

Expert Insights into Studying Vulnerable Communities Online

An Interview with Kyriaki Kalimeri and Yelena Mejova

Publication date: December 24, 2023

One big benefit we derive from digital behavioral data is that they often provide insights into areas that are left blank by more traditional forms of social science data, including sensitive topics like mental health, sexualities or challenging circumstances like communication during wartime or natural disasters. While working with these data offers the opportunity for using ‘data for good’, it also poses methodological and ethical challenges.

We asked Yelena Mejova and Kyriaki Kalimeri, both affiliated with the ISI Foundation in Turin, about their experiences in working with vulnerable communities and studying their online communications. Yelena is a senior research scientist with her main research interests in digital epidemiology and in finding ways to leverage data for social impact. She received her PhD in Computer Science from the University of Iowa. Kyriaki Kalimeri is a research scientist leveraging machine learning and natural language processing to understand decision making from digital data on controversial social issues. She received her PhD in Brain and Cognitive Sciences from the University of Trento and her Diploma in Electrical and Computer Engineering from the Technical University of Crete.

We met Yelena and Kyriaki on June 5, 2023, during the “International Conference on Web and Social Media” (ICWSM-23) in Limassol, Cyprus. The interview was conducted by Leon Fröhling and Indira Sen. The transcript was edited for clarity and length.

Keywords: data ethics, data anonymization, vulnerable communities, humanitarian aid, mental health data, sensitive data, data privacy, reproducibility, machine learning, natural language processing

GESIS: Thanks to both of you for agreeing to this conversation. Our first question is: **What are you working on at the moment, especially with respect to vulnerable communities or sensitive data?**

Yelena Mejova: I have published a couple of papers [1] on the expressions of loneliness online. They range from people just jokingly saying things like “haha, I’m so lonely, someone let’s chat” to real expressions of pain and suicidal ideation. I thought that it would be a small, light-hearted study when I first started looking into it, right before COVID – and then, of course, things got even worse. It was interesting to see that people would turn to social media for alleviating their loneliness and for their self-expression. But it is also interesting to see the lack of support that they get there – at least on Twitter –, how people express themselves, and the kind of responses they get. And even though it is all public data I am very careful in how I write about it. In some sense, especially because we are publishing in communication studies journals, they want to see examples. But before we share any, we make sure to rephrase all of the text of the tweets – that is one safeguard that we try to keep. And, of course, we do not publish the actual datasets.

[The] idea is always the same: there is a population that is in need. These needs can be very different – because of lack of food, because of a natural disaster, or because of simple socio-economic disadvantages or societal stigma. Then we try to bring in our machine learning knowledge and our NLP tools.

Kyriaki Kalimeri: I work on three main research areas. One line of research, very dear to me, is inspired by actual questions that non-governmental organizations – like UNICEF, OCHA [Office for the Coordination of Humanitarian Affairs], the World Food Programme – have [2, 3]. They contact us with specific questions. In this kind of context, we have worked in many different projects, but the idea is always the same: there is a population that is in need. These needs can be very different – because of lack of food, because of a natural disaster, or because of simple socio-economic disadvantages or societal stigma. Then we try to bring in our machine learning knowledge and our NLP [natural language processing] tools, try to generate insights, and thereby give indications to the people that are actually making policies. The other line of research that we follow is machine learning for humanitarian aid. At this point, it is important to define what a vulnerable population is. It depends – and we can never know whether we have the correct definition or refer to the same people. As Yelena said, it could be anyone who is lonely, or it could be a specific group that is isolated somewhere and suffers from something.

Yelena Mejova: Recently we worked with the DEEP platform [4, 5], a tool for humanitarian actors to enter and process their documents. You would think that a simple task like detecting locations would be very basic natural language processing, but we are aware that if we miss a location, we might miss the locations that are most in need. Most detection location tools would never miss New York, but a mention of some camp somewhere will very rarely be caught. A lot of our work therefore is not only building the

tools, but also analyzing them for bias, for coverage, for representativity. Sometimes we might want to favor precision, sometimes recall. What works well is often very context sensitive.

