Anamaria NĂSTASĂ, PhD Candidate
E-mail: anamaria.nastasa@incsmps.ro
National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania
Doctoral School of Sociology, University of Bucharest, Romania

Associate Professor Monica Mihaela MAER MATEI, PhD (corresponding author)
E-mail: monica.matei@csie.ase.ro
Bucharest University of Economic Studies, Romania
National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania

Senior Researcher Cristina MOCANU, PhD
E-mail: cristina.mocanu@incsmps.ro
National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania

ARTIFICIAL INTELLIGENCE: FRIEND OR FOE? EXPERTS' CONCERNS ON EUROPEAN AI ACT

Abstract. *In the last decade, there have been numerous innovations in artificial intelligence technologies in many domains, many innovations more or less favourable. However, artificial intelligence has been and is the subject of multiple controversies, such as the perpetuation of inequalities, discrimination, biased decisions, and other issues regarding transparency and data protection. These problems destroy the trust of citizens and institutions in artificial intelligence. Consequently, the European Commission proposed the AI Act, a regulation for assessing Ai products or services. Our study explores experts' main concerns on artificial intelligence technologies. In the present paper, we analysed the feedback provided by 262 stakeholders on the proposal of the European Commission regarding artificial intelligence through a text mining approach using Latent Dirichlet allocation. The prevalent topics were related to AI applications in industry, transparency and responsibility, and AI technologies testing. The analysis also revealed topic differences based on the type of organisation, especially between consumer organisations and academic/research institutions.*

Keywords*: Artificial Intelligence, AI Act, Text Mining, Latent Dirichlet Allocation*

JEL Classification: O35, D83

DOI: 10.24818/18423264/57.3.23.01

1. Introduction

Artificial intelligence (AI) has become increasingly pervasive in daily life, education and professional training, labour markets, law, and other domains. However, the permeation of life with different AI technologies has begun to raise issues related to the ethics behind algorithmic decisions, the perpetuation of inequalities, issues related to transparency, and many others. The use of AI technologies to bring advantages requires trust in the decisions made by these technologies. In order to achieve this goal, the European Commission's introduced an incipient legislative initiative for increasing citizens' trust in artificial intelligence. The AI Act aims to regulate the use and production of artificial intelligence products and services in the EU area.

1.1. AI Act

One of the first initiatives at the European level tackling artificial intelligence was the White Paper on Artificial Intelligence proposed by the High-level expert group on Artificial Intelligence. The White Paper states that AI must be designed to create two ecosystems: one based on excellence, and one based on trust. The first ecosystem, based on excellence, aims to build collaborations between the public and private sector organisations to stimulate innovations in artificial intelligence (European Commission, 2020). The second ecosystem strives to identify and mitigate the potential threats of AI technologies to citizens. Accordingly, the European Commission wants to create, on the one hand, a legal framework conducive to collaborations between different types of institutions for knowledge production and transfer that will lead Europe to be the global leader in AI technological innovation. On the other hand, EC is committed to building trust among its citizens by creating a regulatory framework to assess AI systems that can raise concerns about the safety of individuals.

The promissory objective of constructing an ecosystem of trust was materialised incipiently through the Artificial Intelligence (AI) Act. The AI Act was proposed in April 2021. The document is a set of rules to regulate all Artificial Intelligence technologies used or produced in the EU. The first section of the proposed regulation lays down the scope of the Act and its definitions (of AI systems, providers, and users of AI systems). Two main objectives can describe the European Commission regulation aims: (1) to assess if AI technologies are safe and harmonised with fundamental human rights, other European regulations, and EU values; (2) to offer a regulatory tool that prevents market fragmentation while fostering AI innovation and investment in EU states (European Commission, 2021; Mazur & Renata, 2023).

The European Commission defines an AI system as "software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with" (European Commission, p. 39, 2021). The listed techniques and approaches are machine learning, logic, knowledge-based, and statistical approaches.

In the second section, the European Commission elaborates a list of forbidden AI practices discerning between three types of risks posed by these technologies: unacceptable, high, and low or minimal. Unacceptable practices clarified in the document are social scoring, manipulation via subliminal techniques that exploit individual or group vulnerabilities, and real-time biometric identification.

The third section firstly discusses the systems classified as high-risk for "health, safety or fundamental rights" (p. 13), simultaneously distinguishing two types of high-risk systems: (1) products or their components regulated already through other European legislative assessments and (2) other AI systems that have implications for fundamental human rights (such as educational, critical infrastructure, biometric identification, essential services, law, employment, justice, and democracy and migration) (Mazur & Renata, 2023, European Commission, 2021). The second chapter of the third section lays out the requirements, and the third chapter sets out the duties for the actors that offer AI technologies or services. The fourth chapter lays the framework for the third parties responsible for assessing AI technologies, and the fifth chapter thoroughly describes the procedures for assessment.

The fourth section presents systems required to comply with transparency obligations, such as deep fake applications, systems that interact with humans, and biometric technologies. The fifth section concerns the obligations of national authorities in creating regulatory sandboxes to verify AI technologies and the potential measures to reduce the obligations of small and medium-sized enterprises.

