Constraint-Based Testing of RESTful Web APIs. *International Conference on Service-Oriented Computing*, (pp. 459-475).

*OpenAPI Specification*. (2023). Retrieved from <https://www.openapis.org>

Richardson, L., Amundsen, M., & Ruby, S. (2013). *RESTful Web APIs*. O'Reilly Media, Inc.

*Schemathesis*. (2023). Retrieved from <https://schemathesis.io>

## **Analysing opinions on the COVID-19 passport using Natural Language Processing (NLP)**

Aguilar Moreno, J.A.<sup>1\*</sup>; Palos-Sanchez, P.R.<sup>2</sup>; Pozo-Barajas, R.<sup>3</sup>.

*\*Corresponding author - presenter*

<sup>1</sup> *e-mail: juanantonio892@gmail.com; Universidad de Sevilla, Spain*

<sup>2</sup> *Universidad de Sevilla, Spain*

<sup>3</sup> *Universidad de Sevilla, Spain*

### **Abstract**

Social networks generate a large amount of information on a daily basis that can be analysed to understand phenomena and implement improvements for society. This article uses natural language processing techniques to identify the general opinion of YouTube comments on videos about the implementation of the COVID-19 passport, one of the most controversial measures to contain the advance of the pandemic while reactivating the economy, the reasons for this and the aspects that generate the most interest.

The Rapidminer software is used to find the most frequent words, positive and negative feelings and a classification by topics.

### **Keywords**

Natural Language Processing; Sentiment Analysis; Topics Analysis; RapidMiner; COVID-19 passport

### **Literature review**

The gradual return to normality after the COVID-19 pandemic has been accompanied by measures to contain a further increase in the number of infections at the same time as the

economy is reactivated, One of the most hotly debated is the introduction of the COVID-19 passport by the European Union, a document required for travel without the need for testing or quarantine and access to certain premises, including workplaces, which gives its holder a nominative accreditation of a higher level of immunity due to having recently overcome the disease or having been vaccinated with the required doses (Gstrein, 2021; Shin et al., 2021).

The political implications of COVID-19 passports are significant, requiring a balance between various interest groups and addressing ethical concerns such as health system inequities and personal choice (Sharun et al., 2021). Previous research analyzing Twitter opinions found a majority with favorable attitudes towards the passports, although concerns centered on the lack of consensus for a standardized certificate and ethical considerations like the digital divide and freedom (Khan et al., 2022). Conversely, a study conducted through focus group interviews revealed that the drawbacks outweighed the benefits, highlighting concerns about temporary immunity, virus mutation, and ethical issues (Fagnoli et al., 2021). Supporters emphasized economic recovery and a return to normalcy.

### **Aim of the study, originality and novelty**

The central objective of this work is to identify the general opinion on the implementation of the COVID-19 passport, the reasons for it and the aspects that generate the most interest to study possible ways of improvement. In line with this objective, the research questions are as follows:

RQ1: Has the COVID-19 passport been perceived as a measure of success or failure among users?

RQ2: What are the main issues of concern or most frequently mentioned in the comments?

In order to answer these questions, this work has used a novel methodology based on NLP. Social media has become an important means of communication for society and an important source of easily accessible data for opinions and emotions, which in situations of major humanitarian disasters can be beneficial for the public sector (Malawani et al., 2020).

To this end, Rapidminer software is used to analyse comments on YouTube videos related to the COVID-19 passport.

Natural Language Processing (NLP) encompasses techniques such as Sentiment Analysis, which allow machines to process human language (López-Martínez and Sierra, 2020). Sentiment Analysis evaluates text content by assigning a numerical polarity based on the connotations of words (Saura et al., 2018), either by machine learning or by lexicon-based sentiment analysis (Kolchyna et al., 2015).

## Methodology

A series of steps were followed: extracting comments with web scraping software, filtering comments in Microsoft Excel, analyse them in RapidMiner software, and finally obtaining the results and visualisations.

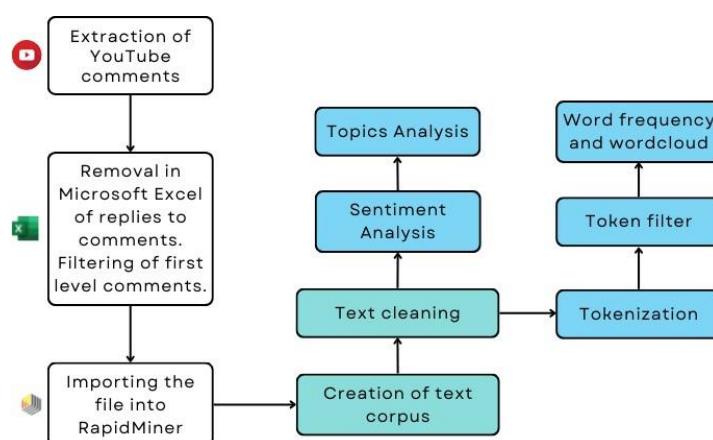


Figure 1. Methodological process

1. Extracting comments: A total of 24,182 YouTube comments were extracted from videos of news related to the COVID-19 passport on British press media channels: BBC, The Sun and Daily Mail. Following the recommendations of Palos et al. (2022), the videos were selected according to subject matter, popularity, number of comments and language. The final set consists of 11 videos with more than 250 comments each. The extraction was carried out with MAXQDA Analytics Pro 2022 software.
2. Filtering of top-level comments: Some of the comments extracted were responses to first level comments, so to avoid the bias they were removed in Microsoft Excel, leaving only the first level comments (10,833).
3. Analysis with RapidMiner: The operators included in the RapidMiner design interface and their usefulness within the process shown in the following image are detailed below:

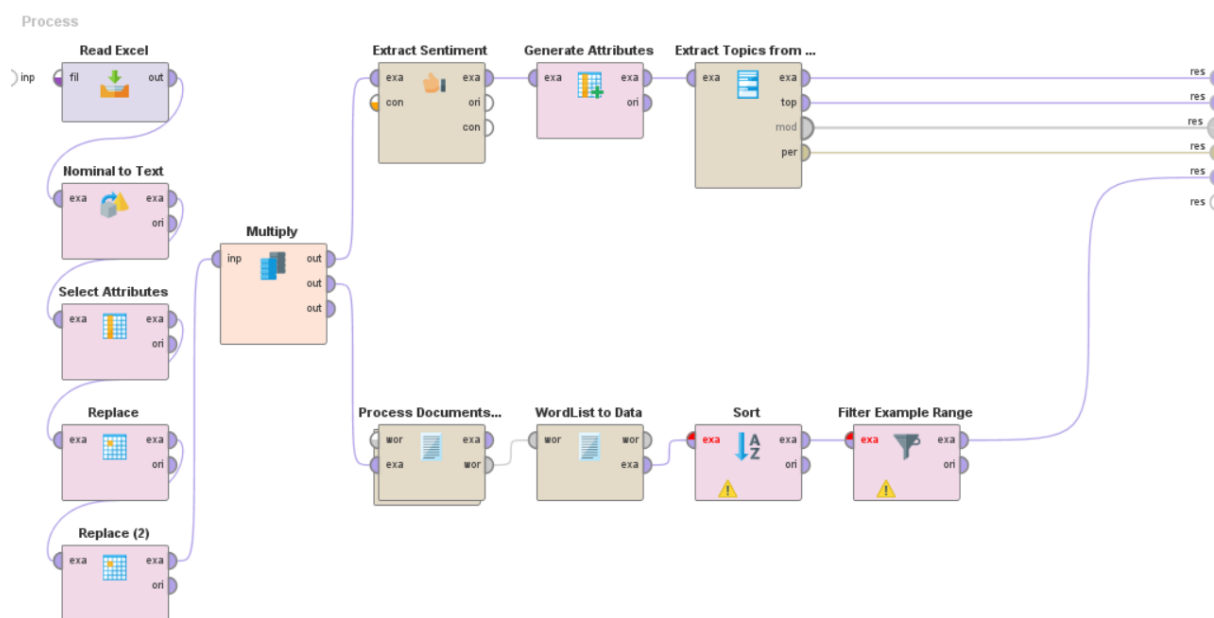


Figure 2. RapidMiner process

<b>Operator</b>	<b>Utility</b>
Read Microsoft Excel	It imports the Microsoft Excel file containing the comments.
Nominal to Text	It converts nominal attributes to string attributes
Select Attributes	It selects the column in .xlsx file to be considered for the analysis (column of comments text)
Replace	It removes unwanted characters to purify text: <code>https?://[-a-zA-Z0-9+&amp;@#/%?~_(!:,;]*[-a-zA-Z0-9+&amp;@#/%?~_()]\RT\s*@[^:]*:\s*[A-Za-z]+</code>
Multiply	It divides the process into two parts, each with its corresponding operators: <ul style="list-style-type: none"> <li>- Sentiment analysis and topic classification (Extract sentiment, Generate Attribute, Extract Topics from Data).</li> <li>- Word Cloud Generation (Process Document from Data, Word List to Data, Sort, Filter Example Range).</li> </ul>
Extract sentiment	It analyses the sentiment expressed in each of the comments using the model indicated in parameters (aylien, sentiwordnet, vader, etc.). Vader has been selected. It detects the relevant words included in each comment and assigns them a score previously established in the dictionary. It performs a global computation of all the words included in the same comment, from which the polarity is obtained.



Generate attributes	A conditional formula translates polarity into the expressions "positive", "negative" and "neutral".
Extract Topics from Data (LDA)	It groups the comments into several topics according to their content and determines the keywords of the topic according to importance.
Process Document from Data	It contains the following operators: 1. tokenise: splits each comment into a sequence of independent words (tokens). 2. Transform cases: converts all tokens to lower case. 3. Filter Stopwords: removes articles, pronouns, etc. 4. Filter tokens by length: filter words by length (4-25 characters).
WordList to Data	It counts the occurrences of each word.
Sort	It sorts the words according to the number of occurrences.
Filter Example Range	It filters the 20 most frequent words.

*Table 1. Operators used in RapidMiner*

## Results

The results are term frequency ranking, topics ranking and sentiment analysis.

### 1. Frequency of terms

Words such as "government", with a total of 734 appearances; "stop"; "freedom"; "conspiracy"; "boris"; "time"; "country", which refer to ethical issues such as freedom, already mentioned in

the bibliographic review as one of the most contradictory aspects, as well as the differences between "country" countries. Also the need to "buy time" to "stop" the virus, seen in a positive sense, or "stop COVID-19 Passport" in a detrimental sense, which also makes sense when including words related to politics and its leaders in the UK, government and conspiracy theories.

<b>No.</b>	<b>Word</b>	<b>Documents</b>	<b>Total</b>
1	vaccine	1678	2247
2	people	1640	2234
3	covid	813	1084
4	passport	776	894
5	passports	664	746
6	government	647	734
7	go	558	633
8	take	524	597
9	want	485	550
10	virus	482	639
11	vaccinated	449	624
12	stop	428	489
13	time	423	458
14	world	392	462
15	freedom	369	434
16	conspiracy	358	404
17	vaccines	355	456

18	country	347	378
19	boris	346	377
20	visit	343	392

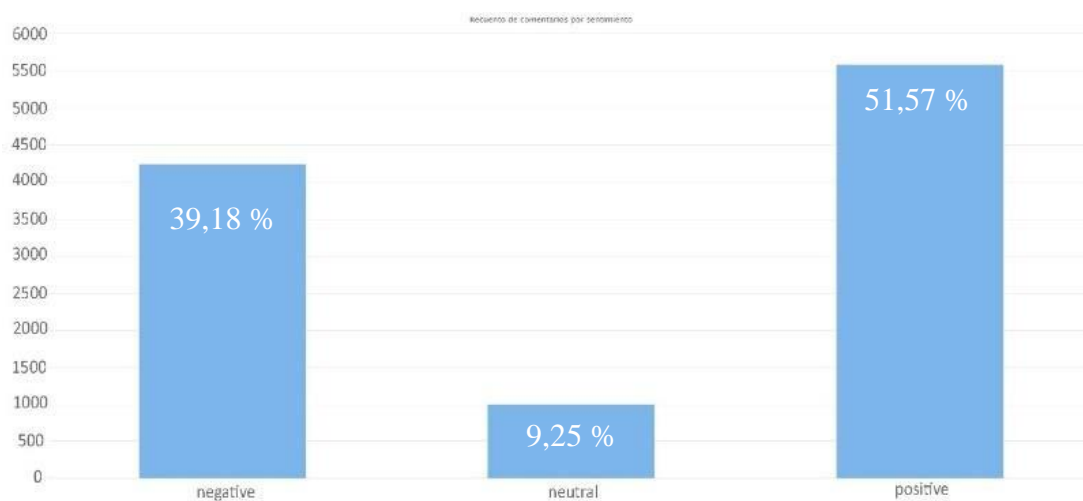
Table 2. Frequency of words in RapidMiner



Figure 3. Word cloud

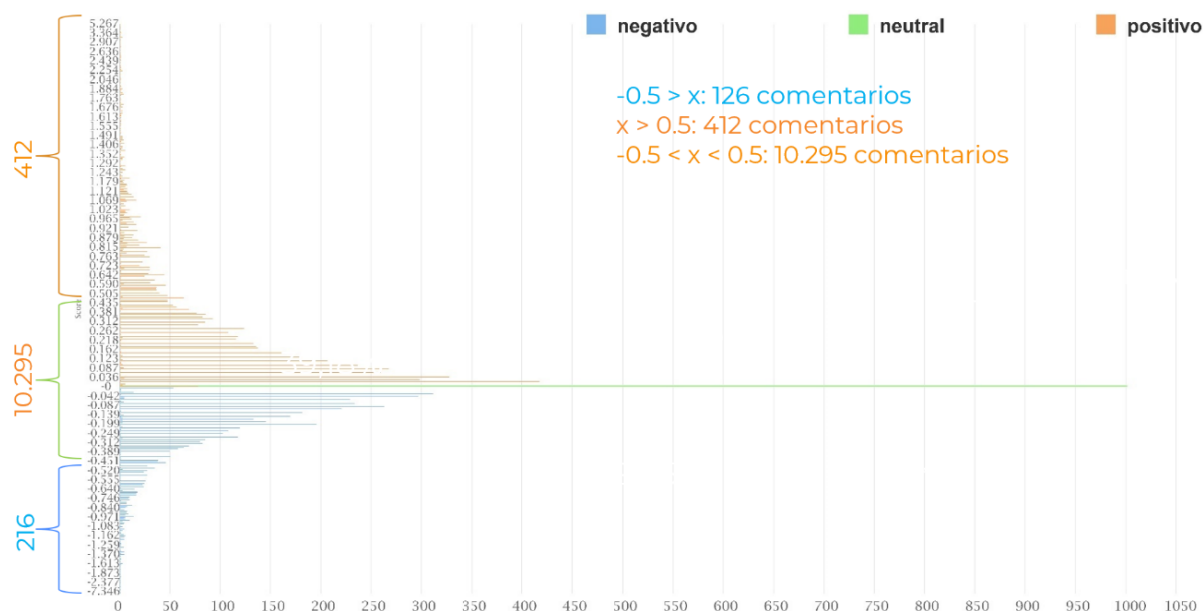
## 2. Sentiment analysis

The sentiment analysis statistics show a majority of positive comments. The following graph shows the count of comments after sentiment analysis: 51.57% positive, 39.18% negative, 9.25% neutral.



*Graph 1. Comment count by sentiment*

However, this information is complemented by the fact that only 126 comments have a score lower than -0.5 and 412 have a score higher than 0.5. The rest of the comments are between -0.5 and 0.5. As can be seen below in the next graph, the majority of comments are in values tending towards neutrality, either towards the positive side or towards the negative side, forming a Gaussian bell. It is necessary to remember that only the following are considered to be neutral scored with an exact 0. The rest are positive or negative whether they have a greater or lesser burden of such a feeling.



