etnološka tribina 48, vol. 55, 2025., str. 55-77

DOI: 10.15378/1848-9540.2025.48.01 original scientific paper, submitted 3. 8. 2025., accepted 31. 8. 2025.

BEING OPEN ABOUT OPEN DATA FROM THE PERSPECTIVE OF ETHNOLOGY AND CULTURAL ANTHROPOLOGY

Olga Orlić
Institute of Ethnology and Folklore Research, Zagreb
orcid.org/0000-0001-9189-5732
olga@ief.hr

Implementing the FAIR principles for data sharing introduces new expectations for the broader academic community. These principles advocate that data should be "as open as possible, and as closed as necessary", acknowledging that not all data requires the same level of accessibility. From the standpoint of ethnologists and cultural anthropologists, however, the push for open data raises unique concerns. This paper explores these challenges, particularly those related to the distinctive nature of cognitive processes and data types within ethnology and cultural anthropology. The article presents perspectives from Croatian ethnologists and cultural anthropologists, highlighting our shared concern about safeguarding personal and sensitive information provided by interlocutors – especially when such data could be linked with published ethnographic and cultural-anthropological material.

Keywords: open science, open data, ethnology and cultural anthropology

> *I believe that sharing raw data should depend on the sensitivity of the topic and the agreement with the interlocutor.*
> (response to an open-ended question in an online survey on the attitudes of ethnologists and cultural anthropologists in Croatia toward open data)

Open science

Open science and the advocacy for the availability and exchange of knowledge and open access to research results (especially those coming from public funding) have been met with near-unanimous approval from the academic community, including in Croatia (Stojanovski 2019). There are various approaches to open science, as well as multiple definitions (cf. Marić, Dumančić and Šarić 2023). UNESCO defined it as such in 2021:

> inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of

> scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community. It comprises all scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences and the humanities, and it builds on the following key pillars: open scientific knowledge, open science infrastructures, science communication, open engagement of societal actors and open dialogue with other knowledge systems. (UNESCO Recommendation on Open Science 2021: 7)

The process of opening up science intensified after the Budapest declaration in 2002[1] and the Berlin declaration in 2003[2] (Marić, Dumančić and Šarić 2023: 122). Some aspects of open science have become so widespread that, for example, open access is now the most important segment of standard academic publishing. A significant outcome of open access is the removal of inequalities in access to research results with closed access, i.e., those available only through payment. Open access enables users to access research results and educational materials freely and quickly. To defend against the negative aspects of open access, such as the emergence of predatory journals, some scholars believe that it is essential to maintain clear, explicit, and conservative standards of academic evaluation (Miller 2012: 386). Fortunately, Croatia has a long-standing tradition of publishing scientific journals within institutions and professional communities, funded by taxpayers' money. This nonprofit model supports a diamond/platinum open access model, in which neither authors nor readers bear any costs. This commitment to a nonprofit model led, already in 2006, to the establishment of the Portal of Croatian Scientific and Professional Journals (*Hrčak,* www.hrcak.hr), for Croatian journals in open access, contributing to Croatia's high position in the European Commission's Open Access Monitor (Stojanovski 2019: 167–168). A new model of funding scientific journals through program agreements with institutions has raised concerns among editorial boards, who believe that receiving funds through a direct single call for proposals is a better solution (ibid.: 170). Besides state funding, the diamond model of open access is also made possible through voluntary editorial and peer-review work (Stojanovski and Mofardin 2025). In the case of open access to monographs and books, Jadranka Stojanovski and Danijel Mofardin argue that the process is progressing more slowly and with much greater caution.[3]

1 Budapest Open Access Initiative.

2 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.