Kyriaki Kalimeri: This happens because most of the available tools were built for commercial purposes and are mostly US- and Europe-centric. If different places have the same name, the one in the US always gets favored – but that is usually not the one that is in need.

GESIS: How did you start working in the field? How did you start in or transition to that area of research?

Yelena Mejova: I think what is interesting to study are controversies. Topics are often controversial because they affect somebody. Abortion in the US affects women, whether or not you vaccinate affects people with compromised immunities. Many things that you think are political are actually affecting people. It is interesting to study questions that are now not decided. There are some questions that we as societies have decided on: children should not starve, and this is not a controversy. The controversies right now are the questions that we are figuring out, they often involve people who cannot necessarily speak or advocate for themselves. I think that is a very interesting way to look at things.

Kyriaki Kalimeri: I am an electronic engineer, with a specialization in telecommunications. Finishing my studies, I felt that even though it was a very interesting domain, I wanted to explore more. I decided to do a PhD in brain and cognitive science, a 180-degree shift from a purely engineering career to something that was all about human psychology and sociology. I was always interested in understanding how people make decisions, which ties very nicely together with what Yelena is saying about controversies. Who we are, what is our background, or our broad personality traits is one thing – but how do we make decisions? We hear the same information but make different decisions. This is what I wanted to study. Coming from a technical background, I focused on how I can do this with social media, with digital data, with telecommunications data.

Yelena Mejova: Using natural language processing.

Kyriaki Kalimeri: Yes, NLP and Machine Learning. Machine Learning is my early career path. When we started collaborating with the UN organizations, I got more into the ‘data for good’ domain, where our contribution is often not as technical. It is not so much about the pure advancement of a technique, of an algorithm, of a platform or of the engineering part in general, but on how to make this all transparent, fair, and usable for people that have lesser knowledge and technologic infrastructure than we do.

GESIS: Since working on these topics, were there any major turning points that changed your perspective on the type of work you do?

Kyriaki Kalimeri: Back when I started, there were no APIs. And if there was one, it was very basic. The only available data were the ones we collected by conducting experiments with sensors. Cell phones were far too expensive for research; they required a lot of funding and budget. I had to go and find my own sensors to do an experiment. I did my PhD on the sociometric badges developed at the MIT [6] kindly offered by Prof. Sandy Pentland; we managed to have 50 participants that was at that time considered a large-scale experiment [7]. Back then, if you were not part of a big company, you had no access to any other type of data. The major turning point was when companies started to give access to their data. We started having more to work with and this was a huge change on what we have been doing.

Yelena Mejova: For me, there were some very drastic turning points. In the beginning I was interested in politics.

But when I got a job in Qatar, I realized I could make a better local impact in the health domain. There, we had a lot of problems, especially with noncommunicable diseases like diabetes or obesity. In that sense, it was just the opportunities around me that turned me towards these behavioral diseases that were talked about quite a lot on social media. But then, slowly, I started drifting back towards politics, and the very basic, very general question of computational social science: “Can we measure something useful in this data?” For a few years I was just asking this question. But now I am asking a slightly higher-level version of that question: “There is something useful here, but how does it actually appear? What are the biases? What are the thought processes? What are the social processes that make us see this type of social data, and that would eventually affect people's health. These turning points for me were always the confluence of different interests, but never just one single thing.

GESIS: Why would you say are we talking about vulnerable communities and their well-being today? Why is there a need to pay special attention to them when doing computational social science research?

Yelena Mejova: Basic ethics.

Kyriaki Kalimeri: These things should be obvious. And not just gender, but everything: ethnicity, religion.

Yelena Mejova: The diversity.

Kyriaki Kalimeri: Exactly, diversity in general.

GESIS: Would you have an answer why that is still the case? Why do we still need to ask such questions to begin with?

Yelena Mejova: It is already a big thing in medicine. For example, in the past drugs were not tested on pregnant women or people from different races. But in the past few decades, these things have come out to be important. It is the same thing in social science. Of course, you cannot lump everyone together. We have to be aware of minorities, of diversities, of diversities within the majorities.