Graph 2. Distribution of comments according to their polarity

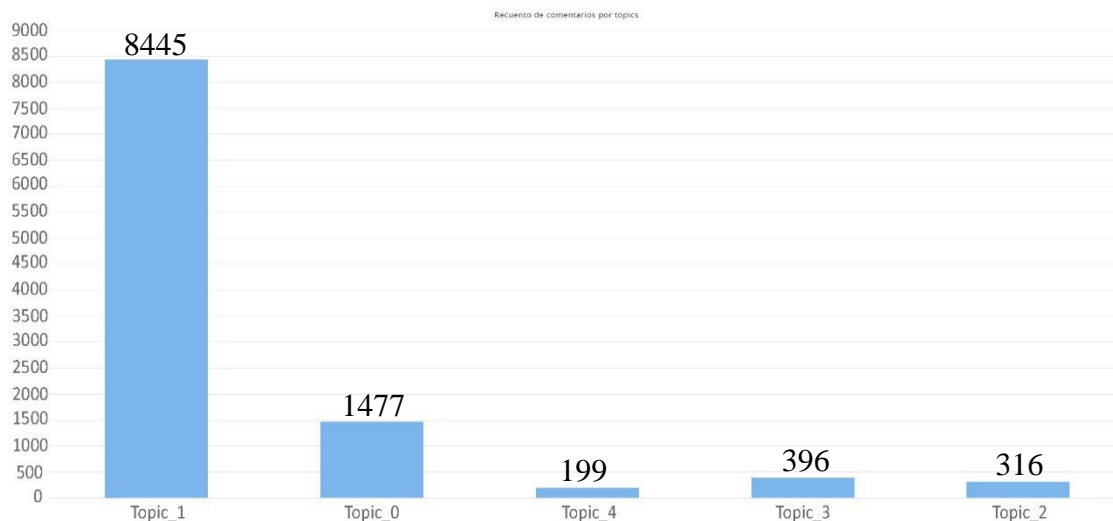
### 3. Classification according to topics

The comments have been automatically grouped according to their content into different themes using the Latent Dirichlet Allocation (LDA) model. The result obtained is a classification into 5 groups of all the comments analysed, as well as a ranking of the most frequent terms in each of these 5 topics.

	<b>Topic 0</b>	<b>Topic 1</b>	<b>Topic 2</b>	<b>Topic 3</b>	<b>Topic 4</b>
<b>Name</b>	Conspiracy theory	General views on obtaining the passport	Health issues	Citizens' rights	Concerns about citizen security

<b>Frequently used terms</b>	Conspiracy, theory, Boris (Prime Minister of the UK), rights	person, vaccinations, passport, want, need, government	immunity, deaths, vaccines, effects, experimental, cases, health, health, immunity, deaths	Freedom, government, control, stop	Crime, offender, punishment, news, news
<b>Likes</b>	8030	39534	1689	1449	694
<b>Comments</b>	1477	8445	316	396	199
<b>Likes/comments</b>	5,43	4,68	5,34	3,65	3,48

*Table 3. Classification by topics*



Graph 3. Comment count per topic

### Conclusion, managerial implications and limitations

In conclusion, the methodology used to analyze comments on the COVID-19 passport has limitations. The polarity of sentiment in the comments may differ from the sentiment towards the passport, and ironic comments can cause prediction errors. Despite these limitations, the interpreted conclusions from the results indicate numerous criticisms both for and against the passport. Supporters argue for the need to resume daily life and revive the economy, while exercising caution against the virus. Opponents express concerns about the vaccine's efficacy. The most concerning issues include adverse vaccine side effects, unequal vaccine availability across countries (impacting the usefulness of the passport in a globalized world), and governmental management of the crisis. Future research should involve a more comprehensive analysis of identified criticisms and concerns, examining these aspects across different demographic groups using algorithms for gender, age, and location detection in social networks, which is an emerging field in artificial intelligence.

Learning from this pandemic, it is crucial to manage it globally to ensure global security. The lack of protection affects not only poor communities, but the entire planet. A clear and transparent policy is needed to address data privacy concerns and combat misinformation on social media, avoiding uncertainty among the population.

## References

- Shin, H., Kang, J., Sharma, A., & Nicolau, J. L. (2021). The Impact of COVID-19 Vaccine Passport on Air Travelers' Booking Decision and Companies' Financial Value: *Https://Doi.Org/10.1177/10963480211058475*, *XX*, No. *X*, 1–10.  
<https://doi.org/10.1177/10963480211058475>
- Sharun, K., Tiwari, R., Dhama, K., Rabaan, A. A., & Alhumaid, S. (2021). COVID-19 vaccination passport: prospects, scientific feasibility, and ethical concerns. *Https://Doi.Org/10.1080/21645515.2021.1953350*, *17*(11), 4108–4111.  
<https://doi.org/10.1080/21645515.2021.1953350>
- Saura, J. R., Reyes-Menendez, A., & Alvarez-Alonso, C. (2018). Do Online Comments Affect Environmental Management? Identifying Factors Related to Environmental Management and Sustainability of Hotels. *Sustainability*, *10*(9).  
<https://doi.org/10.3390/su10093016>
- Malawani, A. D., Nurmandi, A., Purnomo, E. P., & Rahman, T. (2020). Social media in aid of post disaster management. *Transforming Government: People, Process*



*and Policy*, 14(2), 237–260. <https://doi.org/10.1108/TG-09-2019-0088/FULL/PDF>

Lopez-Martinez, R., & Sierra, G. (2020). Research Trends in the International Literature on Natural Language Processing, 2000-2019 — A Bibliometric Study. *Journal of Scientometric Research*, 9, 310–318. <https://doi.org/10.5530/jscires.9.3.38>

Kolchyna, O., Souza, T. T. P., Treleaven, P., & Aste, T. (2015). *Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination*. <https://doi.org/10.48550/arxiv.1507.00955>

Khan, M. L., Malik, A., Ruhi, U., & Al-Busaidi, A. (2022). Conflicting attitudes: Analyzing social media data to understand the early discourse on COVID-19 passports. *Technology in Society*, 68, 101830. <https://doi.org/10.1016/J.TECHSOC.2021.101830>

GSTREIN, O. J. (2021). The EU Digital COVID Certificate: A Preliminary Data Protection Impact Assessment. *European Journal of Risk Regulation*, 12(2), 370–381. <https://doi.org/10.1017/err.2021.29>

Fargnoli, V., Nehme, M., Guessous, I., & Burton-Jeangros, C. (2021). Acceptability of COVID-19 Certificates: A Qualitative Study in Geneva, Switzerland, in 2020. *Frontiers in Public Health*, 9, 682365. <https://doi.org/10.3389/FPUBH.2021.682365/FULL>

## **Artificial intelligence in international business research: a systematic literature review**

Koskinen, J.\*

*johanna.koskinen@haaga-helia.fi; Haaga-Helia University of Applied Sciences & Turku School of Economics, Finland*

**Keywords:** artificial intelligence, AI, international business, international business research, systematic literature review

### **Introduction**

Artificial intelligence (AI) can be described as "a system's ability to correctly interpret external data, learn from that data, and use what it has learned to achieve specific goals and tasks through flexible adaptation" (Kaplan & Haenlein, 2019, p. 17). One of its main benefits is that it can automate a significant amount of not only mechanical tasks, but also cognitive ones (Von Krogh, 2018). Moreover, the new type of AI, "generative AI", such as ChatGPT, is specifically designed to create new content. Experts, such as Bill Gates (2023), are comparing AI's impact on business to the internet, which has had a transformative impact on international business, particularly on newer types of enterprises like International New Ventures (INVs) and Born Globals (Etemad, 2022). Given that AI is already in many parts of our lives, such as smartphones, it is not only important but also critical to understand its role in international business (IB).

My Ph.D. research topic is how AI impacts the internationalization of SMEs at the level of entrepreneurial ecosystems. I am aiming to contribute specifically to IB research. In order to dive deeper into the topic and to advance its theoretical analysis, I conducted a systematic literature review on current AI research in IB. Therefore, my research question is: What is the current state of knowledge regarding AI within IB research?

## Methods

I conducted a systematic literature review adapting the method presented by Kaartemo and Helkkula (2018) by searching and assessing literature, and analyzing and synthesizing the findings. The stages of my research are presented below.

Stage 1: I consulted an IB journal ranking article (Tüselmann et al., 2016) and I used the SCImago Journal Rank (SJR) indicator to assess the quality and impact of IB journals. I chose to focus specifically on the leading IB journals because they attract researchers to publish their highest quality research.

Stage 2: I searched for articles in the Scopus database with the main keywords related to artificial intelligence in the title, abstract or keywords: *artificial intelligence, AI, machine learning, deep learning, neural network, large language models, robots, robotics, natural language processing*. I assessed the results and excluded any articles, which mentioned only one part of a keyword, e.g. learning. I reviewed the full papers to determine whether the article was appropriate for this research. This resulted in 18 papers altogether: 6 theoretical or conceptual, 4 review, 6 empirical quantitative and 2 empirical qualitative papers.

The chosen journals are listed below with the time period covered and amount of publications found in brackets: *The Journal of International Business Studies (1975-8/2023, 4)*, *The Journal of World Business (1997-8/2023, 4)*, *Global Strategy Journal (2014-8/2023, 2)*, *Management and organization review (2005-8/2023, 0)*, *International Business Review (1993-8/2023, 3)*, *Management international review (2005-8/2023, 0)*, *Journal of international management (1998-2023, 3)*, *Thunderbird IB Review (2005-2023, 3)*, *Critical*

*Perspectives on International Business (2005-2023, 0), Transnational Corporations (2005-2023, 0) and Multinational Business Review (2003-2023,1).*

Stage 3: I analyzed the articles by documenting all relevant data in an excel, including journal and author names, year, purpose, method and main findings. I categorized them into three categories, which emerged from the data. The first two ones being the most relevant in terms of my PhD research.

## **Results and findings**

I grouped the articles within this systematic literature review (18) into three categories representing how AI is discussed in IB literature. Two articles were categorized into multiple categories. The main findings are presented below.

### *Category 1: Impact of advanced technologies on the firm (8 papers)*

The first stream of studies are focused mainly on advanced technologies as a whole, which AI is part of, and discusses the implications of it for IB at the firm-level as no studies solely focused on AI's impact were available. These are from a short period of time (2017-2023) indicating its novelty. They mainly emphasize the positive impacts of advanced technologies bringing attention to how they may:

- reduce transaction costs (Strange & Zucchella, 2017; Ahi et al., 2022)
- facilitate international division and management of labor (Strange & Zucchella, A., 2017; Ahi et al., 2022)

- improve productivity, and change the nature of work and how it is structured (Ahi et al., 2022; Brakman et al., 2021) for example via automation and robotization of work (Brakman et al., 2021)
- enhance decision-making (Strange & Zucchella, 2017; Ahi et al., 2022)
- simplify exchange of knowledge (Ahi et al., 2022)
- increase trust in business networks (Ahi et al., 2022)
- impact where to locate manufacturing activities (Strange & Zucchella, 2017)
- enable supply chains to become more resilient and flexible (Gupta & Jauhar, 2023)

Furthermore, it became evident that their application specifically in managing global talent (Malik et al., 2021; Brakman et al., 2021) and supply chains (Strange & Zucchella, 2017; Ahi et al., 2022; Gupta & Jauhar, 2023) has received significant academic attention.

Even though, many raise awareness to the negative impacts of advanced technologies such as cybersecurity risks and ethical concerns around data privacy (Strange, R. & Zucchella, A., 2017; Ahi et al., 2022; Benito et al., 2022), and long-term impact of AI on human development (Ahi et al., 2022), they identify gaps in current research on the negative aspects and encourage IB scholars to research them more (Ahi et al., 2022; Benito et al., 2022). Also, Ahi et al. (2022) emphasize the need to research the negative impacts on work, for example ethical issues, widening income inequality and unemployment.

#### *Category 2: Impact of AI on the business ecosystem (6 papers)*

The second stream of studies focused on the potential implications of AI on the business ecosystem. There is no discussion in the reviewed articles only specifically on this topic, but

they do bring out some important aspects in relation to this topic area. These are recent articles (2020-2022) demonstrating the topic's novelty and need for further research.

Not only is the ecosystem important in international scaling of digital solutions, such as AI (Brakman et al., 2021), but also AI is seen to have importance for the ecosystem. AI is seen to support companies to adapt to their business environments by supporting to tackle grand challenges such as environmental sustainability (Ahi et al., 2022), by being a vehicle for enhancing knowledge spillovers within the ecosystem (Cetindamar et al., 2020), and supporting managers to identify and analyze political risks (Hemphill & Kelley, 2021), while being seen as a grand challenge itself (Benito, 2022). Moreover, it's also important for governments to support different actors in the ecosystem in getting more productivity from existing systems and, thus improving current IT tools, because this focus on algorithm development may have impact on a country's competitive advantage, such as is the case in the U.S. being the largest contributor to algorithm development (Thompson et al., 2021).

### *Category 3: AI as a tool for conducting IB research (8 papers)*

The third stream of studies is focused on both describing the current state of methodological advances in IB research through the use of AI and employing it. This seems to be a topic of wide interest from a longer time period (2007-2023). Referring to the use of AI, these studies not only highlight the usefulness of AI in conducting primarily quantitative IB research (Lindner et al., 2022; Delios et al., 2023; Veiga et al., 2023), but also emphasize that by leveraging these new methods — making use of the deeper and increased amount of data available — is crucial for IB research to remain a relevant field of research in the future as well (Delios et al., 2023).

The majority of studies employ AI in their research (Nair et al., 2007; Brouthers, 2009; Garbe, 2009; Messner, 2022; Vuorio & Torkkeli, 2023). These studies utilize AI in researching collaborative venture formation (Nair et al., 2007), international market selection (Brouthers, 2009), international organizational structure (Garbe, 2009), cultural heterogeneity (Messner, 2022) and predicting early internationalization (Vuorio & Torkkeli, 2023).

## **Conclusions**

This systematic literature review explored the current state of research on AI in IB. The method was adapted from Kaartemo and Helkkula (2018) and this study analyzed a collection of 18 relevant articles and categorized them into three categories. The main contribution of this research is that it demonstrates that there is a lack of high quality research on specifically AI in the IB research field despite the current “hype” around AI. From a temporal perspective, it seems that this is a novel area of research because the majority of discussions have emerged only recently in IB. Despite the low number of publications, the three categories portray well the current state of knowledge.

The first category, "Impact of advanced technologies on the firm," discussed the impacts of AI within the broader context of advanced technologies. Despite the topic's novelty, the studies demonstrated the potential positive impacts of AI adoption, including reductions in transaction costs (Strange & Zucchella, 2017; Ahi et al., 2022), improvements in decision-making processes (Strange & Zucchella, 2017; Ahi et al., 2022), and the transformation of work structures (Ahi et al., 2022; Brakman et al., 2021). Also, it became apparent the majority of research focuses on the positive sides of AI for business and less emphasis is put on the

negative aspects (Ahi et al., 2022; Benito et al., 2022), which calls for more research on this topic area.

The second category, "Impact of AI on the business ecosystem," focused on the impact of AI on the business ecosystem. While the reviewed articles did not solely focus on this topic, they highlighted the significance of AI in facilitating adaptation (Ahi et al., 2022), knowledge sharing (Cetindamar et al., 2020), and identifying political risks (Hemphill & Kelley, 2021) within business environments. Despite the scarcity of research on this topic, the publications do bring insights into how AI is also impacting the business ecosystem.

The third and final category, "AI as a tool for conducting IB research", focused on either the current state of AI utilization in research or employed it themselves in IB research. It brought attention to AI being useful in IB research (Lindner et al., 2022; Delios et al., 2023; Veiga et al., 2023) and many examples of its use were found, which focused on researching for example collaborative venture formation (Nair et al., 2007), and predicting early internationalization (Vuorio & Torkkeli, 2023).

## **Future research**

The majority of articles focus on the impact of AI on the firm or AI as a tool in quantitative research. Less emphasis was paid to AI's impact at the ecosystem level, which I intend to research further in my PhD. Also, the majority of research focuses only on the positive sides of AI for business and less emphasis is put on the negative aspects (Ahi et al., 2022; Benito et al., 2022). AI in qualitative research could also be researched more. Overall, it was evident that it is important to develop IB research both by researching the impact of AI on IB, which



Ahi et al. (2022) bring out as well, and how AI may be utilized in IB research even further.

Based on this review I suggest the following broader topic areas for future research:

- What are the negative impacts of AI at different levels (individual, business, environment)?
- How does AI impact the internationalization of businesses at the ecosystem level?
- How can AI be employed to enhance qualitative IB research?

## Limitations

The main limitation concerned the specificity of AI terminology. Much of the existing IB research seems to focus on advanced technologies on a general level rather than focusing only on AI, a point also raised by Ahi et al. (2022). This brought about the challenge of needing to use more general level technology terminology in order to find articles, but at the same time questioning their relevance for this AI-centric research.