The implementation and administration of the AI Act are discussed in the sixth, seventh, and eighth sections. EC envisions the creation of governance systems at the level of the European system and the member states. The Act also stipulates the designated national authorities, including a national supervisory authority, for handling the processes of implementation and application. The European Data Protection Supervisor will be an institution involved in these two processes. The AI Act also states that there will be a database for high-risk systems where providers must register their products. The obligations for providers regarding monitoring and reporting AI technologies during their entire life cycle are also discussed in the Act. Moreover, market surveillance authorities would be accountable for the compliance of high-risk technologies with the proposed regulation.

The ninth section of the AI Act presents suggestions for minimal-risk AI providers with respect to codes of conduct. The EC encourages these companies to engage in sustainable goals such as climate change, accessibility, and diversity. It is also recommended to include the stakeholders' feedback in the design and development of AI.

The last three sections of the Act present the final conditions of the regulation. It particularly highlights the rules for the confidentiality of information during the implementation process and sets out the penalties for infringement of requirements. It also gives power to the EC in delegating and implementing supplementary acts and provisions needed for the implementation of the Act. On the other hand, the Act prescribes the obligation for the Commission to assess the regulation and update its content while regularly reporting a review of the evaluations.

1.2. Pros and cons of AI

AI ranks on top of the agenda of businesses that aim to take advantage of the vast opportunities emerging in the field, academics that must provide the skills but also the knowledge on the benefits and risks associated with AI, decision makers that must develop the legal framework, NGOs and trade unions organisations having to watch on how legal framework is developed and its effects on workers, clients, societies, individuals, etc.

Enactment of measures developed to regulate AI advancements is essential, but also the definition and measurement of empirical evidence already existing and functioning. AI is a complex technology and can unfold different trajectories in its development (Baruffaldi et al., 2020), and it is essential to define what is legal and what is beyond the scope of producing benefits for society.

Experts consulted on the topic of the future progress of AI were very optimistic about the high rhythm of developments in the field, estimating that in the decades to come, it will reach, to a large extent the human ability (Muller and Bostrom, 2016), pointing to the urging need to understand better the effects that this rhythm will have on societies because inadequately harnessed could have a substantial negative impact on humanity.

Experts and decision-makers are fully aware that a policy in the field co-shapes the field's dynamic, providing the tools to exploit the benefits of AI while controlling the risks. AI development is inevitable (Bareis and Katzenbach, 2022).

There are different sectors, such as healthcare, where developments in the field can be highly beneficial to societies, leading to services of higher quality and delivered to increasing segments of beneficiaries. They can support the current skills shortages in different sectors, but mostly in healthcare, by optimising services and processes. Of course, AI technology requires different skills that professionals in different sectors must develop.

The business sector can benefit significantly from AI development, with transportation and logistics, mining, as well as finance and banking, being in the front of the sectors that will be drastically affected (Nadimpalli, 2017). Companies could use AI to increase the standardisation and quality of products and services and protect human life in dangerous and degrading activities.

As a new and disrupting technology, there are many risks concerning AI, such as replacing workers and increasing redundancies, potential harm to fundamental rights, increasing opportunities to use personal information unethically and transgressing legal regulations, lack and transparency and responsibility to increase profits and surveillance, etc. (Kerr, Barry, and Kelleher, 2020; Henke et al., 2016).

As the disruptions generated by AI technologies are unprecedented, the focus put in different papers and policies on mitigating the possible risks and harm is a must. However, suppose that risks such as privacy violation and manipulation are foreseeable. In that case, some future consequences are still to be experienced, with sectors such as medicine, where any failure can have a huge impact on human life (Cheatham, Javanmardian, and Samandari, 2019).

In the light of possible AI risks and disruptions, our paper aims to explore experts' main concerns on artificial intelligence technologies and policy initiative on AI regulation. For this purpose, we examined the feedback provided by 262 stakeholders on the proposal of the European Commission regarding artificial intelligence (AI Act) through a text mining approach.

2. Methodology

2.1. Data description

Subsequently, after publishing the proposal of the AI Act in February 2021, stakeholders from multiple types of organisations were invited to provide feedback on this document. The period in which the stakeholders could provide their input on this proposal of the European Union was 26 April 2021 - 06 August 2021. A total of 303 feedback stakeholders gave valid feedback. In our analysis, we used 262 documents in English after excluding duplicates, empty feedback, and feedback provided in a language other than English. All the feedback documents were downloaded from the Commission site’ with a .pdf extension.

Almost a third of the documents with feedback on regulation were provided by representatives of companies/business organisations. At the same time, a quarter of these documents were from representatives of business associations, and 20% were from representatives of non-governmental organisations (NGOs).

Table 1. Number of documents by user type

User type	No. of documents	%
Company/business organisation	81	31.0%
Business association	65	24.9%
Non-governmental organisation (NGO)	51	19.5%
Other	19	7.3%
Academic/research Institution	19	7.3%
Trade union	11	4.2%
EU citizen	8	3.1%
Public authority	4	1.5%
Consumer organisation	3	1.1%

Regarding the countries of the organisations that had an input on the proposal of the European Union regarding AI, the majority were from Belgium (31%), Germany (16%), the United States (10%), the United Kingdom (7%) and France (7%). Other countries with a smaller representation among the feedback documents on the regulation were Switzerland, Italy, Sweden, Spain, Ireland, Poland, Finland, Denmark, Japan, Austria, Norway, Czech Republic, Bulgaria, Lithuania, China, Croatia, and Cyprus.