3 The Croatian Ethnological Society (HED) embraced the opportunities offered by open access to monographs and books early on. In 2013, it launched the hed-biblioteka a "publishing series of digital publications in the area of ethnology, cultural anthropology, folklore studies and other related disciplines in the humanities and social sciences, while maintaining all existing standards for publication". Depending on funding received from the Ministry of Science, Education and Youth, some publications are also issued in print; but the primary goal of the hed-biblioteka is electronic publishing, free of charge for both readers and authors. (Hrvatsko etnološko društvo, https://hrvatskoetnoloskodrustvo.hr/o-hed-biblioteci/)

However, open science also advocates and entails openness in data, methodology, evaluation, and educational materials (Marić, Dumančić and Šimunović 2023: 122), as well as citizen science, which aims to involve the public in academic research processes (Ivanjko, Zlodi and Horvat 2024). The focus of open science advocates varies, as Benedikt Fecher and Sascha Friesike have identified five "schools of open science: the democratic, public, pragmatic, infrastructure, and measurement schools" (Fecher and Friesike 2014, as cited in Vlašiček and Flis 2021: 509). The democratic and public schools advocate for allowing access to all scientific processes and results; the pragmatic and infrastructure schools focus on facilitating scientific collaboration aiming to increase the efficiency of scientific discovery (pragmatic) or on developing software solutions necessary to enable open science in a technical sense (infrastructure). The measurement school aims to reform the existing method of "evaluating the quality of scientific research and researchers themselves," advocating a shift away from traditional metrics of academic productivity, such as citations, the h-index, or impact factor, and instead to focus on developing alternative metrics (Priem et al. 2010) that would complement existing ones (Vlašiček and Flis 2021: 507–508). The ideas and motivations behind these efforts are positive and beneficial, as they enhance transparency and the democratization of all processes involved in scientific research and education.

Openness in evaluation focuses on the development of alternative standards of assessment, such as the FOSE standard (Framework for Open Science Evaluation) (Walther and van der Bosch 2012). While proponents of such evaluation emphasize the positive aspects of open review (Kriegeskorte 2012), it is also important not to overlook its potential limitations (van Rooyen et al. 1999). As far as I am aware, such processes have not yet been implemented within the Croatian academic community, despite being advocated (Ukić 2021), so it remains to be seen how they would be applied and what their long-term impact might be.

Šime Ukić (2022: 77) notes the connection between career advancement in academia and the use of traditional metrical tools for evaluation. He interprets researchers' reluctance to share data as stemming from their desire to maintain control over publication, as well as from the lack of incentives for data sharing. Among other things, he mentions the reform of the research evaluation system promoted by the European Commission, which calls for a comprehensive evaluation of researchers and their work, while respecting the diversity of academic disciplines (Cabello Valdes et al. 2017).

I will address the issue of data openness from the perspective of ethnologists and cultural anthropologists in more detail in the remainder of the text, particularly with regard to the confidentiality and the need to protect personal and sensitive data obtained through ethnographic research. I consider this issue important precisely because comparing published texts and research results in small communities with data made available in open access – even when anonymized – could sometimes lead to the unintended identification of interlocutors.

The replication crisis and open data: FAIR and CARE principles

Some authors point out that the so-called replication crisis is more appropriately described as a crisis of trust in scientific results (Baker 2016 in Flis 2023: 227), which emerged when it became clear that a large number of published results from quantitative studies and experiments could not be replicated or verified (Romero 2019: 3; Vlašiček and Flis 2021: 510). Felipe Romero emphasizes that underlying this crisis is the assumption that the very possibility of replicating research results is a fundamental characteristic of science (Romero 2019: 1). The crisis began in 2010, when the inability to replicate certain influential research studies was observed, although some scientists had already been concerned since the 1960s and 1970s about the lack of both implementation and publication of such replication studies (ibid.: 2). Romero provides a detailed analysis of the reasons for the shortage of such studies and explains the proposed reforms – statistical, methodological, and social. One of the solutions was offered by the open science movement through data sharing, which, in addition to enabling verification and comparison, would also provide greater transparency of research processes, accelerate new discoveries, and contribute to cost savings. The demand for greater transparency in research processes spilled over into all scientific disciplines. In 2016, the FAIR data principles were formulated (Wilkinson et al. 2016). The FAIR principles assume that data are Findable, Accessible, Interoperable, and Reusable (ibid.), and the ways of designing and managing data according to these principles have also been thoroughly elaborated and evaluated (e.g., Mons et al. 2017; Celjak et al. 2020; Jacobsen et al. 2020; Landi et al. 2020). Some researchers involved in the process of opening data are primarily focused on data protection, due to reasons such as personal privacy, national security, but competitiveness as well (Brewster et al. 2020). Others draw attention to the fact that some studies collect only sensitive personal data, such as video materials, emphasizing that sometimes the data only have value in that specific form (e.g., in research on sign language communication (Kolbe 2022)). Ethnology and cultural anthropology are a part of the disciplines concerned with the process of data opening.