## References

Ahi, A. A., Sinkovics, N., Shildibekov, Y., Sinkovics, R. R., & Mehandjiev, N. (2022). Advanced technologies and international business: A multidisciplinary analysis of the literature. *International Business Review*, 31(4), 101967.

Benito, G. R., Cuervo-Cazurra, A., Mudambi, R., Pedersen, T., & Tallman, S. (2022). *The future of global strategy*. *Global Strategy Journal*, 12(3), 421-450.

Bill Gates (2023), The Age of AI has begun, The blog of Bill Gates, Retrieved from <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun> on 10.8.2023.

Brakman, S., Garretsen, H., & van Witteloostuijn, A. (2021). Robots do not get the coronavirus: The COVID-19 pandemic and the international division of labor. *Journal of International Business Studies*, 52, 1215-1224.

Brouthers, L. E., Mukhopadhyay, S., Wilkinson, T. J., & Brouthers, K. D. (2009). International market selection and subsidiary performance: A neural network approach. *Journal of World Business*, 44(3), 262-273.

Cetindamar, D., Lammers, T., & Zhang, Y. (2020). Exploring the knowledge spillovers of a technology in an entrepreneurial ecosystem—The case of artificial intelligence in Sydney. *Thunderbird International Business Review*, 62(5), 457-474.

Delios, A., Welch, C., Nielsen, B., Aguinis, H., & Brewster, C. (2023). Reconsidering, refashioning, and reconceptualizing research methodology in international business. *Journal of World Business*, 58(6), 101488.

Etemad, H. (2022). The emergence of international small digital ventures (ISDVs): Reaching beyond Born Globals and INVs. *Journal of International Entrepreneurship*, 20(1), 1-28.

Garbe, J. N., & Richter, N. F. (2009). Causal analysis of the internationalization and performance relationship based on neural networks—advocating the transnational structure. *Journal of international management*, 15(4), 413-431.

Gupta, M., & Jauhar, S. K. (2023). Digital innovation: An essence for Industry 4.0. *Thunderbird International Business Review*.

Hemphill, T. A., & Kelley, K. J. (2021). Artificial intelligence and the fifth phase of political risk management: An application to regulatory expropriation. *Thunderbird International Business Review*, 63(5), 585-595.

Kaartemo, V., & Helkkula, A. (2018). A Systematic Review of Artificial Intelligence and Robots in Value. Co-creation: Current Status and Future Research Avenues. *Journal of Creating Value*, 4(2), 211–228.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business horizons*, 62(1), 15-25.

Lindner, T., Puck, J., & Verbeke, A. (2022). Beyond addressing multicollinearity: Robust quantitative analysis and machine learning in international business research. *Journal of International Business Studies*, 53(7), 1307-1314.

Malik, A., De Silva, M. T., Budhwar, P., & Srikanth, N. R. (2021). Elevating talents' experience through innovative artificial intelligence-mediated knowledge sharing: Evidence from an IT-multinational enterprise. *Journal of International Management*, 27(4), 100871.

Messner, W. (2022). Advancing our understanding of cultural heterogeneity with unsupervised machine learning. *Journal of International Management*, 28(2), 100885.

Nair, A., Hanvanich, S., & Cavusgil, S. T. (2007). An exploration of the patterns underlying related and unrelated collaborative ventures using neural network: Empirical investigation of collaborative venture formation data spanning 1985–2001. *International Business Review*, 16(6), 659-686.

Strange, R., & Zucchella, A. (2017). Industry 4.0, global value chains and international business. *Multinational Business Review*, 25(3), 174-184.

Tatarinov, K., Ambos, T. C., & Tschang, F. T. (2023). Scaling digital solutions for wicked problems: Ecosystem versatility. *Journal of International Business Studies*, 54(4), 631-656.

Thompson, N. C., Ge, S., & Sherry, Y. M. (2021). Building the algorithm commons: Who discovered the algorithms that underpin computing in the modern enterprise?. *Global Strategy Journal*, 11(1), 17-33.

Tüselmann, H., Sinkovics, R. R., & Pishchulov, G. (2016). Revisiting the standing of international business journals in the competitive landscape. *Journal of World Business*, 51(4), 487-498.

Veiga, J. F., Lubatkin, M., Calori, R., Very, P., & Tung, Y. A. (2000). Using neural network analysis to uncover the trace effects of national culture. *Journal of International Business Studies*, 31, 223-238.

Von Krogh, G. (2018). Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing. *Academy of Management Discoveries*.

Vuorio, A., & Torkkeli, L. (2023). Dynamic managerial capability portfolios in early internationalising firms. *International Business Review*, 32(1), 102049.

## **Cognitive Cybersecurity: An approach applied to Phishing**

José Mariano Velo<sup>1\*</sup>; Ángel Jesús Varela-Vaca<sup>1</sup>; Rafael M. Gasca<sup>1</sup>

<sup>1</sup>IDEA Research Group, Universidad de Sevilla, Spain

[josvelmor2@alum.us.es](mailto:josvelmor2@alum.us.es), {ajvarela, gasca}@us.es

**Abstract.** Cognitive systems are computer systems that use the theories, methods, and tools of various computing disciplines to model cognitive tasks or processes emulating human cognition. One of the applications of these systems in Cybersecurity is the detection of phishing emails. Thousands or millions of emails are received daily, all processed by the different security filters to catalogue them in several classes as legitimate emails, spam emails and malware or phishing emails. Despite all the measures taken, there is always a percentage of malicious emails which, for various reasons, can pass these filters. Emails that, ironically, could be easily identified by security analysts if they could review them individually. For this reason, we will define a proposal to improve the detection of such attacks based on a cognitive approach. The proposal focuses on defining the process, potential techniques and metrics using engines for extended decision-making based on decision models.

**Keywords.** Phishing, Cognitive Computing, Deep Learning, AI, Machine Learning

**I Introduction & Background.** Cognitive Computing (Raghavan, 2016) is an emerging field product of the confluence of several “sciences” such as cognitive science, neuroscience, data science and other computer technologies. The union of those sciences and new computing technologies such as Artificial Intelligence (AI), Big Data, Machine Learning (ML), Cloud Computing and others should make it possible to model the process of human cognition

artificially. Computing cognition includes in its processes the modelling of cognition to provide a better understanding of the environment systems that analyse large amounts of data and thus to be able to generate results and predictions that improve decision making.

Cognitive systems are computer systems that employ the theories, methods, and tools of various computing disciplines to model tasks or processes cognitions emulating human cognition (Raghavan, 2016) in them. The human brain is seen as a highly parallelised information processor from the point of view of cognitive computing (Raghavan, 2016). These systems are different from traditional systems since they are adaptive and learn and evolve over time, incorporating the context in computing. They can “feel” the environment, think and act independently and especially they can deal with uncertain, ambiguous, and incomplete information, as well as retain in their memory important data for future use.

E-mail Security Systems Analyses are performed basically by checking blacklisted domains and so on, and they usually work well with global and indiscriminate phishing attacks campaigns since intelligence data associated with them are usually shared from first detection and incorporated into the mentioned reputation repositories for further use and query by other systems. However, those systems usually fail when they parse targeted phishing emails, meticulously prepared to bypass the filters and come from generally “clean” domains.

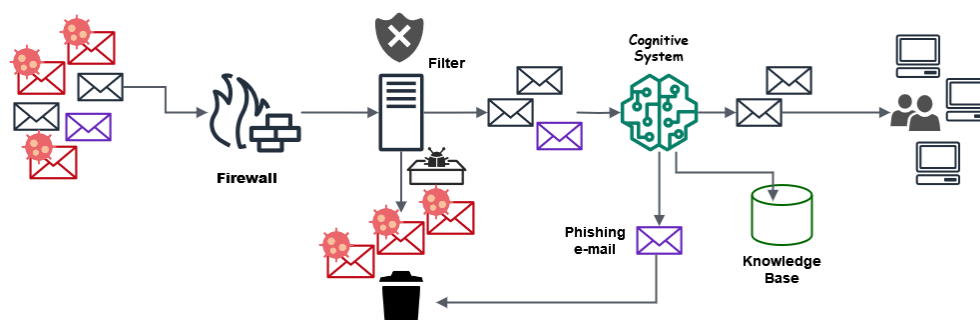


Figure 1. Incoming e-mail flow.

Therefore, we propose a cognitive system that will allow us to improve the analysis performed by the detection system of legitimate or malicious mail and that, if necessary, leave in the hands of the human analyst the last decision. This is, we intend to improve the analysis of emails with elaborate phishing, and for this, we will do it from an approximation based on a cognitive system. Some authors, like (Sumari, 2017), call CAI (Cognitive Artificial Intelligence) this approximation type. Figure 1 shows a scheme of the cognitive system that we propose within the flow of incoming emails in a standard organisation, being able to be observed as a mail with spear phishing (represented in purple colour) that has been able to bypass the initial filter, it is finally discarded after being analysed by the proposed system. Contributions main aspects of this work are:

- Define the complete cognitive process for the case of elaborate phishing.
- Study the metrics and parameters for the processing phase. of information.
- Define new metrics and parameters to analyse emails.
- Proposal for a decision-making system based on decision models.

**II Cognitive System Proposal for Cybersecurity.** Our proposed cognitive system, as can be seen in Figure 2, consists of three fundamental parts or phases divided each of them, into two sub-phases and all of them are related to the modelling of the cognitive process. Details and different phases will be explained below.

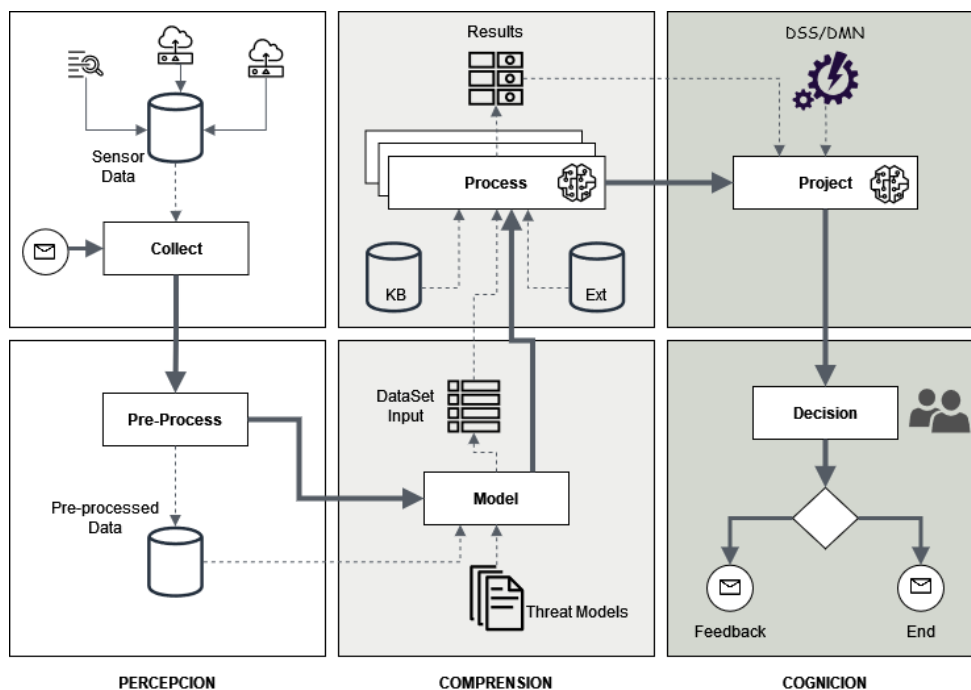


Figure 2. Proposed cognitive model.

**A. Perception.** Corresponds to the data collection phase and their preparation. In this phase, the system “observes” the environment and data is extracted and prepared for its subsequent analysis in the next phase by means of pre-processing. This phase would represent the equivalent of using the different senses in the human being as an initial part of cognition. The phase has two sub-phases. The first is called “Collect”, and the second called “Pre-process”. The first one, “Collect”, is the sub-phase in which we extract the data from the different sensors or sources of raw environmental data but in a controlled manner. As we can see in Figure 2, there are several input sources to a data warehouse where they are collected. Depending on how we apply the system, we can establish which types of data to extract from the environment. Or we can extract all available and then discriminate which are valid for a given task or process. The second of them, the "Pre-Process" sub-phase, is the one that processes the data entrance to clean, order and classify them to the next phase of the system. The mission of this second subphase is to eliminate what is useless and apart from generating new data based on



calculations and simple queries made, so that we can send to the next phase a data set as large and accurate as possible.

**B. Understanding.** In this phase, the data acquired are processed according to the established parameters and the model defined for a certain threat or cybersecurity problem. It also has two sub-phases, as in the previous phase. The first is called “Model” and models and orders the entry of pre-processed data with the “Threat Model” that we want to apply. Here are carried out several calculations and more elaborate queries to present a discretised input dataset as complete as possible to the next sub-phase. The second sub-phase, which we call “Process”, is the one that processes that supplied input and can also incorporate the feedback data obtained in previous iterations, as well as data from external sources such as STIX, Yara rules and others that may be necessary to complement and enrich the chosen “Threat Model”. It should be noted that these data cannot be incorporated into the stage of perception, given that they provide a schema, ontology, or higher-level knowledge of the input data. This second sub-phase takes place in what we have called Multiple Parallel Analysis (AMP). For this reason, it is one of the most important of the whole process. Here, different techniques will be executed in parallel through multiple analyses to produce the various outputs in a block of results, such as AI/ML, NLP (Natural Language Processing), visual analysis and others, as seen in Figure 3. To make a comparison with human cognition, this is where the brain processes all the information it has obtained from the different senses, complementing it with level processes superior as language, intelligence, and others to understand what is happening in the environment.



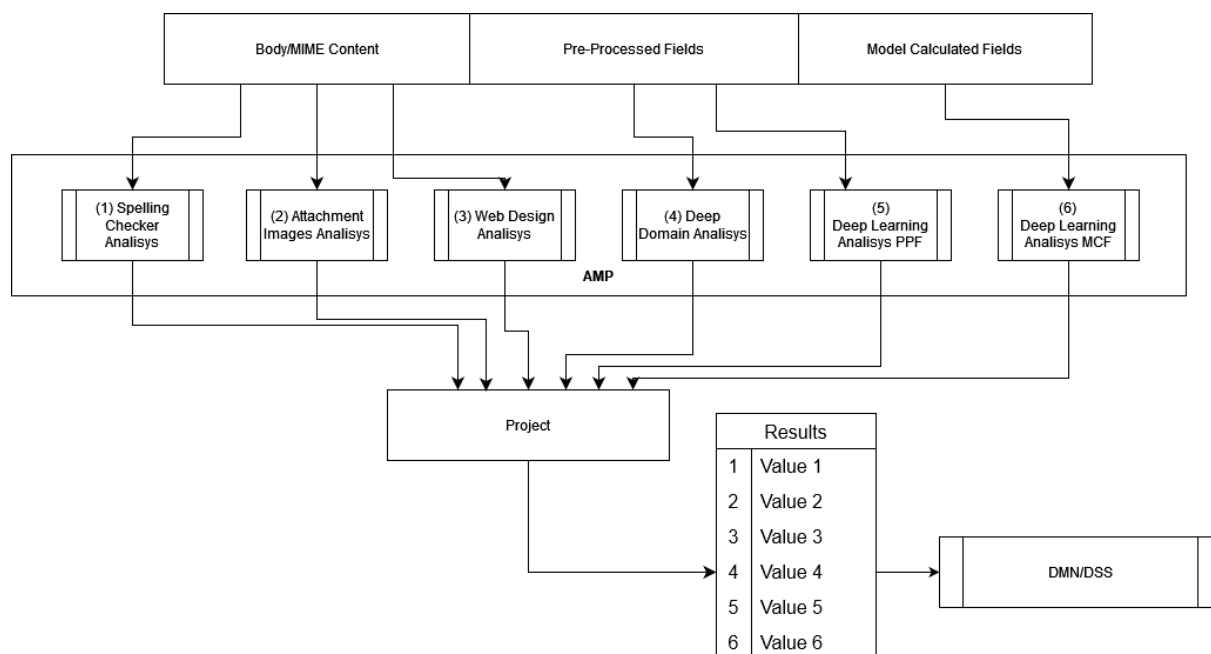


Figure 3. AMP flowchart.

**C. Cognition.** In this last and definitive phase, the results obtained from the previous phase are initially submitted for decision based on a decision support system based on Decision Model Notation (DMN) to obtain the first answer. This response is the system's decision about usability in terms of the analysis carried out in previous phases, and it's a value indicating the probability that a given email is analysed, whether phishing or not. Depending on the percentage obtained is valued next in the sub-phase of decision for the possible values of the same, reaching even to request human intervention in the process in what is called HITL (Human in The Loop) (Wu, 2022). The decision end should provide feedback to the system for its learning and use in future cases.