2.2. Method

In order to extract relevant information from the collection of documents representing feedback on the AI act, a probabilistic approach of a dimensionality reduction technique was employed. A topic modelling instrument, namely Latent Dirichlet Allocation (LDA), provided the latent semantic dimensions revealing the thematic structure of the debate around AI (Blei et al., 2003). The inputs used by LDA are a document term matrix (DTM) and the number of topics to be extracted.

The DTM is a structured representation of our unstructured data or corpus, represented by a matrix with D lines and V columns where D is the number of documents in the corpus and V is the vocabulary size. To be more precise, an element f_{ij} of this matrix shows the number of times term *i* occurs in document *j* (term frequency). Regularly, the vocabulary size is controlled by two means: (i) sparsity and (ii) a weighting system based on term frequency-inverse document frequency (Tf-idf) scores.

Sparsity is a value characterising a DTM, computed for each tern as:

$$sparsity_i = 1 - \frac{a_i}{D}$$

where a_i is the total number of documents where term *i* occur and D is the corpus size. The terms defined by higher values are eliminated. This means that the terms appearing in only a small part of the documents, generating zero cells in DTM, are excluded from the vocabulary. Removing sparse terms with a threshold of s% will remove those terms that do not appear in at most 1-s% of the documents.

This procedure will reduce the number of lines in DTM by eliminating those with many zero cells.

A weighting system is used to quantify the relevance of a word. This system gives a larger weight to terms encountering higher frequency in a document but a reduced frequency in the corpus. The values satisfying this condition are known as term frequency-inverse document frequency (Tf-idf). The tf-idf scores are computed by the formula:

$$tfidf_{ij} = f_{ij} \times \log\left(\frac{D}{1 + d_i}\right)$$

where: D is the number of documents in the corpus, f_{ij} is the frequency of word *i* in document *j*, and $d_i$is the number of documents containing term *i*.

This study's vocabulary selection was based on investigating the tf-idf distribution. Finally, we have used only terms with tf-idf scores above the median value.

The second input of an LDA function, the number of topics, was selected based on two criteria. The first metric, CaoJuan2009, is based on topic density (Cao et al., 2009). This method uses the distances among topics to establish the best LDA model. The second metric, Deveaud2014, maximises the distances among the topics (Deveaud et al., 2014).

In topic modelling, word occurrences and co-occurrences in the corpus are used to reveal a synthesis of the documents through their latent semantics. To be more precise, words describing the same concept are clustered within topics.

A latent structure behind the corpus could be captured by a singular-value decomposition of the tf-idf matrix. If D is the size of the corpus, and V is the vocabulary size, the decomposition will reveal: a matrix representing a document-topic matrix and the term-topic matrix:

$$\hat{A}_{D\times V} = \Theta_{D\times K} S_{K\times K} \Phi_{K\times V}$$

Where K is the number of the largest singular values retained.

This approach was extended to a probabilistic model, where these matrices have the same interpretation, although they reflect probability distributions (Blei et al., 2003). The suited probability distribution for modelling discrete data such as words and documents is the multinomial. If we have V possible outcomes, the probability mass function for a multinomial distribution with parameter $\beta = (\beta_1, \beta_2 \dots, \beta_V)$ is given by:

$$p(x|\beta) = \prod_{i=1}^{V} {\beta_i}^{x_i}$$

In the LDA framework, each document is characterised by a multinomial distribution with parameter θ over the K topics, and each topic is represented as a multinomial distribution with parameter φ over V words. The goal is to estimate θ for each document and φ for each topic. This is achieved through Bayesian inference that requires prior beliefs. In general, a mathematically convenient prior is used, namely a conjugate prior. A distribution is considered a conjugate if, used as a prior in the Bayes rule, produces a posterior distribution in the same parametric family (Murphy, 2012). The conjugate prior of a multinomial distribution is called a Dirichlet distribution. A Dirichlet distribution is a multivariate generalisation of the beta distribution with a probability density function:

$$p(\theta|\alpha) = \frac{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)}{\prod_{i=1}^{K} \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \dots \theta_k^{\alpha_k - 1}$$

Where α is the vector of the Dirichlet parameters and gamma $\Gamma(x)$ is gamma function:

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u}\, du$$

LDA is a generative probabilistic model. Therefore, the generation process of a document, behind LDA implies to:

(1) Sample from a Dirichlet distribution (Dir(α)) to extract a document-specific distribution over topics.

(2) Sample the words in the document from these topics. A conjugate Dirichlet prior (Dir(β)) is used.