From a computational perspective, the difficulty is to work on overlapping minority attributes.

Intersectionality.

Kyriaki Kalimeri: From a computational perspective, the difficulty is to work on overlapping minority attributes.

Yelena Mejova: Intersectionality.

Kyriaki Kalimeri: That is the most difficult thing from a computational perspective.

Yelena Mejova: Because things get sparse quickly.

Kyriaki Kalimeri: There is also the ethical component to this because vulnerable populations are identifiable in the data.

Yelena Mejova: And this raises privacy concerns.

Kyriaki Kalimeri: Often these are attributes that are not shared or collected. In a previous project about diversity and inclusion in employment in Italy [8], we considered only the gender, because we could not have access to ethnicity or other attributes. I am not even talking about sexual orientation. We were discouraged from asking even about ethnicity, for instance, “Are you open to people that are from different countries?” Things often are a taboo in society, and this propagates also to the analysis that you are doing. For me, our society needs to mature, be much more open, more substantial.

GESIS: You talked about demographics, and gender is kind of the first thing that people think about when they think about biases and de-biasing. And maybe now there is some discussion about ethnicity. But it feels like they formulate ethnicity in a way that is very US-centric. And since you both work on Italian contexts where maybe there are not exactly the same ethnic categories, do you also find challenges with how it is studied currently and how you would like for it to be studied?

Yelena Mejova: In Italy specifically, I have never seen ethnicity considered in datasets. It is not a topic compared to the United States. It is not being tracked.

Kyriaki Kalimeri: Yes, it is not something that I have ever encountered in data.

When you look at, for example, Twitter's terms of service for the data usage, you cannot even try to guess user demographics.

Yelena Mejova: Mostly we do not generate our own data. We scrape it, it is generated by somebody else. It is the design of those systems that predetermines what we can do. For example, Facebook does have an elaborate selection system for gender, i.e., there were more than two selections [9]. If this was surfaced somehow to researchers, we would be able to work on it. But if there is not even a button for it, a selection, then there is little we can do. When you look at, for example, Twitter's terms of service for the data usage, you cannot even try to guess user demographics. Of course, there is a good reason for it – privacy issues. But if you are not able to target a specific group, we cannot study them either. So, we have to find the middle ground with protection on one side. And especially if people choose to reveal such information about themselves, if they are okay with it, then we should support it.

Kyriaki Kalimeri: Sometimes they do so to advocate for certain things that are important to them.

Yelena Mejova: Just like in this talk, in our workshop [10] on people disclosing their preferred pronouns. And more and more people do it including heterosexual cisgender people. So, it is becoming a norm and it should be at least not stifled.

GESIS: Moving to some of your more recent work, “Authority without Care: Moral Values behind the Mask Mandate Response” [11] which is also going to be presented at this conference [ICWSM-23] [12]. Maybe we could use this as an example for you to explain how you approach such a kind of project and how you in the end derive policy implications from that. How do you get from asking a research question and answering it to drawing wider conclusions that could potentially inform policy making?

Kyriaki Kalimeri: It is not a linear process. We started with much more complex research questions about prosocial behaviors, and it is a continuous back and forth of what kind of data we can have access to and what kind of questions we can answer with this data. We usually start with really cool research questions but then we need to adjust them according to the data availability. We cannot have the holistic view of society that we would like to have so our final projects are an interplay between important research questions and solid answers we can provide.

Especially in the humanitarian research, there is a very high percentage of cool things that we would love to do but we have no data; it is not unlikely that the communities we work with have no access to internet for instance. We can only have access to these people through humanitarian experts in the field. There are different kinds of implications. From our side we are always trying to connect our research to questions that reflect actionable insights about our society with a potential positive impact.

Yelena Mejova: Of course, it does not come out of nothing. Often, we already know some surveys or articles. The masking issue was obviously very politicized.

Kyriaki Kalimeri: The research questions are often driven by our intuition.