**III The Phishing Problem. Cognitive approach.** Our proposal focuses on analysing e-mails with spear phishing and elaborated phishing threats since they are emails that have managed to pass security controls corporate and can be identified by a human. This is the "gap"

we want to cover, and for this reason, we propose our cognitive system. The analysed data and parameters we propose in each phase are, as shown in Figure 4:

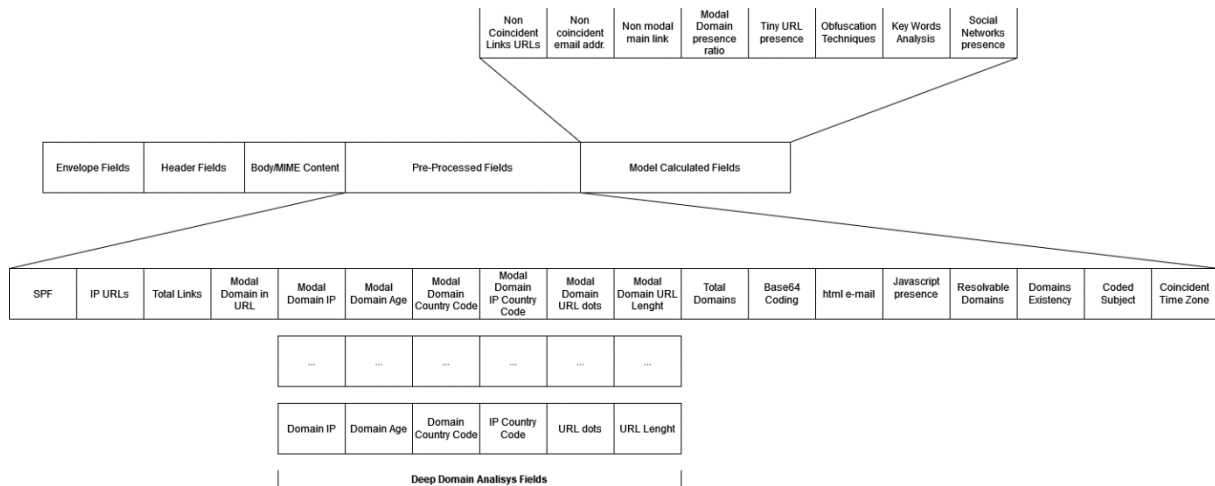


Figure 4. Proposed dataset.

**A. Perception.** Based on MIME and S/MIME specs, we gather all possible information from envelopes, headers, and body fields of incoming e-mails. And send the following to be pre-processed: SPF, IP type URLs, Domains IP, Domain’s age, number of links, number of domains and sub-domains, URL dots, URL length, Modal domain country, domains/IP countries, Base64 encoding, html e-mails, resolvable domains to real IP, Javascript in the e-mail, UTF-8 coded e-mail subject and coincident time zone.

**B. Understanding.** This second phase, as we have seen previously, is where a data set is collected and prepared. It’s subdivided into two sub-phases, “Model” and “Process”. First, let’s add the following data to the model of data by calculating certain parameters, all aimed at providing a data set as accurately as possible to the engines of the following sub-stage. We propose to obtain, calculate, and incorporate into the set the following parameters: non-

coincident URL Links, non-coincident e-mail addresses, non-modal main link, modal domain presence ratio, Tiny URL presence, obfuscation techniques, keywords analysis, social networks Links presence.

**C. Cognition.** In the last phase, the cognitive process culminates. For this reason, we have a first sub-phase, "Project", where we are going to project all the results obtained in the previous sub-phase (Process) and will generate, through a decision-making system based on in DMN models as proposed by (Valencia-Parra, 2021), an output with the result of the entire process, expressed as a percentage of probability as a function of the weights assigned to the different decision model tables. DMN (Decision Model and Notation) is an OMG standard based on a description statement (“if-then”) of a decision model. Those weights are shown in Figure 5.

INPUTS DMN/DSS						
SCA	AIA	WDA	DDA	DLA PPF	DLA MCF	OUTPUT
			Result > 70	Result > 70	Result > 70	<b>Malicious/Phishing</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	Result > 70	Result > 70	Result > 70	<b>Malicious/Phishing</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	30 < Result < 70	30 < Result < 70	Result > 70	<b>Malicious/Phishing</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	30 < Result < 70	Result > 70	30 < Result < 70	<b>Malicious/Phishing</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	Result > 70	30 < Result < 70	30 < Result < 70	<b>Malicious/Phishing</b>
			30 < Result < 70	30 < Result < 70	30 < Result < 70	<b>Doubtful</b>
2 < Misspelling < 5	5 < Reference Img < 10	5 < Comments < 15 100 < Html Tags < 250	30 < Result < 70	30 < Result < 70	Result < 30	<b>Doubtful</b>
2 < Misspelling < 5	5 < Reference Img < 10	5 < Comments < 15 100 < Html Tags < 250	30 < Result < 70	Result < 30	30 < Result < 70	<b>Doubtful</b>
2 < Misspelling < 5	5 < Reference Img < 10	5 < Comments < 15 100 < Html Tags < 250	Result < 30	30 < Result < 70	30 < Result < 70	<b>Doubtful</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	Result < 30	Result < 30	30 < Result < 70	<b>Doubtful</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	Result < 30	30 < Result < 70	Result < 30	<b>Doubtful</b>
Misspelling > 5	Reference Img < 5	Comments < 5 Html Tags < 100	30 < Result < 70	Result < 30	Result < 30	<b>Doubtful</b>
			Result < 30	Result < 30	Result < 30	<b>Legitimate</b>
Misspelling < 2	Reference Img > 10	Comment > 15 Html Tags > 250	Result < 30	Result < 30	30 < Result < 70	<b>Legitimate</b>
Misspelling < 2	Reference Img > 10	Comment > 15 Html Tags > 250	Result < 30	30 < Result < 70	Result < 30	<b>Legitimate</b>
Misspelling < 2	Reference Img > 10	Comment > 15 Html Tags > 250	30 < Result < 70	Result < 30	Result < 30	<b>Legitimate</b>

Figure 5. DSS/DMN Table.

**IV Conclusions and future work.** Cognitive systems should play an important role in the future of security analysis, and therefore the so-called cognitive cybersecurity will establish a new paradigm. They will also provide support and help in making decisions for security analysts when detecting, assessing, and dealing with incidents that risk the assets of a certain company or organisation. Even though we are in an early stage of our work, we estimate that the proposal made can substantially improve the classification of incoming mail from a given organisation as it provides a system of behaviour like the human at the time scanning emails for phishing and other threats. The next step (in which we are working now) is to obtain a dataset that is as complete as possible to carry out an exhaustive study of the variables analysed and measure their suitability and relevance, as well as the applied model. It will also tune the parameters of the tables of the decision model used in the last phase. Likewise, incorporating other aspects to improve the model, such as the analysis of visual or mail quality, the use of NLP for analysis of urgent language in the body or subject of the email and the analysis of the visual resemblance of domains, among others.

## **V References.**

- V. V. Raghavan, V. N. Gudivada, V. Govindaraju, and C. R. Rao, *Cognitive computing: Theory and applications*. ISBN: 9780444637444. Elsevier, 2016.
- A. D. W. Sumari and A. S. Ahmad, “The application of cognitive artificial intelligence within c4isr framework for national resilience” in 2017 Fourth Asian Conference on Defence Technology-Japan (ACDT). IEEE, 2017, pp. 1–8.
- X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, and L. He, “A survey of human-in-the-loop for machine learning” *Future Generation Computer Systems*, 2022.
- A. Valencia-Parra, L. Parody, A. J. Varela-Vaca, I. Caballero, and M. T. Gómez-López, “Dmn4dq: when data quality meets DMN,” *Decision Support Systems*, vol. 141, p. 113450, 2021.

## Human-Computer Interaction in the Age of GenAI: Implications of Simulacra for Service Business

Tuomi, A.<sup>1\*</sup>, Ascenção, M.P.<sup>2</sup>

*\*Corresponding author - presenter*

<sup>1</sup> [aarni.tuomi@haaga-helia.fi](mailto:aarni.tuomi@haaga-helia.fi); Haaga-Helia University of Applied Sciences, Finland

<sup>2</sup> [mariopassos.ascencao@haaga-helia.fi](mailto:mariopassos.ascencao@haaga-helia.fi); Haaga-Helia University of Applied Sciences, Finland

**Keywords:** artificial intelligence, generative AI, service sector, simulacra, human-computer interaction

### Background of the study

We are living in a time of rapid technological change, whereby new approaches to and applications of information communication technology (ICT) are transforming the way traditional business are operated. Most recently there has been a renaissance of sorts in one technology in particular, artificial intelligence (AI), whereby emerging AI tools are able to generate seemingly realistic-looking content, from images to virtual 3D environments. (Davenport & Mittal, 2022). The increase in technological capability, most notably generative AI (GenAI) to automatically generate high-quality content sees the traditional business management disciplines, including services marketing and management, enter what some scholars are calling the ‘era of falsity’ (Plangger & Campbell, 2022) and ‘post-truth’ (Keyes 2004).

First, it should be noted that the concept of ‘falsification’ is not new. As soon as new digital tools have emerged, businesses have tried to make the most of them (Bower, 1998).

Consumers have also been surprisingly quick to adapt to changes in technology, whereby for example when Photoshop – one of the most prominent photo-editing software stacks globally – was released in 1990, there was a real fear that it would lead to grave misuse and large amounts of false content, ultimately eroding people’s trust in visual images altogether (Leetaru, 2019). That did not happen, of course, and despite these fears, the concept of ‘photoshopping’ images and of an image being ‘photoshopped’ soon entered the common

lingo. However, at the same time it should be noted that the increase in visual literacy does not automatically mean a decrease in the effectiveness of manipulated images to influence behaviour, whereby for example in their experimental study Lazard et al. (2020) demonstrated that even when participants knew an advertisement photo to be manipulated, they preferred it regardless over a non-manipulated photo.

French philosopher Baudrillard (1994) referred to such 'falsified' content as 'simulacra'. For example, in the context of hospitality and tourism, much of the marketing material put out by brands is digitally manipulated, and the effectiveness of using e.g. F&B related images to boost sales has been known for long (Hou et al., 2017). Novel technologies capable of generating high-quality content represent a continuation to this tendency.

However, what academics are noting is that the degree of falsification is rapidly increasing, whereby content (e.g. code, text, images, video) generated by technology has increasingly started to pass as content generated by human software engineers, authors, photographers or animators. For example, Tuomi (2021) used AI to generate human-like restaurant reviews, demonstrating that several AI-generated reviews passed as human-written reviews when their authenticity was evaluated by human judges.

## **Aim of the study**

To that end, this study aims to conceptually explore the types of human-technology interactions that emerge as a result of novel human-technology interaction, specifically the use of novel GenAI tools, paying particular attention to the context of service business. Specifically, the paper aims to address the following research question: What are the key implications (costs and benefits) of adopting generative AI applications for different service business stakeholders (businesses, workers, education)?

## **Methods**

Adopting an integrative literature review approach, the study conceptually maps out parallels between advances across three distinct fields of scholarship: service business, AI-human teams, and the concept of authenticity. Particularly the study draws on Baudrillard's (1994) notion of simulacra.

In his seminal work, Baudrillard proposed that the postmodern society has replaced reality and meaning with symbols and signs, and that, in most situations, human experience is a simulation of reality. More specifically, Baudrillard argues there to be four layers of simulacra: first order simulacra, which are faithful copies of the original, and where the “judge” of authenticity can still discern the reality from the imitation; second order simulacra, which represent a perversion of reality and where the judge of authenticity believes the copy to be an improved, more desirable version of the reality; third order simulacra, which illustrates a pretence of reality where there is no clear distinction between reality and its representation, and in consequence, the copy replaces and improves upon the reality; and fourth order simulacra: pure simulation, whereby the connection to reality is fully severed. These are signs that simulate something that never existed in a form of hyperreality in which the copy has become truth.

According to Bukatman (1993, 2) “it has become increasingly difficult to separate the human from the technology”. Following Baudrillard, the study will conceptualise emerging forms of human-computer interaction (HCI) in the age of GenAI across the four different layers of simulacra. Doing so will significantly contribute to the current theoretical understanding of HCI, as well as provide service business practitioners, particularly strategic and operational decision-makers, insight into how best integrate GenAI into organisational strategy, service processes, and workflows (i.e. business design, service design, job design).

## **Results/Findings**

The results of this conceptual study will help service business stakeholders to better understand the opportunities and challenges – that is, the costs and benefits – of adopting novel GenAI tools as part of day-to-day operations. To illustrate, practical examples of emerging digital content creation tools has recently skyrocketed. Codex automatically generates lines of code, DALL-E 2, Midjourney generate high-resolution images, ElevenLabs generates voice from existing samples, GPT-4 generates short- or long-form text, Autodraw is the “autocorrect” for digital drawing, Fontjoy generates font pairings, Namelix generates business names and logos, and Make-A-Video and D-ID generate animated videos from text prompts. In addition to new tools, generative AI capabilities are also increasingly being added

to existing software. For example, Photoshop recently launched a feature called Generative Fill to its photo editing suite of technologies to enable new types of photo editing capabilities.

Novel digital content creation tools bring new affordances to stakeholders involved in value creation in service business, from workers to businesses and the educational sector.

For the worker, emerging technology offers new tools, enabling new forms of creativity to emerge in human-technology collaborative teams, e.g. in marketing or sales. The automation of more tasks changes workflows, whereby the importance of assessing when and how the human should be kept in-the-loop becomes imperative (Ivanov, 2021). New tools also bring the need for training, and the overall shift to more automated work processes poses questions around upskilling and re-skilling of staff (Tuomi et al., 2020). For service businesses, novel forms of human-technology interaction offer more ease of access, more control and the possibility to do more in-house. However, technology adoption becomes the likely bottleneck, whereby businesses will have to consider the costs and benefits of different emerging tools. Which tools to use, which to ignore? How to use them the most effectively? Finally, for the educational sector, the indistinguishability of digitally generated content from human-generated material presents a double-edged sword. On one hand, new tools mean an abundance of higher quality content, but on the other, an authenticity crisis might follow. How to assess content created through human-technology collaboration? How to deal with plagiarism and intellectual property in the context of AI-generated content?

## Conclusion

Overall, from a sociotechnical perspective, research on human-computer interaction (HCI) and collaboration have witnessed significant developments and prominent theories over the last decades. Recently, a key development has been the increasing integration of AI and machine learning into HCI systems, most notably generative AI, enabling e.g. more adaptive and personalised user experiences. This has led to the emergence of new types of human-computer interactions, prompting a need to re-conceptualise HCI.

Besides Baudrillard's (1994) simulacra, previous theories like Activity Theory and Distributed Cognition have played crucial roles in understanding collaboration in HCI.

Activity Theory emphasises the social and contextual aspects of human-computer interaction,



highlighting the interplay between users, tools, and the environment (Bonnie, 1995). Distributed Cognition, on the other hand, focuses on how cognitive processes are distributed across individuals, artefacts, and the environment, emphasising the importance of collaboration and shared understanding in achieving complex tasks (Hollan et al., 2000). Additionally, theories such as Social Presence Theory have contributed to understanding communication and collaboration in computer-mediated environments. Social Presence Theory explores the degree to which technology can convey a sense of being present with others (Gunawardena, 1995).

Noting the development of previous theory on human-computer interaction, and drawing further inspiration from simulacrum, we argue that increasingly, Baudrillard's (1994) notions of simulacra meet the everyday reality in human-technology interaction, prompted recently through the increase in capability of and access to powerful new GenAI systems. As consumption moves further to artificial and digital environments such as the metaverse, many consumers, workers, students, and managers will experience ever-increasing versions of the simulacra instead of 'reality'. Conceptually and empirically exploring emerging forms of human-technology interaction in the context of service business and generative AI is thus important and timely. Future research should aim to adopt varied ontological, epistemological and methodological approaches. Particularly useful might be studies that effectively bridge the gap between theory and practice and demonstrate measurable effects in real-world environments, for example field experiments, multiple case studies or action research.