The algorithm used to estimate the posterior distribution is Gibbs sampling. This Markov Chain Monte Carlo method approximates the distribution by repeatedly sampling from conditional distributions. The Gibbs sampler is used when we need to sample from a complicated joint distribution. Therefore, we cannot sample directly from the probability density function $p(x_1, x_2, \ldots, x_K)$, but we can sample from conditional distributions: $p(x_i|x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_K)$, producing a Markov chain, through the following steps (Hastie et al., 2009):

1. Initialisation: $x_k^{(0)}$, k=1,2…K
2. Generate $x_k^{(t)}$ from $p(x_k^{(t)}|x_1^{(t)}, \ldots x_{k-1}^{(t)}, x_{k+1}^{(t-1)}, \ldots x_K^{(t-1)})$, t=1,2…

Repeat step 2 until the joint distribution of $\left(x_1^{(t)}, x_2^{(t)} \ldots x_K^{(t)}\right)$ does not change.

For LDA, this algorithm will compute the probability that a topic k is represented by a specific word, given all other topic assignments to all other words. (Darling, 2011).

Obviously, the topic structure LDA produces depends on the number K of topics extracted. In this study, the selection of K was driven by two criteria. The first (Cao et al., 2009) relies on the correlation between two topics, computed as the standard cosine distance:

$$cor(\beta_i, \beta_j) = \frac{\sum_{v=1}^{V} \beta_{iv}\beta_{jv}}{\sqrt{\sum_{v=1}^{V}(\beta_{iv})^2}\sqrt{\sum_{v=1}^{V}(\beta_{jv})^2}}$$

Where β_i represents word-topic distribution vector for the topic i.

The stability of the topic structure is measured by the average cosine distance between each two topics:

$$stability(K) = \frac{\sum_i \sum_{j=i+1}^{K} cor(\beta_i, \beta_j)}{K(K-1)/2}$$

A structure is more stable if the stability value is smaller since the topics are more independent. The proposed method uses two concepts: topic density and model cardinality. Topic density represents the number of topics within the radius of r from topic $Topic_i$, denoted by $Density(Topic_i, r)$. The distance is computed by the average value of the cosine distance. The cardinality (C_n) of a topic model is given by the number of topics having a density less than n.

Using these concepts, the parameter K will be updated through the formula:

$$K_{n+1} = K_n + f(r) \times (K_n - C_n)$$

where $f(r)$ is the changing direction of r.

It starts with an initial model, where the number of topics is K_0, and sequentially train LDA models, computing the average cosine distance (r) and the densities of the model’s topics (Cao et al., 2009).

The second criterion used to select the number of topics is based on their word distributions, denoted by $p(w|k)$ (Deveaud et al., 2014). It computes the divergence between all topics of an LDA model by the Jensen-Shannon diverge measure:

$$D(k||k') = \frac{1}{2}\sum_{w \epsilon W_k \cap W_{k'}} p(w|k) log\left(\frac{p(w|k)}{p(w|k')}\right) + \frac{1}{2}\sum_{w \epsilon W_k \cap W_{k'}} p(w|k') log\left(\frac{p(w|k')}{p(w|k)}\right)$$

where W_k is the set of the most relevant words of the topic k, having the highest probabilities.

For the set of K topics extracted by LDA, denoted by T, the number of topics will be given by:

$$\hat{k} = argmax \frac{1}{K(K-1)} \sum_{(k,k')\in T} D(k||k')$$

LDA is based on the hypothesis that words and documents are exchangeable, meaning that word order is ignored. Hence, the model architecture relies on de Finetti’s theorem, stating that the distribution of a sequence of exchangeable random variables is a mixture of independent and identically distributed random variables.

Therefore, the main limitation in LDA comes from the fact that the topics are assumed to be independent, reducing its prediction capacity. This drawback could be eliminated by employing a new approach where the topic distribution is assumed to be the logistic normal. This is called the Correlated Topic Model (Blei & Lafferty, 2006). Future work will consider this approach to investigate experts' main concerns on artificial intelligence technologies.

3. Results

The investigation was undertaken on a corpus of 262 documents representing the feedback received by the European Commission in the public consultation process with respect to the AI initiative. The final document term matrix (DTM) used as input in the LDA model has the following characteristics: the size of the vocabulary is 1044 terms, and the sparsity is 71%. This DTM was obtained after performing the preprocessing operations: converting tokens to lower case, removing punctuation, numbers, stop words, and sparse terms. Moreover, the initial vocabulary was reduced by retaining only the terms with higher tf-idf scores. The threshold selection was based on the distribution of the tf-idf weights.

The findings are extracted from 12 topics produced by an LDA model. The selection of the number of topics was guided by the metrics explained in Section 2.2

and presented in Figure 1. However, the final decision was based on the topic interpretation aspects.