Yelena Mejova: In this particular case I remember even some Pew Research Center articles [13] talking about masking. In that sense it was a straightforward connection. But I was really thinking about it politically because that is my interest. But then Kyriaki thought to look at their values and it goes beyond just politics.

Kyriaki Kalimeri: Yes, because I was interested in how people reason about this. It is not just “I wear a mask” versus “I do not wear a mask”. But it was also the whole reasoning of a simple behavior which turned into a political stance. It was like an indication of group membership rather than being an action that you take towards a public health issue. It was more like “I stand with this or that politician” and an emergence of how you build communities and the resistance in society through a very simple, non-invasive act. We observed the evolution of this transformation while it was happening. Vaccination, for instance, is an invasive intervention; I understand people’s concerns to vaccinate their children or themselves. It is something that they are afraid of for a series of reasons. Wearing a mask, on the other hand, has absolutely no impact on your life. It is just not elegant and perhaps you cannot breathe as easily as you would breathe without it, however there is absolutely no risk, and yet the conversation was so heated.

Yelena Mejova: Yes! Somehow people had strong feelings about it. Where did they come from?

Kyriaki Kalimeri: They were really arguing about something that does not cost you much and it does not harm you in any way.

We can come up with all these cool algorithms, but they are not reaching their potential if they cannot be used by people without technical know-how.

GESIS: *Maybe shifting to a slightly different topic which is that in this research community with work on vulnerable communities and their wellbeing, do you see any common misconceptions or misperceptions?*

Yelena Mejova: There is definitely a disconnect between the engineers and the actual policy implementers. Just as we have had all of our wonderful talks and academics and graduate students presenting their work, towards the very end, one of the speakers said: “If you don't have funding for it, how are we supposed to do anything?” If the tools are not usable it is a big problem. We can come up with all these cool algorithms, but they are not reaching their potential if they cannot be used by people without technical know-how. There is a lot more non-technical work that has to happen before this. Unfortunately, it is

really hard to get it into the hands of people who are implementing it on the ground. They are usually very busy and skeptical about the future.

Kyriaki Kalimeri: Yes, there are misconceptions on both sides. Sometimes policymakers think we do magic. It is also a responsibility from our side to communicate things in a way that is widely understood, not just graphs, values, and formulas, but render the concepts understandable to people from totally different backgrounds. The same way I would not understand anything if I went to a lawyers' conference. You have to be able to adjust the communication of what you do to your audience's background. That is an important skill to teach younger researchers, the more interdisciplinary the field becomes. When I was at the university, we were all telecommunication engineers with a common language, but then as a visiting PhD scholar at MIT, Sandy Pentland taught me the importance of proper and thorough communication of science. This was a great lesson for me.

What do you call a 'vulnerable population'?

We need to communicate with people that come from different backgrounds, and we need to make sure our messages are not misinterpreted. Often because the same word means different things in different disciplines. Our latest workshop was on vulnerable populations. A highly esteemed collaborator from UNICEF argued: "What do you call a 'vulnerable population'? And who is that you mean that is vulnerable?" – A whole super interesting discussion on what is vulnerability and who is vulnerable arose. Us, as computer scientists and engineers, never gave that much thought on it. Yet, it was a very important discussion which broadened our understanding and perception on the field.

Yelena Mejova: This is extremely important now with large language models. The difficulty is that they are so easy to use, and they are very useful in many ways, but what they lack is being deterministic, having an actually good predictability. It is this sort of uncertainty that is difficult to communicate. A difficulty we had during COVID, when people noticed we were wrong. We just did not know, we were guessing. Now we know. It is the nature of the thing, and it is hard to communicate to people that I am concerned that tools will be used in ways that they are not suitable for.

GESIS: That is a great takeaway here, that it is not only important to solve a task or answer a question, but also to communicate it effectively and responsibly and make sure that it reaches people.

Kyriaki Kalimeri: Because this is what in the end creates all the misconceptions. If you have the slightest doubt about what you are hearing, you are going to be very skeptical. Transparency and digital literacy are very important in all disciplines. Science, journalism, policy making. As things move more towards generative AI, it is going to be really hard to distinguish what is truth and what is not truth and how much of a truth is in there.