## References

- Baudrillard, J. (1994). *Simulacra and simulation*. Ann Arbor: University of Michigan Press.
- Bonnie, N. (1995). *Context and consciousness: activity theory and human-computer interaction*. Cambridge, MA: MIT Press.
- Brower, K. (1998). Photography in the age of falsification. *The Atlantic*, 5/1998.
- Bukatman, S. (1993). *Terminal identity. The virtual subject in postmodern science fiction*. London: Duke University Press.
- Davenport, T., Mittal, N. (2022). How generative AI is changing creative work. Available at: <https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work> (Accessed 6th July 2023).

- Gunawardena, C. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications* 1(2/3), pp. 147-166.
- Hollan, J., Hutchins, E., Kirsh, E. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction* 7(2), pp. 174-196.
- Hou, Y., Yang, W., Sun, Y. (2017). Do pictures help? The effects of pictures and food names on menu evaluations. *International Journal of Hospitality Management* 60, pp. 94-103.
- Ivanov, S. (2021). Robonomics: The rise of the automated economy. *ROBONOMICS: The Journal of the Automated Economy* 1.
- Keyes, R. (2004). *The Post-Truth Era: Dishonesty and Deception in Contemporary Life*. New York: St. Martin's Press.
- Lazard, A., Bock, M., Mackert, M. (2020). Impact of photo manipulation and visual literacy on consumers' responses to persuasive communication. *Journal of Visual Literacy* 39(2), pp. 90-110.
- Leetaru, K. (2019). Deep fakes are merely today's photoshopped scientific images. Available at: <https://www.forbes.com/sites/kalevleetaru/2019/08/24/deep-fakes-are-merely-todays-photoshopped-scientific-images/> (Accessed June 21<sup>st</sup> 2023).
- Plangger, K., Campbell, C. (2022). Managing in an era of falsity: Falsity from the metaverse to fake news to fake endorsement to synthetic influence to false agendas. *Business Horizons*, 65(6), pp. 713-717.
- Tuomi, A., Tussyadiah, I., Ling, E.C., Miller, G., Lee, G. (2020).  $x=(tourism\_work)$   $y=(sdg8)$  while  $y=true$ : automate (x). *Annals of Tourism Research* 84, 102978.
- Tuomi, A. (2021). Deepfake consumer reviews in tourism: Preliminary findings. *Annals of Tourism Research Empirical Insights*, 100027.

## **INFORMATION SOCIETY AND SOCIO-ECONOMIC CHARACTERISTIC IN EUROPE: A TYPOLOGY OF REGIONS**

Purificación Crespo-Rincón<sup>1</sup>; Rosa Jordá-Borrell<sup>2</sup>; Francisca Ruiz-Rodríguez<sup>2</sup>

<sup>1</sup>[purcrerin@alum.us.es](mailto:purcrerin@alum.us.es)

<sup>2</sup>Universidad de Sevilla, España.

### **KEYWORDS**

Information and communication technologies (ICTs), households and individuals (H+i); socioeconomic and technological factors, NUTS 2.

### **INTRODUCTION**

The digital development of a region is considered as the process that drives the transformation of the region towards the information and knowledge society, and makes it possible to modify conditions through the generation and processing of information and knowledge in order to improve competitiveness, innovation and the adoption of information and communication technologies (ICTs) (González-Relaño, R. et al., 2021; Hernández & Maudos, 2021).

ICT access and usage has given rise to the so-called digital society. In this society, the public, companies and institutions establish their relationships (social, administrative, employment, consumption, etc.), while also generating data and information via digital devices and platforms, giving rise to digital development (Ruiz-Rodríguez et al., 2018). Therefore, digital technologies are a constant component of the lives of persons and companies, with these being dependent on digital technologies and their specific infrastructures.

At present, empirical research on the information and communication society at the NUTS 2<sup>1</sup> regional level in the European Union (EU) is insufficient, especially the analysis of its relationship with the socioeconomic environment. Moreover, when this topic has been addressed, studies have focused above all on the factors that determine basic digital development (Cruz-Jesús et al., 2016). Identification of this research gap has led to the proposition of the following hypotheses, with their corresponding research questions:

1. The differences between regions in terms of Internet access for the European population are small, as EU and national government policies on communications equipment and infrastructure have helped to significantly reduce these potential inequalities. This raises the question: Are regions with good telecommunications infrastructure coverage, such as the availability of broadband, those that most use advanced ICTs?
2. The distance between regions, in terms of the use of social networks by the population, is small, while the dissimilarities between regions in the use of advanced ICTs are important. Are these latter differences associated with the socio-economic characteristics of the regions? Which socio-economic factors or elements are more important? Is the innovation and business capacity of a region associated with the advanced use of ICTs by the population of that region?

In this context, the objective of this study is, firstly, to identify the socio-economic (related to training, enterprise and regional wealth) and technological (innovation, digitalisation, etc.) factors that may define the underlying structure of the digital development of households and individuals; and secondly, to characterise the socio-economic environments that are associated

---

<sup>1</sup> Nomenclature of territorial units for statistics, a hierarchical system for dividing up the economic territory of the EU, and the spatial system of the European Union

with advanced ICT usage by households and individuals (ICT\_H+i) at the NUTS 2 regional level in Europe.

## **THEORETICAL FRAMEWORK**

In recent decades, the widespread diffusion of ICTs has led to a transformation of the world into an information society. Thanks to ICT infrastructure and equipment, individuals and governments now have much better access to information and knowledge than ever before in terms of scale, scope and speed.

The rapid advancement of ICTs worldwide is arousing the interest of many researchers, motivating them to focus their research on the impact of ICT diffusion on socioeconomic growth. According to Stanley et al. (2018), the socioeconomic characteristics of a region may or may not favour advanced ICT access and usage in that region. In this context, prominent modern theories such as neo-Schumpeterian theories (Schumpeter & Nichol, 1934; Malacarne, 2018; Garbin & Marini, 2021) and neoclassical growth theory have pointed out the positive relationship between the socioeconomic characteristics of a territory/region and the development of technology, including ICTs.

Consequently, these theories suggest that ICTs are a key input for a region to improve the production process, modernise technology and improve the skills of the workforce. They also mean that competitiveness and complex technology are essential elements of the regional business fabric (Jordá-Borrell, 2021).

## **METHODOLOGY**

In order to meet the objectives and verify the hypothesis, the methodology used is as follows:

### **SCOPE OF THE STUDY.**

A population of 333 NUTS 2 regions in Europe was used, including the NUTS 2 regions of the 27 EU countries, the 26 regions of Turkey (country in EU accession negotiations), the 41 regions of the United Kingdom (country belonging to the EU at the time of the study), the 7 regions of Norway (country belonging to the European Economic Area), and the 8 regions of Switzerland (member of the European Free Trade Association-EFTA).

### **VARIABLES AND DATA.**

Fifteen variables have been used that show a relationship between the socioeconomic characteristics of a territory (level of wealth, level of education, innovative capacity and entrepreneurial dynamics) and the access and use of ICT\_H+i. The variables and data for these variables have been obtained from Eurostat regional statistics for 2019.

### **METHODS OF ANALYSIS.**

The following have been applied; a) Factor Analysis (FA) to identify the socio-economic and technological factors that define the underlying structure of digital development of households-individuals at the NUTS 2 level; and b) Cluster Analysis (CA) to characterise the socio-economic environments associated with advanced ICT usage.

## RESULTS

### FACTOR ANALYSIS

Four factors were obtained with the FA (Table 1) with a KMO of 0.858, determinant of 0.00003601 and a total explained variance of 70.036 %:

Factor 1. This encompasses all the variables of Internet access and usage, associated with GDP/pc, employment in high-tech sectors and population with higher education (bachelor's degree, master's, doctorate). This factor could be called ICT\_H+i and GDP/pc.

Factor 2. This comprises the positive association between the variables of patent generation and R&D expenditure in all sectors, and could be called R&D and Patents.

Factor 3. This combines the population with secondary and tertiary education (medium-high), households with broadband and the unemployment rate. When the unemployment rate decreases, the number of households with broadband and the population with a medium-high level of education increases. This factor could be designated human capital and ICT infrastructure.

Factor 4. This associates the growth rate of telecommunications employment and net growth in the number of companies (business fabric), and is not directly associated with the use of ICTs by individuals. This factor could be called business growth and telecommunications employment.

Rotated component matrix				
	Component			
	1	2	3	4
Individuals who used the Internet for online banking	0.924			
Individuals who access the Internet on a daily basis	0.881			
Individuals who used the Internet to interact with public authorities	0.857			
Population aged 25-64 with tertiary education	0.840			
Households with Internet access	0.790			
Individuals who used the Internet for the sale of goods and services	0.759			
GDP per capita compared to the EU28 average in PPP	0.738			
Employment in High Technology Sectors	0.572			
Patent applications as a share of GDP		0.853		
Intramural spending on R&D		0.782		
Working population aged 15-64 with secondary and tertiary education			0.769	
Households with broadband			0.702	
Unemployment rate in the population aged 15-74 years old.			-0.610	
Net Growth of the Number of Companies				0.795
Growth rate of Employment in Telecommunications				0.790

Extraction method: principal component analysis.  
 Rotation method: Varimax with Kaiser normalisation  
 a. The rotation has converged in 5 iterations

Table 1: Factor Analysis

## CLUSTER ANALYSIS

A cluster analysis was carried out with the factors resulting from the FA, identifying 5 homogeneous groups of European regions, which are arranged in order from the lowest to the highest intensity of access and advanced use of ICT\_H+i (Map 1).

**CLUSTER 1.** This includes eastern European regions, as well as Turkey, northern and central Italy, Portugal and the Spanish region of Castile-La Mancha (32.52% of regions). This cluster is defined by a below average Factor 1 (ICT\_H+i and GDP/pc), a lower level of socio-economic development than the rest, a higher than average unemployment rate, a low telecommunications employment growth rate and an advanced use of ICTs (online sales and interaction with the



public authority) well below the average, although these regions have an extensive access infrastructure.

CLUSTER 2. This group is defined by Factor 3. Here we differentiate two subgroups:

Cluster 2.1 (5.75% of regions). This is composed of southern European regions of Greece, Italy, Andalusia and Extremadura, regions with a low socio-economic level: GDP/pc below the European average and a high unemployment rate. 99% of households have Internet access and human resources with tertiary education, but the use of advanced ICTs is deficient.

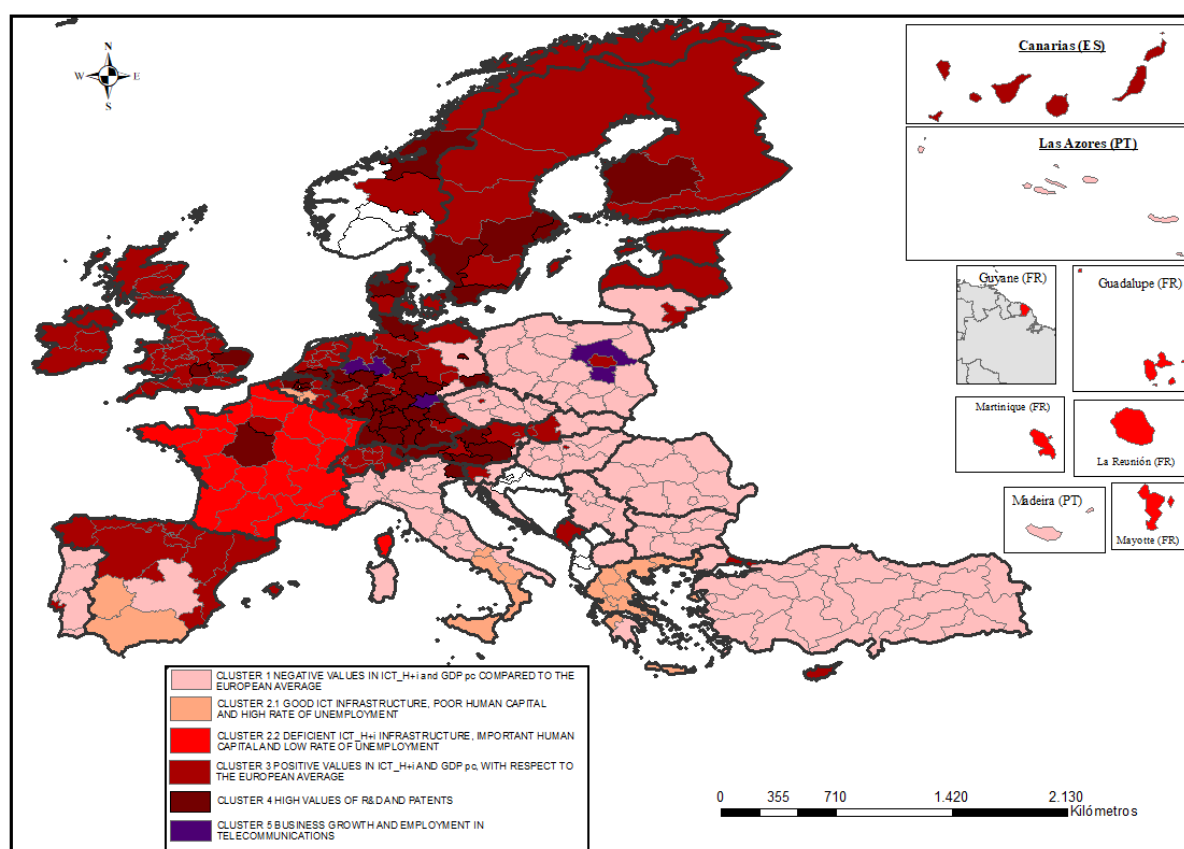
Cluster 2.2 (14.23% of regions). This corresponds to the French regions, except those in the centre, characterised by a GDP/pc close to the European average, and a working population with a medium-high level of education, slightly below the European average. However, they have not developed digital skills because they do not have full coverage of ICT infrastructures.

CLUSTER 3. (38% of regions) Regions positively defined by Factor 1 (ICT\_ H+i and GDP/pc). It includes regions in Sweden, Finland, Estonia, Latvia, Denmark, northern regions of Germany, the Netherlands, Belgium, the United Kingdom, Ireland, north-eastern Spain (plus the Canary Islands), Austria, Switzerland, the Czech Republic and some capitals and regions of Eastern Europe. These regions have a slightly lower level of ICT access infrastructure than the European average, while GDP/pc, business growth, workforce with tertiary education and employment in high-tech sectors are all slightly above the European average.

CLUSTER 4. (13.98% of regions). This includes many of the regions of Germany, Austria, southern Finland and Sweden, the central core of France and the British region of East Anglia-Berkshire. This group is mainly delimited by Factor 2 (R&D and Patents). These are regions with high levels of ICT access and advanced ICT usage, with a very high GDP/pc compared to

the European average, a telecommunications employment growth rate and employment rate in high-tech sectors that are almost double the European average. All this is based on a high level of investment in R&D, which gives rise to a high number of patents.

CLUSTER 5. (1.22% of regions). This comprises the German regions of Oberfranken, Münster and Detmold and the Polish region of Mazowiecki regionalny, characterised by Factor 4. These are socio-economically advanced regions, with a GDP/pc and an employment growth rate in the telecommunications sector above average, as well as a highly skilled workforce and a high degree of ICT access and advanced use.



Map 1. Typology of regions according to the interaction between ICT and socio-economic characteristics in Europe.

## DISCUSSION AND CONCLUSIONS

The analyses carried out and the results obtained have made it possible to confirm that good telecommunications coverage (fixed and mobile networks, broadband, etc.) does not guarantee advanced ICT use. EU national and regional policies have been aimed at deploying ICT access infrastructures to achieve a homogeneous and uniform regional map (Ruiz-Rodríguez et al., 2020). It has also been shown that a higher educational level of citizens increases the potential of regions to both generate and absorb advanced technology (Wang et al. 2020).

The FA corroborates the starting hypothesis, which is that European regions with a high socioeconomic level reinforce their investment in the advanced use of ICT\_H+i and this is what will mark part of the digital divide and regional heterogeneity in Europe (Jordá-Borrell & Lopez-Otero 2020).

The CA shows that European regions that have a higher level of GDP/pc also have a strong business dynamic and a significant volume of telecommunications employment, which predisposes these regions to invest more in R&D and patent development. The concentration of wealth, technological capabilities and knowledge in the central European area is evident, and it is these regions that are leading the digital revolution in Europe (the regions of Germany, Austria, southern Finland and Sweden, and the central core of France) (Pradhan & Arvin, 2020).

This research, therefore, is a clear contribution to the study of digital, technological and socioeconomic inequalities in European regions, and highlights the concentration of wealth, technological skills and knowledge in the central area of Europe and the need to increase the diffusion of these to peripheral areas with a lower socioeconomic level, which has not allowed them to reach an advanced use of ICT.