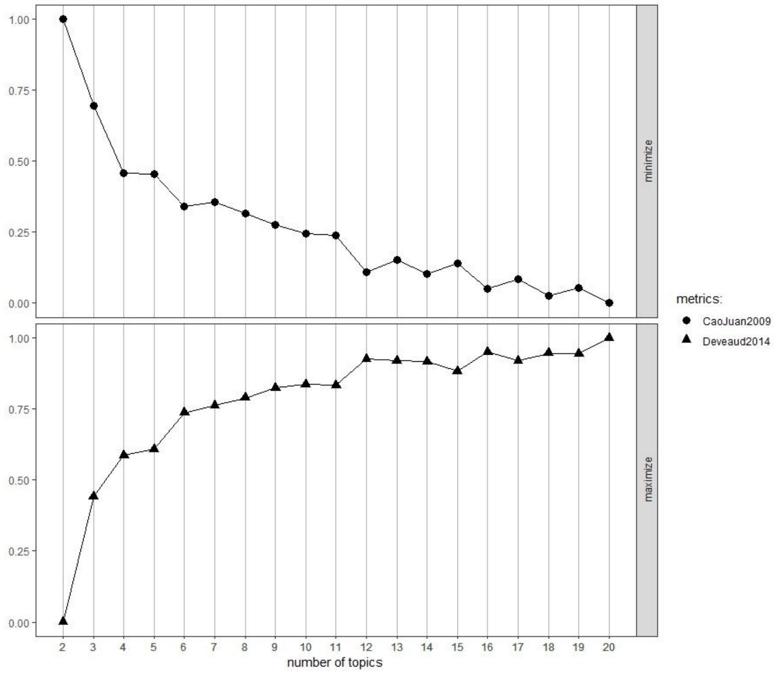


Figure 1. Criteria for the number of topics selection

The findings presented below are drawn from the investigation of theta (θ) distributions showing the probability of topic k occurring in document d and beta (β) distributions revealing the words' probability of belonging to that topic. Hence, the outputs illustrate which topics are prevalent in documents and which terms best represent each topic.

The themes dominating the discussion are those encountering higher θ probabilities in the 262 documents of the corpus. By descending sorting these probabilities, we have extracted the top 3 topics for each document. The ones emerging with the most significant frequency in this ranking are presented below through the terms depicting the highest Beta probabilities. Because of the large frequency of feedback documents received from companies and business associations, these three topics primarily explain the feedback received from these stakeholders.

The first topic, "Application of AI in the industry", has a twofold meaning. On the one hand, it implies the application of AI as a technology in the industry to foster innovation. On the other hand, it also suggests the application of AI ACT as a regulation in the industry, its standards, its potential effects, and its harmonisation with other legislative initiatives. Therefore, this topic is particularly representative for the companies that use AI technologies in their products and their main concerns regarding the initiative. The concerns voiced might be related to the regulation (“rules”, “standards”, “legal” requirements) and its complementarity with other existing regulations in the industry, but also its potential effects on the “companies” and the “development of their ”products (“applications”), their “innovation”, on the overall “market”. The EU initiative might pose constraints for some companies focusing on technology because it introduces new obligations and compliances for them besides those already existing on the market. Therefore, it creates some resistance from these companies to introducing the AI Act.

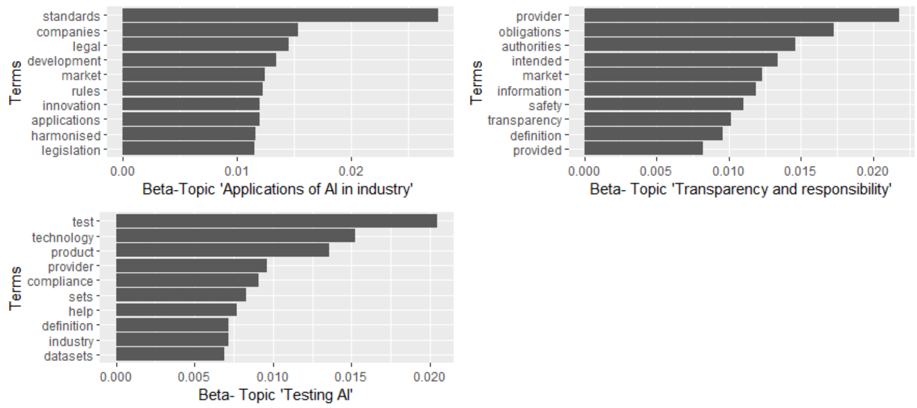


Figure 2. Prevalent topics - word assignment

The second topic synthesises the "Transparency and responsibility" issues and requirements. In scientific and public discourse, transparency is emphasised as an essential element of technology development. The transparency of technology regarding the disclosure of relevant information can increase consumer trust and engagement. Moreover, it is also a critical part of the AI Act because it is a prerequisite element of citizen trust. Under the AI regulation, providers that offer some types of AI technologies (deepfakes, technologies that interact with humans, use biometric identification, or make social categorisation) have transparency obligations. This topic is mainly crystallised on companies' feedback on this part of the AI Act. Accordingly, the topic explains the companies' worries regarding transparency and safety issues in relation to their obligations as providers of AI technologies on the market and authorities' control. Some of the organisations' arguments may be related to the definitions and pieces of information offered in the Act, but also the clearness of these transparency obligations. The obligations may also pose worries related to threats related to business secrets and product-related information for some organisations.