Yelena Mejova: Although we are already moving towards that. Many schools and educational programs have digital literacy classes and hopefully they will incorporate a little bit of the latest technologies in there as well.

GESIS: Yelena, you touched upon that before, when you talked about some of the steps that you would do before publishing data, like rephrasing it, for example. Maybe you could reiterate on that and give some practical advice to researchers who want to do responsible research in vulnerable communities. What are some of the things that you have to care for with that type of research?

Yelena Mejova: I have not been a part of universities. Most of my research has not gone through the institutional review board. And in some places, they would even say that you do not have to submit your work because it is public data and you do not interact directly with the participants. But especially when you start talking about people who post personal things like health, you realize that the right to be forgotten does not exist anymore when you solidify it into a dataset. We have published some annotated datasets of political, general topics that are not necessarily very personal. But with the personal ones, it is hard. Usually, we have the content but not the ID or something similar, or as Twitter suggests [read now: as Twitter *used to suggest*], we release the tweet IDs and then the tweets have to be recollected, but these might be unavailable because they have been deleted. This limits the reproducibility of our research. There is the usual tension between keeping people safe and not disclosing too much information. In that sense having information brokers who would keep data and verify who is accessing it would be better for research. But the question of how to organize this would come up.

GESIS: Who is going to make the decisions?

Yelena Mejova: This is never going to happen. But for example, when I was working on a paper on abortion [14], it was very difficult to find a list of abortion clinics in the US. I am guessing, they do not want to be found that easily. But it is a tension – they should be found by somebody if somebody wants to use them. I had to go to a research group that collected a list, but they did not want to publish it. After I signed a contract that I would not use it in an unethical way I received this list, and it was very helpful. But at this point, it seems like the researchers themselves have to police how certain information gets released. This is something I will be speaking about a little bit during the panel on ethics at this conference [15]. A very nice idea that is, however, difficult to implement. Think of datasets in post publication, there should be some way to encourage people to shepherd the data afterwards. We should be able to track it. Who uses it and how? What is the metadata for it like? Is the licensing appropriate? Should it be changed some time down the line? All of these things, it would be great if the authors themselves were keeping track of it. But, of course, they do not because there is no reward in it, and it is really hard.

Kyriaki Kalimeri: And it becomes expensive because you need space, you need storage, you need to monitor these things.

Yelena Mejova: In a perfect world we would do that. But we need infrastructure, we need reward systems, we need many things to encourage us to do it. It still could be improved a lot.

There is always this trade-off between reproducibility and ethical concerns.

GESIS: Kyriaki, would you also like to talk about some of the safeguards that you take when working on this type of research?

Kyriaki Kalimeri: In the past I worked also with biomedical information, even more sensitive data. We had to ensure the integrity, confidentiality, and of course data availability. Even if the server was safe enough, there are many things to be careful with when you are dealing with such sensitive information. In those studies, of course we obtained ethical approval, and provided informed consent, however still we need to be cautious. Social media data are less sensitive, but still, we need to anonymize, and make sure all our collaborators are aware of the proper use of them. A tricky area is data reproducibility. As Yelena said, to reproduce exactly our findings, other researchers need to have access to the raw data. There is always this trade-off between reproducibility and ethical concerns.

Yelena Mejova: We do not deal with that much data that something like differential privacy could help with. A lot of things we are measuring are on a very personal level. I am not an expert on it, but there are huge datasets that Facebook publishes. Some of the things I have done cover at least a few thousand people in a network somewhere. We could have tried applying some more sophisticated privacy-preserving things and then releasing them. But this would be a paper in itself.

Kyriaki Kalimeri: Exactly, if you have access only to the aggregated data, usually you cannot reproduce the same result of the paper. The solution is not as easy; even with anonymization, and I am referring to proper sociodemographic and behavioral anonymization not just hashing the name or ID, often reproducibility is impacted too.

Yelena Mejova: It could even be pointless.

Kyriaki Kalimeri: Hence often is a blind, “trust me, I know what I am doing” kind of situation. This is why I personally believe in trusting partnerships between academia and industry.