Nevertheless, this study has its limitations, given that the number of variables included in the analysis is conditioned by the availability of regional data provided by Eurostat at the NUTS 2 level. Therefore, further research is recommended to increase the number of study variables related to the use of advanced digital technologies. To this end, it would be desirable for Eurostat to publish more data on ICT\_H+i, and especially on companies at NUTS 2 level, as this is the main level used by the European Union for the enactment of new territorial-regional policies for development and reduction of inequalities.

**FUNDING:** This article is a result of the R+D+i project PID 2019-107993GB-I00 funded by MCIN/ AEI/10.13039/501100011033.

**Thesis by compendium of publications, in which we already have this article published in the** *Revista Cuadernos Geográficos (2023), entitled “INFORMATION SOCIETY AND SOCIO-ECONOMIC CHARACTERISTICS IN EUROPE. A TYPOLOGY OF REGIONS.* <http://dx.doi.org/10.30827/cuadgeo.v62i1.25334>, *Scopus Journal*

## REFERENCES

Cruz-Jesus, F., Vicente, M. R., Bacao, F., & Oliveira, T. (2016). The education-related digital divide: An analysis for the EU-28. *Computers in Human Behavior*, 56, 72–82. [doi:10.1016/j.chb.2015.11.027](https://doi.org/10.1016/j.chb.2015.11.027)

Garbin, M. H., & Marini, M. J. (2021). Análise do conjunto normativo aplicável à Ciência, Tecnologia e Inovação do município de Pato Branco/PR. *X Seminário Internacional sobre Desenvolvimento Regional*. Retrieved from <https://online.unisc.br/acadnet/anais/index.php/sidr/article/view/20989/1192613286>

González-Relaño, R., Lucendo-Monedero, Á. L., Ruiz-Rodríguez, F. (2021). Information and Communication Technologies of households and individuals, geographical proximity and regional competitiveness: distribution, clusters and spatial patterns of technological capacity in Europe. *Boletín de la Asociación de Geógrafos Españoles*, (90), 7. [doi: 10.21138/bage.3118](https://doi.org/10.21138/bage.3118)

Hernández, L., & Maudos, J. (2021). *Competencias digitales y colectivos en riesgo de exclusión en España*. COTEC. Retrieved from <https://cotec.es/proyecto/competencias-digitales/51a02688-a11f-4fee-b047-41288ea0e0ac>

Jordá-Borrell, R., & Lopez-Otero, J. (2020). Economic growth factors in developing countries: the role of ICT. *Boletín de la Asociación Española de Geografía*, (86). [doi: 10.21138/bage.2979](https://doi.org/10.21138/bage.2979)

Jordá-Borrell, R (2021). La digitalización y/o la transformación digital de la empresa en Andalucía. In Mateu Bellés, J.F. y Furió A. (Eds.), *A Vicenç M. Rosselló, geògraf, als seus 90 anys* (pp. 433-448) [ISBN 978-84-9133-428-6](https://doi.org/10.21138/bage.2979). Valencia, España: Universitat de Valencia.

Malacarne, M. A. (2018). *Uma análise do desempenho econômico internacional do setor de tecnologia de informação e comunicação (tic) no brasil (2000-2017)*. LUME, repositório digital. Retrieved from <http://hdl.handle.net/10183/189723>.

Pradhan, R. P., Mallik, G., & Bagchi, T. P. (2018). Information communication technology (ICT) infrastructure and economic growth: A causality evinced by cross-country panel data. *IIMB Management Review*, 30(1), 91-103. [doi: 10.1016/j.iimb.2018.01.001](https://doi.org/10.1016/j.iimb.2018.01.001)

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterisation of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. *Telecommunications Policy*, 42(3), 187-211. [doi: 10.1016/j.tele.2019.05.002](https://doi.org/10.1016/j.tele.2019.05.002)

Ruiz-Rodríguez, F., González-Relaño, R., & Lucendo-Monedero, Á. L. (2020). Comportamiento espacial del uso de las TIC en los hogares e individuos. Un análisis regional europeo. *Investigaciones Geográficas (Esp)*, 73, 57-74. [doi: 10.14198/INGEO2020.RRGRLM](https://doi.org/10.14198/INGEO2020.RRGRLM)

Stanley, T. D., Doucouliagos, H., & Steel, P. (2018). Does ICT generate economic growth? A meta-regression analysis. *Journal of Economic Surveys*, 32(3), 705-726. [doi: 10.1111/joes.12211](https://doi.org/10.1111/joes.12211)

Schumpeter, J. A., & Nichol, A. J. (1934). Robinson's economics of imperfect competition. *Journal of Political Economy*, 42(2), 249-259. [doi:10.1086/254595](https://doi.org/10.1086/254595)

Wang L., Luo, G.L., Sari, A., Shao, X. F. (2020). What nurtures fourth industrial revolution? An investigation of economic and social determinants of technological innovation in advanced economies. *Technological Forecasting & Social Change*, 161. [doi: 10.1016/j.techfore.2020.120305](https://doi.org/10.1016/j.techfore.2020.120305)

## **LEGAL ANALYSIS OF USE OF ROBOTS IN MEDICINE**

Author Marina Galvin

*\*Corresponding author - presenter*

<sup>1</sup> *e-mail address cndmarina@gmail.com*

<sup>2</sup> *Univerisity of Seville (PHD in law of the university of Seville)*

**KEY WORDS:** Medicine, Robots, Artificial Intelligence, Health and Quality of life.

In this paper were going to try to explain which are the effects and consecuencies of use of robots in medicine and its legal aspects, in other words, what would its effects the advantages in technology and artificial intelligence and its applications in the field of medicine and health and have improved the quality of life, curing diseases and performing diagnoses that until recently were not possible. The autonomy of the AI must be antropocentric and respect at all times tha autonomy and the fundamental rigths of the patients, having to reject at all times the decision making based on automatization processes because these decisions could empty the content the patients decisions and ethical judgment of the professionals raising problems at the same time of determination of responsability, in relation to legal regulation of the Biomedicine we can point that the critical argument exposes that the AI effects structurally to the personal autonomy where the consideration of good arises from intersubjetive recognition and provides the necessity material conditions so that individual choice may be possible recognizing structural inequalities what should legal regulation for an appropriate way that avoids the

algorithmics challenges that produces risks in AI that through normative stipulations affects to the people relational autonomy (NEDELSKY, J. 2013).

In another vein the determining factors that have driven AI are the Big Data and the Computational Technological Development that allows the data to be object of treatment through complex algorithmics models, in this legal analysis we can't forget the right to information and explanation to the patient, its reach we can find it in European Data Protection Regulation (article 22, pulled apart 1&4 )above all terms of profiling of patients or making automated decisions that can affect to the capacity of people at the time of access to good or services.

An another element to consider is the derived responsibility of the use of AI in the healthcare given that produce damages and injures derived from decisions-making by the intelligent systems but we must ask ourselves thisd question they could be charged AI systems of the result of the damage produced? as a reply we can say that the direct responsible is the professional who intervenes as a support of the artificial systems to finish tihis abstract we will mention of the AI and the health insurance, so that is the main challenge who is facing the insurance health field is the risk assessment, the competiton or customer support, the medical attention and general expenses produced by a legislation what is increasing in most countries.

Thereby the ethical issues in healthcare where in United States top legislative activity in various regulations about, for instance the Self Drive Act H.R. 3388 or the Future of Artificial Intelligence Act 2017, H.R. 4625 among others. We highlight the european regulatory framework in this matter about the civil responsibility in the European Resolution (2015/2103 (INL)) or the regulation 85/374/ECC about defect products and the cibersecurity in protection data and the public healthcare.



AIM OF THIS PAPER:

The aim of this research work is as I said in the abstract, we try to explain which are the effects and consequences of use of robots in medicine and its legal and ethic aspects, in other words, what would be its effects the advantages in technology and artificial intelligence and its applications in the field of medicine and health and that have improved the quality of life, curing diseases and performing diagnoses that until recently were not possible. From an innovative and original perspective as humanity has put its work and intelligence to fight against the diseases this inevitably characterizes the past, the present and the future of the humankind, in this sense the advances of Biomedicine and the Biotechnology open horizons to the treatment of severe diseases even increase life expectancy, however these advances have sparked a political debate, ethic, scientific, philosophical and legal about the positive and negative implications of these that have been very notable in Spain where they have been produced a few legal reforms they have raised antagonistic issues as such as the stem cell research, legalization from supernumerary embryos of in vitro fertilization processes, the authorization of genetic diagnosis of therapeutic cloning through the Biomedical Investigation Law 14/2007 July, 3.

Nevertheless and despite the positive aspects of technology in Biomedicine, the technological imperative should never be prioritized as an alternative to the humanism defined by (GRACIANO G, 2004) as an alternative and overcome the ancient ethic imperative who wonders what is correct or not, but the technically feasible and viable does not have to be morally legitimate and legally licit unlike pure academic science or instrumental that only seeks to enrich with their knowledge being ethically neutral, but the union between the technological innovation and the scientific investigation result in technoscience that has transformed scientific knowledge to

achieve other ends that are sometimes silenced as well as control the activity of the scientists by society that draw the lines should never be traded and the scientists themselves have to control through the ethics responsibility as a result among others the possible dangers generated by the Biotechnologies advances, should be observed at all times the principles of caution and responsibility, the principle of responsibility has been treated in the play titled "Das prinzip Verantwortung " by JONAS H. 1965, Interconnected with this principle we can find the the caution principle that applies in the face of possible scientific uncertainties in connection with the severe damages that could give rise take measures to prevent the risks articles 22 & 2 form Biomedical Investigation Law (BIL), the positive impact of the technoscience it depends on the destiny that is given (Value Neutrality Theory, defined by (OLIVER, L. 2021) Being therefore impossible to guarantee the full security. Also the autonomy of the AI must be anthropocentric and respect all the times the autonomy and the fundamental rights of patients, having to reject at all times the decision making based on automatization processes because these decisions could empty the content the patients decisions and the ethical judgment of the professionals raising problems at the time of determination of the responsibility in relation to legal regulation of Biomedicine we can point that the critical argument exposes that the AI effects structurally to the personal autonomy and the individual relationships conditioning the relational autonomy where the consideration of good arises from intersubjective recognition and provides the necessary material conditions, so that individual choice may be possible recognizing structural inequalities what should legal regulation look for, in an appropriate way that avoids the algorithmics challenges that produces risks in AI that through normative stipulations affects to the people relational autonomy, that was dealt by NEDELSKY, J. 2013.

**METHODOLOGY:**

To write this paper we have used the scientific-logic method, framing a process to collect data and other sources of information in a methodological and descriptive context, from internet and books on matter.

**RESULT OF RESEARCH:**

the result of this research is to give a conceptual and scientific -legal idea of use of robots in medicine, keep in mind the problems it can cause their use in public healthcare and its ethical repercussions from the point of view of the patients and the protection of their personal data in the performance of this activity legislation provides the figure of the person responsible for data processing and his liability towards third parties , in the misues of patients data, definetly to give a deeper insight into regulations within the framework of civil liability regime in this matter.

**CONCLUSION:**

The limitations and managerial, consist in how have you tried is to give positive vision about the robotics and AI in medicine and other aspects of life because we can not deny that the robotics and AI have made progress in all walks of life although the negative thoughts of humankind about this technology, we can't forget either the need to regulate this technology before the possible risks and illegitimate meddling in our lifes through our electronic medical

records, or devices and applications that knows more about us than ourselves (heart rate, blood sugar index and other data from us).

Therefore the regulations is needed through laws that adapt to technological changes at the same time create an ethical framework to help control out the scientifics, engineers and manufacturers activities and their responsibility against possible damage that could cause the use of robots in all fileds of medicine and we must understand that the law does not advance at the same pace than technology turning to law in sometihing inappropriate to regulate the robotics and IA and their different uses.

It was planted a logical-scientific-structural path about the use of robots in medicine and its legal, ethical, social and economic repercussions at European, national, and International level as well as the effects of use of AI in medicine and the consecuencies have been worked from the point of view of risk theory and liability regime in an attempt to fill the legal looploses existing in this matter so controversial and difficult to regulate, we also dealt with the repercussions of the cibersecurity configured as an essential element in digital infrastructures clinic, being an elusive goal in a near future.

#### LITERATURE:

BORH, A. "The Artificial Intelligence in the ambit of the Healht", published by Elsevier, 2021.

GRACIANO, G. "Bioethical logbook", published by The Official Magazine of Spanish Bioethics and Ethic Medical Association, 2004.

JONAS, H "Das Prinzip Verantwortung", published by Suhrkamp Verlag, 1965.

MEMARZADEH, K. "The Artificial Intelligence in the ambit of the health"published by Elsevier, 2021.

NEDELSKY, J. "Laws relations: A relational Theory of Self Autonomy and Law", published by OUP US, 2013.

OLIVÉ L. "Are Ethically Neutral the science and technology?: Value Neutral Theory, 2021.

## Mining user sentiment on healthcare during COVID-19

Aunimo, L.<sup>1\*</sup>; Oprescu, A. M.<sup>2</sup>; Muñoz-Saavedra, L.<sup>3</sup>

*\*Corresponding author - presenter*

<sup>1</sup> *lili.aunimo@haaga-helia.fi; Haaga-Helia University of Applied Sciences, Finland*

<sup>2</sup> *aoprescu@us.es; Department of Electrical Technology, Universidad de Sevilla, Spain*

<sup>3</sup> *lmsaavedra@us.es; Architecture and Computer Technology, Universidad de Sevilla, Spain*

**Keywords:** sentiment analysis, user generated data, Twitter data mining, COVID-19, healthcare

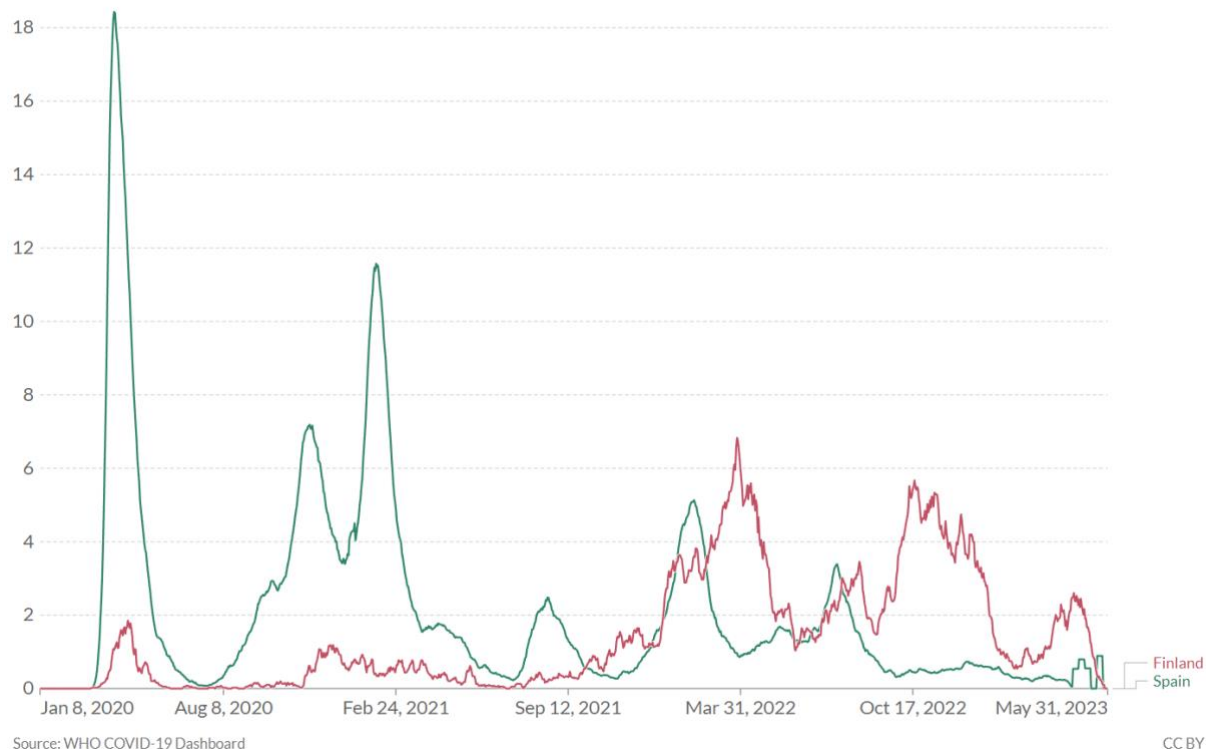
### Introduction

The COVID-19 pandemic has had a significant impact all over the world. Different regions and countries have experienced different impacts and outcomes during the COVID-19 pandemic. Specifically, Finland and Spain have had contrasting experiences about pandemic outcomes. They have adopted different pandemic response strategies and different mortality rates (see Figure 1) have been experienced in these countries. The countries have distinct population sizes, social life, culture, demographics, and healthcare systems. Healthcare systems in both countries are universal, however different in structure and capacity. User views on the healthcare system are of a high relevance to monitor the quality of the services and the satisfaction experienced. COVID-19 was an unexpected incident that affected the healthcare system vastly. It is important to study how it affected the public opinion on it and if there are differences in the public opinion in Spain and in Finland. Additionally, it is important to study how well sentiment mining techniques applied on user generated data in the social media can be used to monitor public opinion on a topic. Being able to use automated methods on freely available and already existing data may provide cost savings when compared with traditional methods such as patient surveys and interviews. Data mining -based methods may also provide additional information that can not be captured by the traditional methods.