The third topic is associated with "AI testing". This theme is mainly associated with data sets used to test AI technologies before they are placed on the market and their entire life cycle. This topic is also related to AI Act provisions regarding the testing and verifying technologies on their entire life-cycle on the market (especially high-risk AI technologies). Some of the most prevalent terms for this topic are "test", "product", "technology", "datasets", and "sets".Therefore, in relation to the Act, among companies who voiced their feedback, there are significant worries about the datasets used in testing AI products. As in the case of transparency obligations, some organisations may be concerned about data leakage, which may be conducive to threats to trade secrets. Therefore, some companies may accentuate these risks and the need for safe environments for testing.

There are other topics that are very well represented in the feedback sent by NGOs, trade unions, or academic institutions. Hence, in the subsequent figures, we present the topics where theta distributions revealed significant differences depending on the respondents' category.

The topic "consumer education" is more prevalent among Trade Unions. The topic is represented by the words: "consumer", "education", "public", "national", "access", "right", "tools", "social", "protection", and "applications". This topic is focused on concerns related to the lack of skills and education among consumers regarding the use of artificial intelligence and their rights and protection against potential risks posed by AI technologies. The lack of digital literacy among citizens (and consumers of technology) might put them in a vulnerable position in the face of technologies that permeate many life aspects. This topic may also point out the risks of inadequate use of AI technologies in the educational sector.

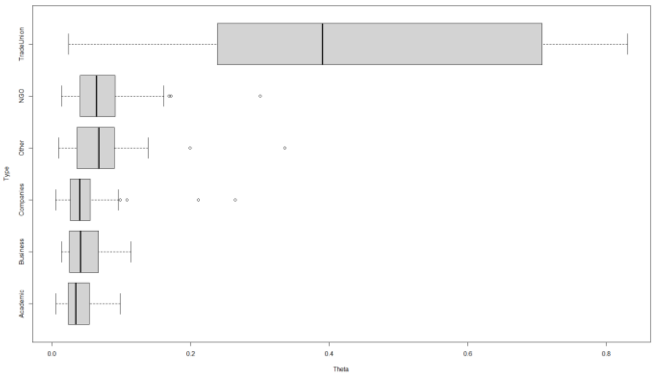


Figure. 3. Topic "Consumer Education"

On the other hand, based on the feedback received, academic/research institutions representatives are more likely to discuss on issues regarding "law enforcement", "fundamental rights", and "ethical concerns", meanwhile non-governmental organisations representatives are more prone to mention issues related to "ethical concerns" and "law enforcement". Therefore, compared to companies, academic and research representatives and NGOs involved in the Act's consultation process are more familiarised and connected with potential ethical concerns imposed by AI technologies (such as surveillance and through facial and other biometric identification techniques), AI technologies' risks to fundamental human rights, and law enforcement. Academics, researchers, and NGOs representatives are more likely to mention arguments about AI technologies' dangers and risks to individual users or consumers.

The topic of law enforcement is associated with ideas regarding the application of the AI Act to protect citizens and their fundamental rights. The most representative words related to "Law enforcement" theme are: "right", "fundamental", "legal", "protection", "law", "public", "enforcement", "framework", "practices", and "control This topic emphasises the role of the AI Act regulation in mitigating control practices and fundamental human rights through law enforcement.

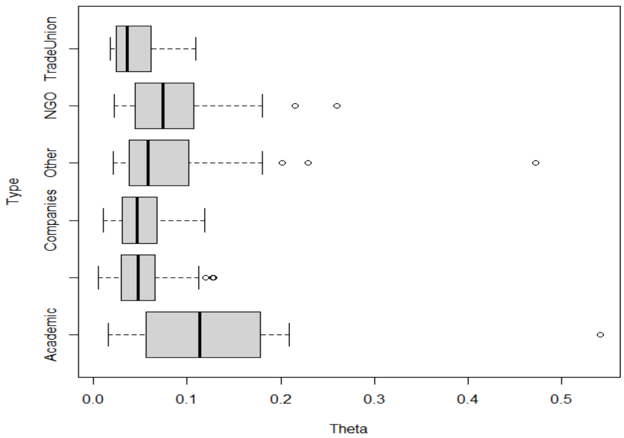


Figure 4. Topic “Law enforcement”

The topic of "fundamental rights" is more focused on data protection and personal privacy concerns. On the one hand, there are discussions regarding the complementation and overlapping of GDPR with the AI Act. On the other hand, issues regarding data protection and privacy (especially those related to General Data Protection Regulation) are increasingly common public discourse in relation to AI technologies, especially due to the presence of many controversies regarding personal data breaching events. Therefore, many organisations stress out the importance of this initiative in protecting personal data and protect human fundamental rights. The words related to this topic are: "right", "GDPR", "human", "decision", "impact", "fundamental", "law", "assessment", "personal", and "privacy".