GESIS: That is a great summary of some of the tensions working with sensitive data and vulnerable communities.

*The super important thing is to be transparent about everything.
You share the data; you say what data you share.
You scrape, you say what you have scraped.*

Kyriaki Kalimeri: The super important thing is to be transparent about everything. You share the data; you say what data you share. You scrape, you say what you have scraped. Say how you used it and why. When the Cambridge Analytica story [16] came out, people had no idea how their data were being treated and of course it was a scandal. Being upfront with the data, methods, but also objectives, we can create a relationship of trust with the broader audience. This comes back to how much people trust science and how we communicate. From my past experience, if you are transparent and you communicate properly your ideas and objectives people are willing to collaborate and contribute their data. It takes of course much more time and effort.

GESIS: This relationship between transparency and trust – the lack of it can lead to erosion of trust.

Kyriaki Kalimeri: While nurturing it will allow us to advance science, because otherwise it will be a black box and people do not know whether they could trust you or not.

GESIS: For these research topics, if someone was getting started, what are the type of resources that you would suggest to them in terms of tutorials or workshops? We will obviously include your workshop here, but are there are other handbooks, papers, any courses that you think are interesting?

Kyriaki Kalimeri: Well, there are tutorials about data mining, natural language processing, machine learning. Together with Yelena we have also given two tutorials on the topic [17]. Then, there is plenty of information on social sciences, too.

Yelena Mejova: Most of the students might find “The Missing Semester of Your CS education” [18] helpful. How do you use Linux terminal? How do you do the very basics? Often people come from many different places. This one is very basic but essential. I think there is a very nice paper on data annotation “Grounded theory” [19]. In computing we are very used to classification and annotation for classification, but sometimes you do need to understand what is in the data. So, the 'grounded theory' thing is very important. There is a very nice overview of how to do it.

GESIS: If you were able to make a wish to the universe for something that would help you with your research on vulnerable communities. What would that be?

*If you were able to make a wish ...
Data!
[... and a] data broker.*

Kyriaki Kalimeri: Data!

Yelena Mejova: The thing that I mentioned would be nice. The thing about a data broker. Because it is ambitious, but it is not impossible. You could put the sages of some discipline together, or even many disciplines, or even pay them as long as they would give you some kind of access to particular datasets. Right now, it is about connections.

For example, if you just extend general data protection regulations and individuals could agree to leave their data open for research.

GESIS: Democratizing data access. I think that is a really nice note to end our interview on. Thank you!

References

- 1 For instance: Mejova, Y., Weber, I., & Fernandez-Luque, L. (2018). Online health monitoring using facebook advertisement audience estimates in the United States: Evaluation study. *JMIR Public Health and Surveillance*, 4(1), e7217. <https://doi.org/10.2196/publichealth.7217>
- 2 [with UNICEF] Sartirano, D., Kalimeri, K., Cattuto, C., Delamónica, E., Garcia-Herranz, M., Mockler, A., Paolotti, D., & Schifanella, R., 2023. Strengths and limitations of relative wealth indices derived from big data in Indonesia. *Frontiers in Big Data*, 6, 10.
- 3 [with OCHA] Forooshani, M.K., van den Homberg, M., Kalimeri, K., Kaltenbrunner, A., Mejova, Y., Milano, L., Ndirangu, P., Paolotti, D., Teklesadik, A., & Turner, M.L., Towards a global impact-based forecasting model for tropical cyclones. *EGUsphere*, 2023, 1-29.
- 4 DEEP – a collaborative analysis platform for effective aid responses (s.d.). <https://thedeep.io/> [retrieved Nov 1, 2023]
- 5 [with DataFriendlySpace/DEEP] Belliaro, E., Kalimeri, K. and Mejova, Y. (2023). Leave no Place Behind: Improved Geolocation in Humanitarian Documents. *Proceedings of the 2023 ACM Conference on Information Technology for Social Good* (pp. 31-39).
- 6 Kalimeri, K. (2013). *Traits, States and Situations: Automatic Prediction of Personality and Situations from Actual Behavior*. PhD Thesis, University of Trento. <https://doi.org/10.13140/2.1.2847.9043>
- 7 Lepri, B., Staiano, J., Rigato, G., Kalimeri, K., Finnerty, A., Pianesi, F., ... & Pentland, A. (2012). The sociometric badges corpus: A multilevel behavioral dataset for social behavior in complex organizations. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing* (pp. 623-628). IEEE. <https://doi.org/10.1109/SocialCom-PASSAT.2012.71>