The research questions are as follows:

- Did COVID-19 affect the public sentiment of tweets related to healthcare services? If so, how can we detect and quantify these changes using data mining and sentiment analysis techniques?
- Is there a difference on the views among citizens in Spain or Finland? If so, can data mining and sentiment analysis techniques be used to detect and quantify these differences?
- Is there variation in sentiment during the different phases and events related to the COVID-19 pandemics? If so, can we detect it based on tweets?

There has been prior work on mining public opinion on digital applications for health services from social media (Pai & Alathur, 2018) or vaccination (Tavoschi et al., 2020). To the best of our knowledge, there is no prior research on mining social media for public opinion on healthcare services in Finland or Spain.



**Figure 1: Confirmed deaths from COVID-19 per million people in Finland (red) and Spain (green) from 8 January 2020 to 31 May 2023. Credit: (Our World in Data, 2023)**

## Related Work

There is some previous work related to sentiment and topic mining from user-generated content in healthcare (Zunic et al., 2020). For sentiment analysis, the topic of healthcare services emerged as one of the most popular and widely discussed topics among online communities. Before the COVID-19 pandemic, researchers in (Byrd et al., 2016) explored Twitter data for real-time detection and surveillance of influenza. A longitudinal study aimed to characterize patient sentiments across the United States over a four-year period (Sewalk et al., 2018). . Regarding COVID-19 specifically, (Tsai & Wang, 2021) studied people's attitudes toward public health policies and events through Twitter opinion mining. To gain a deeper understanding of public sentiment, the collected tweets were meticulously annotated across a spectrum of sentiments, ranging from very negative and negative, through neutral, to positive and very



positive. The researchers found that staying at home and social distance policies caused a strong negative impact among the public. (Ainley et al., 2021) used Twitter data to gain insight into people's perceptions of healthcare during the pandemic in the United Kingdom. Thematic and sentiment analysis was used to parse Twitter data. Public opinion on telemedicine during the pandemic was studied in (Pool et al., 2022) and (Pollack et al., 2021). This type of care was found to be positively received by most of the public.

Sentiment analysis from social media data has been done using rule-based methods, such as VADER (Hutto & Gilbert, 2014) or transformer-based language models such as (Quoc Nguyen et al., 2020). Cliche (2017) present work where they use deep learning methods (CNNs and LSTMs) to create a sentiment analyser for Twitter data in English. There is prior work on sentiment analysis from Twitter data for Finnish. This work is based on lexicon-based approaches using Finnish lexicons (Vankka et al., 2019) or on first translating Tweets into English and then using VADER sentiment analysis software (Strengell & Sigg, 2018). For sentiment analysis in Spanish there is an open-source software available for sentiment detection from Spanish social media text (Pérez et al., 2021).

The body of prior work on sentiment analysis from Spanish Tweets includes the study (Dubey, 2020) where the author compares Twitter data from different countries, including Spain, to analyse the emotion of tweets. A lexicon-based approach based on the NRC lexicon is being used. The NRC lexicon is based on a classification of words into 8 basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and into positive or negative. (Lopez & Gallemore, 2021) present an open data set containing tweets gathered worldwide during the COVID-19 pandemic. The data set also contains Spanish tweets and an analysis for both named entities and sentiment (Lopez & Gallemore, 2021).

## **Methodology**

We used data scraped from the social media during COVID-19. Although the start of the COVID-19 pandemic was officially declared by the World Health Organization on 11th of March 2020 (World Health Organization, 2020), it is important to acknowledge that at that time there were already significant indicators of the ongoing crisis on both traditional media and social media platforms. The World Health Organization declared on the 5th of May 2023 that the COVID-19 no longer represents a global health emergency. As it is not meaningful to use the official starting and ending dates for the COVID-19 pandemic, we collect data from a longer period of time. Our study does not focus on sentiment during the COVID-19 pandemic, but rather on the public opinion on healthcare during that time. Because of this, also tweets from the time period prior and after the pandemics are studied.

Twitter was selected as the data collection platform due to its popularity and large quantity of user-generated content. The Python Snsrape library is used to access and retrieve Twitter data. This tool allows scraping tweets based on search queries and date ranges, among other parameters.

Once the data collection using Snsrape has been completed, the data will be pre-processed to ensure the quality and reliability of the dataset. The next step will be to perform sentiment analysis on it. We plan to analyse the sentiments of the Finnish tweets using the VADER sentiment analysis tool. This tool is especially made for analysing social media data (Hutto & Gilbert, 2014). The Finnish Tweets are first translated into English with Google translate. This is an approach also used by (Strengell & Sigg, 2018). Another alternative would have been to use a Finnish sentiment lexicon, such as SELF/FEIL (Öhman, 2021) and perform

a lexicon-based sentiment analysis on Finnish tweets. However, the advantage of using VADER is that it is especially tailored for social media text. For sentiment analysis in Spanish, we use the pysentimento toolkit (Pérez et al., 2021), which is especially created for analysing Spanish tweets.

To assess the quality and uniformity across Finnish and Spanish languages, the results of the automatic sentiment analysis are evaluated with a manually annotated data set. This data set is a representative sample of the retrieved Finnish and Spanish tweets, and it is annotated by the researchers using two annotators for each tweet and a third one in case of disagreement.

## **Conclusions and expected results**

This paper presents ongoing research activities. The expected results of the research will add knowledge on how public opinion can be monitored using user generated data and automated sentiment analysis methods. In this research Twitter data from Spain and Finland concerning healthcare services is under study. As a representative sample of the data is manually annotated and the results of the automatic analysis are compared with it, we expect to be able to draw conclusions and to provide recommendations on the quality and best practices concerning the use of automated sentiment analysis to monitor public opinion. To the best of our knowledge, this is the first time that Finnish and Spanish Twitter data concerning healthcare services is used to mine public opinion.

The research method has some limitations. As we know, the number of sources for user generated data are several. To gain more reliable results, we could extend the data collection also to other social media platforms such as Facebook, Instagram and various discussion forums to name just a few. Additionally, user surveys and interviews could be implemented in parallel to gain additional insight on user opinion on health care during COVID-19.

In addition to the methodological contributions mentioned above, the study aims to shed light on the public opinion on healthcare services during the different phases of the COVID-19 pandemic and to highlight differences in Spain and Finland.

## References

- Ainley, E., Witwicki, C., Tallett, A., & Graham, C. (2021). Using Twitter Comments to Understand People's Experiences of UK Health Care During the COVID-19 Pandemic: Thematic and Sentiment Analysis. *Journal of Medical Internet Research*, 23(10), e31101. <https://doi.org/10.2196/31101>
- Byrd, K., Mansurov, A., & Baysal, O. (2016). Mining Twitter data for influenza detection and surveillance. *Proceedings of the International Workshop on Software Engineering in Healthcare Systems*, 43–49. <https://doi.org/10.1145/2897683.2897693>
- Dubey, A. D. (2020). Twitter Sentiment Analysis during COVID-19 Outbreak. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.3572023>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/ICWSM.V8I1.14550>
- Lopez, C. E., & Gallemore, C. (2021). An augmented multilingual Twitter dataset for studying the COVID-19 infodemic. *Social Network Analysis and Mining*, 11(1). <https://doi.org/10.1007/S13278-021-00825-0>
- Öhman, E. (2021). *SELF & FEIL: Emotion and Intensity Lexicons for Finnish*. <http://arxiv.org/abs/2104.13691>
- Our World in Data. (2023). *Our World in Data*. <https://ourworldindata.org/>
- Pai, R. R., & Alathur, S. (2018). Assessing mobile health applications with twitter analytics. *International Journal of Medical Informatics*, 113, 72–84. <https://doi.org/10.1016/j.ijmedinf.2018.02.016>
- Pérez, J. M., Giudici, J. C., & Luque, F. (2021). *pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks*. <https://arxiv.org/abs/2106.09462v1>
- Pollack, C. C., Gilbert-Diamond, D., Alford-Teaster, J. A., & Omega, T. (2021). Language and Sentiment Regarding Telemedicine and COVID-19 on Twitter: Longitudinal Infodemiology Study. *Journal of Medical Internet Research*, 23(6), e28648. <https://doi.org/10.2196/28648>

- Pool, J., Namvar, M., Akhlaghpour, S., & Fatehi, F. (2022). Exploring public opinion about telehealth during COVID-19 by social media analytics. *Journal of Telemedicine and Telecare*, 28(10), 718–725.  
<https://doi.org/10.1177/1357633X221122112>
- Quoc Nguyen, D., Vu, T., Tuan Nguyen, A., & Research, V. (2020). *BERTweet: A pre-trained language model for English Tweets*. 9–14.  
<https://doi.org/10.18653/v1/2020.emnlp-demos.2>
- Sewalk, K. C., Tuli, G., Hswen, Y., Brownstein, J. S., & Hawkins, J. B. (2018). Using Twitter to Examine Web-Based Patient Experience Sentiments in the United States: Longitudinal Study. *Journal of Medical Internet Research*, 20(10), e10043.  
<https://doi.org/10.2196/10043>
- Strengell, N., & Sigg, S. (2018). Local emotions - Using social media to understand human-environment interaction in cities. *2018 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2018*, 615–620. <https://doi.org/10.1109/PERCOMW.2018.8480364>
- Tavoschi, L., Quattrone, F., D'Andrea, E., Ducange, P., Vabanesi, M., Marcelloni, F., & Lopalco, P. L. (2020). Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Human Vaccines & Immunotherapeutics*, 16(5), 1062–1069.  
<https://doi.org/10.1080/21645515.2020.1714311>
- Tsai, M. H., & Wang, Y. (2021). Analyzing Twitter Data to Evaluate People's Attitudes towards Public Health Policies and Events in the Era of COVID-19. *International Journal of Environmental Research and Public Health*, 18(12), 6272.  
<https://doi.org/10.3390/ijerph18126272>
- Vankka, J., Myllykoski, H., Peltonen, T., & Riippa, K. (2019). Sentiment Analysis of Finnish Customer Reviews. *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 344–350.  
<https://doi.org/10.1109/SNAMS.2019.8931724>
- World Health Organization. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
- Zunic, A., Corcoran, P., & Spasic, I. (2020). Sentiment Analysis in Health and Well-Being: Systematic Review. *JMIR Medical Informatics*, 8(1), e16023.  
<https://doi.org/10.2196/16023>

## **Summary of Analyzing How Process Mining Reports Answer Time Performance Questions**

Capitán-Agudo, C.<sup>1,2\*</sup>; Salas-Urbano, M.<sup>2</sup>; Cabanillas, C.<sup>2,3</sup>; Resinas, M.<sup>2,3</sup>

*\*Corresponding author - presenter*

<sup>1</sup> *ccagudo@us.es; University of Seville, Spain*

<sup>2</sup> *SCORE Lab, Spain*

<sup>3</sup> *13US Institute, Spain*

Keywords: process mining · time performance · qualitative analysis · quantitative analysis · grounded theory

### **Background of the study**

Business process management comprises the set of methods, techniques and tools for discovering, analysing, redesigning, executing and monitoring processes (Dumas, La Rosa, Mendling, & Reijers, 2018). The information of a process is usually available in the form of a log, which is a record of the actions and information managed in the process. Process mining (PM) is the discipline that uses techniques to extract value from event logs, which are records of the actions and information managed in processes. For example, some techniques check that processes are executed correctly (Van der Aalst, 2016). Owing to its importance, methodologies and case studies have been developed to understand and guide the use of PM techniques. For example, *PM<sup>2</sup>* (Van Eck, Lu, Leemans, & Van Der Aalst, 2015) identified the typical stages in these projects.

By contrast, other authors have focused on understanding the use of PM techniques in specific contexts. For example, in (Klinkmüller, Müller, & Weber, 2019) the authors evaluated qualitatively reports submitted to an annual competition called the BPI Challenge (BPIC), in which is provided a real-life event log by an organization along with business questions. The participants submit reports with their answers together with the analysis performed. In the

evaluation, the authors focused on understanding the use of visual representations and their information. Alternatively, in (Zerbato, Soffer, & Weber, 2021) the authors performed an empirical study to observe how analysts tackled the initial exploration of PM.

### **Aim of the study and novelty**

In our previous study published in the 20<sup>th</sup> International Conference on Business Process Management (Capitán-Agudo, Salas-Urbano, Cabanillas, & Resinas, 2022) we conducted a systematic analysis of BPIC reports answering time performance questions. We focused on them because it is one of the most common aspects in PM. The aim of the study was to discover what operations (e.g., manipulation methods, graphical representations) are used to address this type of questions, relationships between them, and differences in the context in which they are used (e.g., types of answers, objectives) to help and to ease time performance analysis of business processes. We complemented the previous work in (Klinkmüller et al., 2019) by not focusing our study to representation operations. Also, we complemented the work in (Zerbato et al., 2021) by focusing on the analysis.

### **Methodology**

We applied a mixed method approach similar to (Emamjome, Andrews, & ter Hofstede, 2019), which combines qualitative and quantitative research methods. In this summary we have included the following research questions of our previous study:

- RQ1: What operations are used to answer the questions on time performance?
- RQ2: What types of answers to time performance questions can be identified?
- RQ3: How does the context affect the similarity of the answers to time performance questions?

To answer them, the first step was the selection of BPICs that addressed time performance business questions. We found a total of 7 questions in 4 different editions: 2015, 2017, 2019 and 2020. They were classified depending on the goal to be achieved: differences, whose goal is finding differences in the throughput of different processes, fragments whose goal is calculating the throughput of process parts; and bottlenecks, whose goal is to find bottlenecks in the process. We also identified that some questions could achieve more than one goal (see Table 1). For our study, we considered the reports that had a specific section to respond them.

Goal	Questions
Differences	Q15, Q17, Q20A
Fragments	Q17, Q19, Q20A, Q20B
Bottlenecks	Q20C, Q20D

*Table 1: Questions per each goal (e.g., Q20A represents question 1 of BPIC 2020)*

Once we had identified these answers, we coded them following an inductive category development based on several iterations, inspired by the methodology of Grounded Theory (Stol, Ralph, & Fitzgerald, 2016). We followed this approach, because Grounded theory is commonly used with qualitative data to generate concepts, which could be the operations used in the answers in our case.

Initially, we applied an open coding to the answers, which consisted in reading them and making annotations of parts related to operations. Afterwards, we grouped the annotations related to the answers to some questions to identify the used operations. For example, we found similar annotations in two different reports (see Table 2). In both cases, different filters of traces are applied, so we grouped them in an operation called “Filter traces”. Nevertheless, the criterion used in each filter is different, so we distinguished them with another subcode called



variant, which represented how the operation was implemented. The coding process was iterative, and it finished when no new operations and variants were obtained.

Text in the report	Annotation	Operation	Variant
For the former case, filtering was performed by designating A_Pending as forbidden	Filter traces depending on the lack of A_Pending	Filter traces	Filter traces by activities
By filtering all cases that did not have a project number in Disco	Filtering of traces without project number in Disco	Filter traces	Filter traces by organizational units

*Table 2: Examples of coding*

Using this data, we created a dataset that related the identified operations and variants to each answer of the BPICs. We applied quantitative methods in this dataset to answer the research questions, such as a machine learning algorithm and a measure frequently used in them. For RQ1, we applied descriptive statistics, because they can provide an overview of the operations. For RQ2, we applied the KMeans clustering algorithm, whose input was a Boolean matrix where the rows represented the answers, the columns represented operations, and the cells represented whether an operation was performed (1) or not (0). We decided to use this algorithm, because it allows to discover groups in the data, which could be perceived as types of answers. We applied the mean Silhouette index to determine the optimal number of clusters, which was particularly used, because it does not need a training set to evaluate the results, making it appropriate for our analysis since we did not know which groups could be in the answers. For RQ3, we calculated the Sørensen-Dice coefficient (Sørensen, 1948) between the operations of pairs of answers, because we were interested in determining whether the questions were being answered differently depending on various aspects (e.g., their objectives).