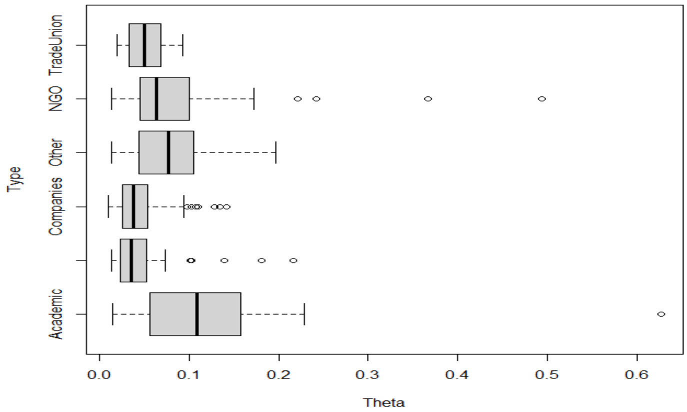


Figure 5. Topic "Fundamental rights"

The topic of "ethical concerns" emphasises issues related to biometric identification (especially face recognition) and the potential consequences and harms for citizens regarding public surveillance. The stakeholders extensively discuss biometric identification and social scoring, racial profiling, and other mass surveillance practices. They emphasise the risks and harms of these practices for the citizen. The terms associated with this topic are: "biometric", "right", "recognition", "fundamental", "identification", "enforcement", "human", "surveillance", "harm", and "public".

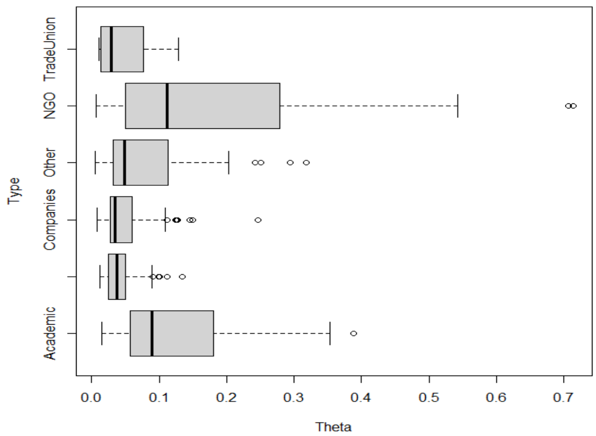
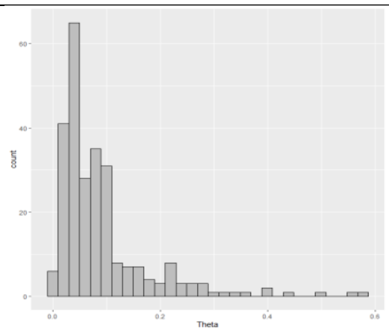
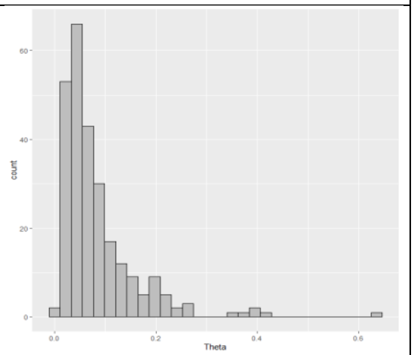
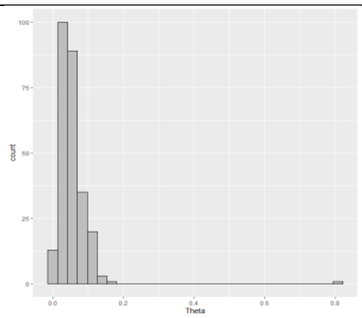
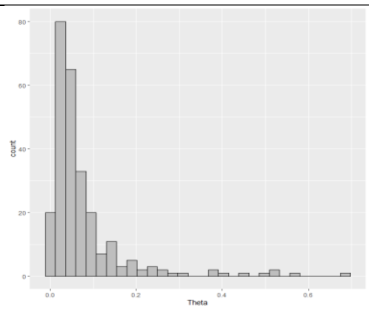
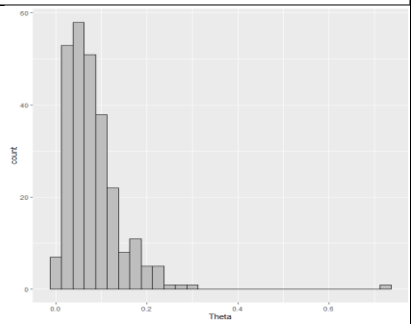


Figure 6. Topic "Ethical concerns"

Our LDA analysis also revealed other topics in the feedback for the proposed AI Act (Table 2). One of these topics concerns AI applications in the finance and banking industry. This topic reveals terms related to statistical techniques and approaches that companies use for business decisions.

Table 2. Other topics

Topic name	Specific terms	Theta distribution
AI in finance and banking industry	1. definition 2. financial 3. credit 4. approaches 5. statistical 6. techniques 7. regulatory 8. application 9. regulated 10. decision	
Health management systems	1. health 2. digital 3. healthcare 4. patients 5. development 6. ethical 7. potential 8. access 9. business 10. trust	
Safety of personal data	1. service 2. provider 3. processing 4. customer 5. GDPR 6. protection 7. application 8. security 9. law 10. legal	

Topic name	Specific terms	Theta distribution
Manufacturing of medical devices	1. device 2. medical 3. software 4. product 5. management 6. safety 7. MDR 8. existing 9. manufacturers 10. notified	
Research and innovation	1. technology 2. research 3. innovation 4. sandboxes 5. regulatory 6. response 7. plan 8. areas 9. key 10. new	

Other themes revealed by our analysis are related to health. These topics are especially representative of organisations from the medical sector. One topic shows medical organisations' concerns about the AI applications developed and used in this sector and their potential consequences for patients (for example, ethical concerns). The other topic focused on health and showed a dialogue related to the manufacturing of AI medical devices. The topic reveals ideas about medical products, devices, and software in relation to the AI Act and other European regulations in the field (such as MDR). Another topic is similar to the one related to fundamental rights, because it reveals worries about the safety of personal data processing and protection in relation to GDPR. And finally, our analysis showed the presence of a discussion topic on the AI Act related to the impact of this regulation on research and innovation. This topic mainly illustrates the inputs on the regulatory sandboxes discussed in the commission's proposal, which aimed at assessing the technologies and, at the same time, increasing innovation.