- 8 Urbinati, A., Kalimeri, K., Bonanomi, A., Rosina, A., Cattuto, C., & Paolotti, D. (2020). Young adult unemployment through the lens of social media: Italy as a case study. *International Conference on Social Informatics* (pp. 380-396). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-60975-7_28
- 9 Goldman, R. (Feb 13, 2014). Here's a List of 58 Gender Options for Facebook Users. *ABC News*. <https://abcnews.go.com/blogs/headlines/2014/02/heres-a-list-of-58-gender-options-for-facebook-users> [retrieved Nov 1, 2023]
- 10 4th Workshop on the Data for the Wellbeing of the Most vulnerable. At the [17th] International AAAI Conference on Web and Social Media (ICWSM). June 5, 2023. <https://sites.google.com/view/dataforvulnerable23> [retrieved Nov 1, 2023]
- 11 Mejova, Y., Kalimeri, K., & Morales, G. D. F. (2023). Authority without Care: Moral Values behind the Mask Mandate Response. *ArXiv*. <https://doi.org/10.48550/ARXIV.2303.12014>
- 12 The 17th International AAAI Conference on Web and Social Media (ICWSM) (2023). <https://www.icwsm.org/2023/index.html/> [retrieved Nov 1, 2023]
- 13 Kramer, S. (Aug 27, 2020). More Americans say they are regularly wearing masks in stores and other businesses. Pew Research Center. <https://www.pewresearch.org/short-reads/2020/08/27/more-americans-say-they-are-regularly-wearing-masks-in-stores-and-other-businesses/> [retrieved Nov 1, 2023]
- 14 Mejova, Y., Gracyk, T., & Robertson, R. E. (2022). Googling for Abortion: Search Engine Mediation of Abortion Accessibility in the United States. *Journal of Quantitative Description: Digital Media*, 2. <https://doi.org/10.51685/jqd.2022.007>
- 15 Panel Discussion at 17th International AAAI Conference on Web and Social Media (ICWSM) (2023): Computational Social Science Research & Data Ethics. <https://www.icwsm.org/2023/index.html/> [retrieved Nov 1, 2023]
- 16 Confessore, N. (April 4, 2018). Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. *The New York Times*. <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html> [retrieved Nov 1, 2023]
- 17 2019: Modeling Human Values with Social Media, <https://sites.google.com/site/ic2s2humanvalues/> [retrieved Nov 18, 2023]; slides: <https://www.slideshare.net/kkalimeri/ic2s2-tutorial-on-modeling-human-values-via-social-media-data> [retrieved Nov 18, 2023]; 2023: <https://propaganda.math.unipd.it/ic2s2-tutorial/> [retrieved Nov 18, 2023]
- 18 The Missing Semester of Your CS Education (s.d.). <https://missing.csail.mit.edu/> [retrieved Nov 1, 2023]
- 19 Nelson, L. K. (2020). Computational Grounded Theory: A Methodological Framework. *Sociological Methods & Research*, 49(1), 3-42. <https://doi.org/10.1177/0049124117729703>

Suggested citation

Kalimeri, K. & Mejova, Y. (2023). *Expert Insights into Studying Vulnerable Communities Online. An Interview with Kyriaki Kalimeri and Yelena Mejova* (GESIS Guides to Digital Behavioral Data, 4). Cologne: GESIS – Leibniz Institute for the Social Sciences.

Series editors

Maria Zens, Katrin Weller, Claudia Wagner

Publisher

GESIS Leibniz-Institute for the Social Sciences, Cologne

License

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)