Specifically, we opted for this similarity measure, because it considered the whole answer in the similarity calculations, and this is more representative for our case (e.g., Jaccard Index only contemplates the intersection of elements for the similarity calculations).

## Results

In this section we describe the main results of the questions. The result of RQ1 is a catalogue of 55 operations and 137 variants, which we have classified in 6 different groups (see Table 3).

Group	Description
Analyse time	Temporal analysis
Manipulate data	Reorganization of elements of the process
Calculate statistics	Application of descriptive statistic measures
Represent	Display insights
Identify elements	Highlight aspects based on criteria

Others	Do not fit in any other group
--------	-------------------------------------

Table 3: Summary of the catalogue of operations and variants

In RQ2 we obtained the best groupings for 4 clusters, which had an average Silhouette Index of 0.12. This value means that the clusters are unstructured, but it makes sense due to the variety of operations applied in the answers. The resulting clusters represented different ways to answer time performance questions of business processes with different goals, and they were characterised depending on the operations used along with their frequencies (see Table 4).

Type of answer	Main objective	Main answer Characteristics
Straight forward	Gain insights quickly	Low number and low variety of operations
Exhaustive	Obtain different results	High number and high variety of operations
Difference finder	Find time performance differences	High number of time performance comparisons
Manipulatory	Obtain limited data subsets	High number of manipulation operations

Table 4: Types of answers in the clusters

In RQ3, we found that the greatest average similarity generally occurred between answers that belonged to the same question. Secondly, we identified that the questions of BPIC 2020 were more similar to questions of other challenges than itself, if we consider them as a single question. This suggests that the objective is more important than the event log in the similarity

of the answers. Thirdly, we discovered high similarities between questions that shared answer types.

## **Conclusion and limitations**

In this work we have applied qualitative and quantitative techniques to provide an overview of the current state-of-practice to answer time performance questions related to business processes. As a result, we have discovered a catalogue of operations that are used, different ways to perform them and insights about the factors that could influence in them.

However, our work has some limitations. The first limitation is the risk of subjective bias in the annotations. To mitigate that, two authors made annotations and performed the coding of subsets of reports independently, and then shared the results. If they disagree in some coding aspect, the four authors discussed the problem to find a solution. Secondly, our work addresses typical time performance questions, but it does not cover all questions. Third, not all operations used by the analysts might appear in the reports, but since they are limited in space, we believe that the most relevant are the ones that appear.

## **References**

- Capitán-Agudo, C., Salas-Urbano, M., Cabanillas, C., & Resinas, M. (2022). Analyzing How Process Mining Reports Answer Time Performance Questions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13420 LNCS, 234–250. [https://doi.org/10.1007/978-3-031-16103-2\\_17](https://doi.org/10.1007/978-3-031-16103-2_17)
- Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). *Fundamentals of business*

- process management: Second Edition. In *Fundamentals of Business Process Management: Second Edition*. <https://doi.org/10.1007/978-3-662-56509-4/COVER>
- Emamjome, F., Andrews, R., & ter Hofstede, A. H. M. (2019). A case study lens on process mining in practice. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11877 LNCS, 127–145. [https://doi.org/10.1007/978-3-030-33246-4\\_8](https://doi.org/10.1007/978-3-030-33246-4_8)
- Klinkmüller, C., Müller, R., & Weber, I. (2019). Mining Process Mining Practices: An Exploratory Characterization of Information Needs in Process Analytics. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11675 LNCS, 322–337. [https://doi.org/10.1007/978-3-030-26619-6\\_21](https://doi.org/10.1007/978-3-030-26619-6_21)
- Sørensen, T. (1948). A method of establishing group of equal amplitude in plant sociobiology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *American Journal of Plant Sciences*, 5, 1–34.
- Stol, K. J., Ralph, P., & Fitzgerald, B. (2016). Grounded theory in software engineering research: A critical review and guidelines. *Proceedings - International Conference on Software Engineering, 14-22-May-*, 120–131. <https://doi.org/10.1145/2884781.2884833>
- Van der Aalst, W. (2016). Process mining: Data science in action. In *Process Mining: Data Science in Action*. <https://doi.org/10.1007/978-3-662-49851-4>
- Van Eck, M. L., Lu, X., Leemans, S. J. J., & Van Der Aalst, W. M. P. (2015). PM2: A process mining project methodology. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9097, 297–313. [https://doi.org/10.1007/978-3-319-19069-3\\_19/COVER](https://doi.org/10.1007/978-3-319-19069-3_19/COVER)
- Zerbato, F., Soffer, P., & Weber, B. (2021). Initial Insights into Exploratory Process Mining



## Bridging Researchers- Developing Transferable Skills & Competences for Next-Gen Researchers

Practices. *Lecture Notes in Business Information Processing*, 427 LNBIP, 145–161.

[https://doi.org/10.1007/978-3-030-85440-9\\_9](https://doi.org/10.1007/978-3-030-85440-9_9)



## **What role do Information and Communication Technologies play in fuelling animosity? Delving into the transmission channels through affective polarisation is transmitted online**

Romero-Martín, G.<sup>1\*</sup>; Caraballo-Pou, M.Á.<sup>2</sup>; Merchán-Hernández, C.<sup>3</sup>

\*Corresponding author – presenter.

<sup>1</sup> gonrommar@alum.us.es; Department of Economics and Economic History, University of Seville, Spain.

<sup>2</sup> Department of Economics and Economic History, University of Seville, Spain.

<sup>3</sup> Department of Sociology, University of Seville, Spain.

### **Keywords**

Affective polarisation, animosity, disinformation, Information and Communication Technologies, social media.

### **Background of the study**

The world is facing a crisis based on disinformation and fuelled by Information and Communication Technologies (ICT) that is confronting society on crucial issues that will determine our future (Collomb, 2014). According to Iyengar & Westwood (2015), hate towards dissidence has even overcome racial prejudice nowadays, depicting a relevant problem that threatens dialogue and consensus. In addition, Reiljan (2020) highlights that this social phenomenon has sharply increased in Western democracies over the last few years.

Although ICT have democratised access to information, they have also fuelled one of the major problems facing contemporary societies: information disorder and fake news (Taddeo et al., 2022). The use of the Internet threatens deliberative democracy as it increases selective exposure to like-minded content and avoids ideologically dissonant information. From a sociological approach, in the early stages of the Internet era, virtual communication was thought to enlarge mentality and engage citizens in a deliberative process that would lead to a mutual

uplifting of minds (Habermas, 2006). Nevertheless, ICT, since their outbreak, not only have not increased mutual understanding, but they have also weakened citizens' tolerance towards political differences. We now discuss less and worse, and the variety of the information consumed is increasingly biased or fake (Yarchi et al., 2020).

### **Aim of the study, originality, and novelty**

While there is a wide range of academic literature about hate and ICT, the links with affective polarisation (AP) remain diffuse. AP refers to distrust and dislike towards out-partisans (out-group polarisation) and, at the same time, a sense of recognition and affection towards partisans (in-group polarisation), (Iyengar et al., 2012). Current work about the issue at hand show that there is a need for delving into the factors that justify the rise of animosity nowadays (Reiljan, 2020).

Consequently, this work provides an updated review and relates analytically the principal studies on the link between ICT and the rise of AP. By this, we aim to contribute to a better understanding of how some online dynamics can boost AP in the real world. The current paper can be useful for social researchers who need a clear and holistic approach to the role that ICT play when observing the global rising trend of animosity. Also, this work is a response to the constant demand over time from the scientific community to provide more theoretical and empirical work about the specific connection between ICT and AP (Yarchi et al., 2020).

### **Methodology**

Based on the structure of a systematic review, this publication aims to select the most remarkable existing literature by following a replicable and transparent process. The systematic



review started by outlining the principal research questions. In this sense, questions such as “How do ICT operate when fuelling and spreading AP?”, “Which are the main transmission channels through which AP is disseminated?”, and “Which are the interrelations among these transmission channels?” formed the basic criteria of selection of the literature for inclusion in our database.

In the second step, we considered a list of descriptors which aimed to identify related publications that included any of the chosen terms in the title, abstract, or keywords. Based on the results obtained, the selection criteria were established: 1) Publications needed to be published in an academic journal, in the press, or be part of a book chapter. 2) The selected cases needed to be connected to the determinants, transmission channels, sources and consequences of AP in the online context. 3) These publications must have been written in English, French or Spanish.

In the third step, we conducted different searches focusing on the title, the abstract and the keywords to enable replication of the search procedures in all databases. The search and selection of the publications applying the inclusion and exclusion criteria were performed by two different researchers, seeking to maximise precision in the process.

The fourth step aimed to capture the literature that was important to solve our research questions. The selected contributions were further refined and classified into three categories related to the following strands in the fifth step: anonymity in social networks and online group participation; opinions interpreted as information; AI, hate and denialist movements; and other publications related to animosity and social media. Finally, in the sixth step, the selected literature is discussed and interpreted.

## **Results/findings and argumentation**

Regarding the aforementioned three strands, we considered a democratic society that interacts offline and online. Both dimensions can be regarded as a mirror, so what happens in the real world has an impact on the Internet and vice versa. Based on the preceding sections, six sources of AP when referring to ICT can be identified: AP due to disinformation, AP due to hate speech (anonymity), AP due to personal opinion, AP due to influential people's opinion, AP due to online group formation, and AP due to denialism/extremism.

Starting with AP due to disinformation, it is produced by AI systems, spread by ICT, and consumed by individuals and groups. Disinformation is classified and exposed to each individual/group according to their online behaviour, (Marozzo & Bessi, 2018). As represented, politicians and social media maintain a mutually beneficial relationship (Kertysova, 2018). Disinformation can be regarded as an input for individuals and groups and their opinions are based on what they consume online. This way, a direct source of AP (based on individuals' opinions) is found (Lelkes et al., 2017). Also, as Stier et al. (2018) highlight, others' opinions can also be considered accurate information.

In the case of AP due to anonymity, it produces an "online disinhibition effect" as individuals' behaviour is not publicly linked to themselves but to an impersonal profile. Anonymity online is widely conceived as a threat to consensus-building as it breeds incivility and undermines the possibility of a mutual uplifting of minds (Suk et al., 2022). Finally, we find AP due to disinformation, where groups of individuals become deniers of a sensitive social issue about which there has long been a tacit social consensus (environment, gender, identity, COVID-19...). Their reactions increase AP, as the group is going to express disapproval openly in

relation to this issue and reject those who do not think the same. In opposition, people who are for it will regard the first group as a threat to some achieved rights (Kertysova, 2018).

## **Conclusion, implications and limitations**

While the study of AP has been well-documented through surveys in offline contexts, little research has been done in the digital world. The literature review and analytical framework presented in this work provide a roadmap for further research about the role of ICT in fuelling AP. It aims to be useful to understand and map actors, transmission channels and relationships among variables. As has been shown, three strands of literature can be identified when analyzing the relationship between AP and ICT. Firstly, anonymous profiles in social networks and online groups. Secondly, opinions disguised as information. Lastly, Artificial intelligence, disinformation and denialist political movements. Based on these three strands of literature, six transmission channels can be found: disinformation, anonymous hate speech, personal opinions, people's opinions, online group formation and denialism/extremism. The classification of the three thematic blocs and transmission channels provides social scientists new horizons of research in the study of animosity in digital contexts.

The analysis of AP and ICT cannot be separated from what happens offline: the irruption of multiparty systems (Wagner, 2021), the use of polarising speeches as an electoral weapon (Druckman & Levendusky, 2019), and the study of the psychology of the voters' profile (Rogowski & Sutherland, 2016) are some of the individual and institutional dimensions of AP that explain the rise of hatred in the online world. In this sense, an important gap in the literature concerns how political parties use their institutional and fake online accounts to spread fake news and attack political opponents. In this context, parties are sometimes helped by

ideologically-biased journalists and radical online movements who swirl rumours about political rivals as a way to erode their public image. There are plenty of cases of online hatred campaigns towards politicians that have ended in court. By the time the vilified person has been able to demonstrate their innocence, their honour has been so eroded that they cannot continue exercising their mandate anymore. Also, further efforts should be done to understand under which circumstances ICT promote freedom of speech or, on the contrary, which factors explain animosity among individuals (Oz & Cetindere, 2023).

Landon-Murray et al. (2019) also highlight the importance of cyberethics when combatting fake news, another relevant outcome of this work. Also, Thompson et al. (2022) outline the role that ICT have in restoring integrity online, conceiving blockchain as a key tool to fact-check information. Another interesting research line would be how AP violates the principles of impartiality and objectivity when referring to each country's press and which media prioritise sensationalism over facts. Is AP a profitable business? Has the mass media business experienced a paradigm shift in its mission? Either way, the quality of mass media is also the quality of democracy.

## References

- Collomb, J.-D. (2014). The Ideology of Climate Change Denial in the United States. *European Journal of American Studies*, 9(1). <https://doi.org/10.4000/ejas.10305>
- Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization? In *Public Opinion Quarterly*, 83(1), 114-122. <https://doi.org/10.1093/poq/nfz003>

- Habermas, J. (2006). Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. *Communication Theory*, 16(4), 411-426. <https://doi.org/10.1111/j.1468-2885.2006.00280.x>
- Iyengar, S., & Westwood, S. J. (2015). Fear and Loathing across Party Lines: New Evidence on Group Polarization. *American Journal of Political Science*, 59(3), 690-707. <https://doi.org/10.1111/ajps.12152>
- Kertysova, K. (2018). Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered. *Security and Human Rights*, 29(1-4).
- Lelkes, Y., Sood, G., & Iyengar, S. (2017). The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect. *American Journal of Political Science*, 61(1), 5-20. <https://doi.org/10.1111/ajps.12237>
- Marozzo, F., & Bessi, A. (2018). Analyzing polarization of social media users and news sites during political campaigns. *Social Network Analysis and Mining*, 8(1). <https://doi.org/10.1007/s13278-017-0479-5>
- Oz, M., & Oz Cetindere, E. (2023). Perceived Social Sanctions and Deindividuation: Understanding the Silencing Process on Social Media Platforms. *International Journal Of Communication*, 17, 22. <https://ijoc.org/index.php/ijoc/article/view/20116/4170>
- Reiljan, A. (2020). ‘Fear and loathing across party lines’ (also) in Europe: Affective polarisation in European party systems. *European Journal of Political Research*, 59(2), 376-396. <https://doi.org/10.1111/1475-6765.12351>
- Rogowski, J. C., & Sutherland, J. L. (2016). How Ideology Fuels Affective Polarization. *Political Behavior*, 38(2), 485-508. <https://doi.org/10.1007/s11109-015-9323-7>

- Stier, S., Bleier, A., Lietz, H., & Strohmaier, M. (2018). Election Campaigning on Social Media: Politicians, Audiences, and the Mediation of Political Communication on Facebook and Twitter. *Political Communication*, 35(1), 50-74.  
<https://doi.org/10.1080/10584609.2017.1334728>
- Suk, J., Coppini, D., Muñiz, C., & Rojas, H. (2022). The more you know, the less you like: A comparative study of how news and political conversation shape political knowledge and affective polarization. *Communication and the Public*, 7(1), 40–56.  
<https://doi.org/10.1177/20570473211063237>
- Taddeo, G., de-Frutos-Torres, B., & Alvarado, M.-C. (2022). Creators and spectators facing online information disorder. Effects of digital content production on information skills. *Comunicar*, 30(72), 9–20. <https://doi.org/10.3916/C72-2022-01>
- Thompson, R. C., Joseph, S., & Adeliyi, T. T. (2022). A Systematic Literature Review and Meta-Analysis of Studies on Online Fake News Detection. *Information*, 13(11), 527.  
<https://doi.org/10.3390/info13110527>
- Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2020). Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, 38(1-2), 98-139.  
<https://doi.org/10.1080/10584609.2020.1785067>
- Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, 69, 102199. <https://doi.org/10.1016/j.electstud.2020.102199>