Some of the topics produced by our analysis are interwoven because many of these topics addressed by stakeholders are also overlapping in the AI Act. At the same time, to some extent, these topics mirror the AI Act initiatives in the topics addressed. Still, they also focus on issues of interest for the types of organisations that gave feedback.

4. Conclusions

The present research aim was to explore the main concerns on artificial intelligence technologies of the stakeholders involved in the AI Act public consultation. To fulfill the objective of our paper, we analysed the feedback provided by 262 stakeholders on the proposal of the European Commission artificial intelligence (AI) Act using a text mining approach based on Latent Dirichlet Allocation. The analysis revealed 12 topics. The topics related to AI applications in industry, transparency, responsibility issues, and AI technologies testing were prevalent. The analysis also revealed topic differences depending on the type of organisation, especially between companies and consumer organisations, NGOs, and academic/research institutions. Consumer organisations were more likely to discuss the lack of skill and education among consumers regarding the use of artificial intelligence. Moreover, academic/research institutions were more likely to raise concerns about fundamental rights and ethical issues.

Acknowledgements*: The study was developed within the NUCLEU Project PN 22_10_0103.*

REFERENCES

[1] **Bareis, J., Katzenbach, C. (2021),** ***Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics***. *Science, Technology & Human Values*, 47(5), 855-881;

[2] **Baruffaldi, S., van Beuzekom, B., Dernis, H., Harhoff, D., Rao, N., Rosenfeld, D., Squicciarini, M., (2020),** ***Identifying and measuring developments in artificial intelligence: Making the impossible possible***. *OECD Science, Technology and Industry Working Papers, No. 2020/05*. OECD Publishing, Paris, https://doi.org/10.1787/5f65ff7e-en;

[3] **Blei, D., Lafferty, J. (2006),** ***Correlated topic models***. *Advances in neural information processing systems, 18,* 147;

[4] **Blei D.M., Ng A.Y., Jordan, M.I. (2003),** ***Latent Dirichlet Allocation***. *Journal of Machine Learning Research*, 3, 993—1022;

[5] **Cao, J., Xia, T., Li, J., Zhang, Y., Tang, S. (2009),** ***A density-based method for adaptive LDA model selection.*** *Neurocomputing*, 72(7-9), 1775-1781;

[6] **Cheatham, B., Javanmardian, K., Samandari, H. (2019),** ***Confronting the risks of artificial intelligence.*** *McKinsey Quarterly*. *McKinsey Global Institute*. (2019);

[7] **Darling, W.M. (2011),** ***A theoretical and practical implementation tutorial on topic modeling and gibbs sampling***. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, 642-647;

[8] **Deveaud, R., SanJuan, E., Bellot, P. (2014),** ***Accurate and effective latent concept modeling for ad hoc information retrieval***. *Document numérique*, *17*(1), 61-84;

[9] **European commission (2021),** ***Proposal for a regulation of the European parliament and of the council: laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts,*** https://eur-lex.europa.eu/legal content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN;

[10] **European Commission (2020), White Paper on Artificial Intelligence - A European approach to excellence and trust**. *European Commission, Brussels*, https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf;

[11] **Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H. (2009).** ***The elements of statistical learning: data mining, inference, and prediction***. Vol. 2, 1-758. New York: Springer;

[12] **Henke, N. et al. (2016),** ***The Age of Analytics: Competing in a Data-Driven World***. *McKinsey Global Institute*;

[13] **Kerr, A., Barry, M., Kellejer, J.D. (2020),** ***Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance***. *Big Data & Society*, 1-12;

[14] **Mazur, J., Wloch, R. (2023),** ***Embedding digital economy: Fictitious triple movement in the European Union's Artificial Intelligence Act.*** *Social & Legal Studies*, https://doi.org/10.1177/09646639231152866;

[15] **Muller, V.C., Bostrom, N. (2016),** ***Future Progress in Artificial Intelligence: A Survey of Expert Opinion***. *In: V.C. Müller (ed.), Fundamental Issues of Artificial Intelligence, Synthese Library 376, Springer International Publishing Switzerland, DOI 10.1007/978-3-319-26485-1_33;*

[16] ***Murphy, K.P. (2012). Machine learning: a probabilistic perspective***. *MIT Press;*

[17] **Nadimpalli, M. (2017),** ***Artificial Intelligence Risks and Benefits***. *International Journal of Innovative Research in Science Engineering and Technology*, 6(6).