It is important to keep in mind that data openness according to FAIR principles does not necessarily mean full disclosure, but rather availability under "well-defined conditions" (Landi et al. 2020). This principle is also included in the EU Horizon 2020 program's open science guidelines, which state that data should be "as open as possible, as closed as necessary." Barend Mons et collaborators outlined different levels of data openness in relation to the FAIR principles, ranging from fully open to fully closed (Figure 1, Mons et al. 2017: 53).

Thus, Mons et al. (2020: 1) highlight that

> FAIR is not a new standard; is not a top down requirement; is not an all-or-nothing binary state (FAIR or not FAIR). The FAIR principles were conceived and designed as a resource for optimal choices to be made during many aspects of data and tool generation as well as (re)use and long term stewardship.

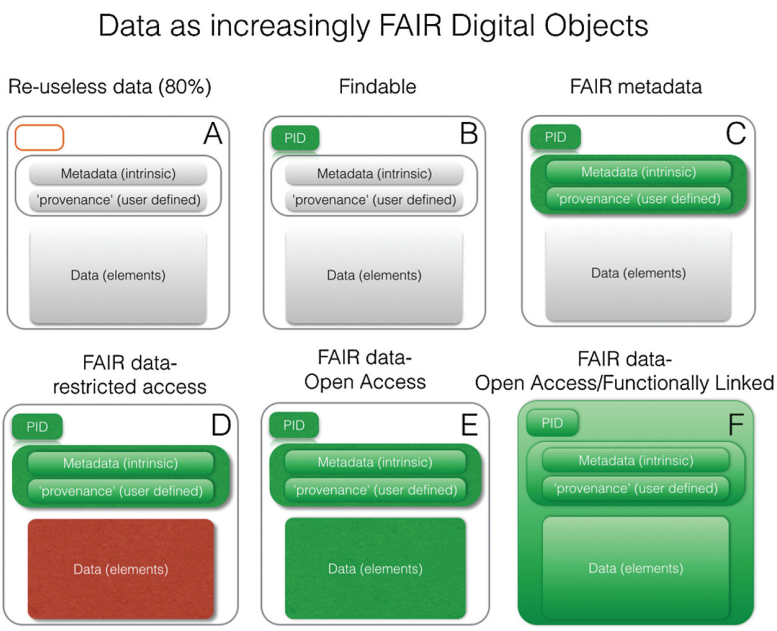

Fig. 1. Varying degrees of FAIRness. As elements become coloured, they have become FAIR. For example, adding a persistent identifier (PID) increases the fairness of that component. Coloured elements in green are FAIR and open, coloured elements in red are FAIR and closed. In the final panel, the mechanism for expressing the relationship between the ID, the metadata, and the data, is also FAIR (i.e. follows a widely accepted and machine-readable standard, such as DCAT or NanoPublications) and interlinked with other related FAIR data or analytical tools on the Internet of FAIR Data and Services.

Figure 1. Degrees of FAIR principles, taken from Mons et al. 2017: 53)

Carol Tenopir et al. (2011) observed a gap between theory and practice when it comes to open data, between the expressed willingness to share data and what actually occurs in practice, which happens only upon request, if at all. Their research also concluded that the type of data to be shared influences the willingness to share it. For example, as many as 90% of scientists in the atmospheric sciences[4] and 85% in the biological sciences are willing to share their data, compared to 65% of scientists in the medical sciences and 58% of those in the social sciences (Tenopir 2011: 12). Heinz Pampel and Sünje Dallmeier-Tiessen (2014) further elaborated on this reluctance. The creation of additional documentation required to prepare data for publication (in the case of personal or sensitive data this includes not only anonymization – an irreversible process – or pseudonymization – a reversible process – but an entire procedure aimed at preventing the identification of research participants) is a demanding, time-consuming, and expensive process that researchers must undertake alongside a range of other activities and priorities (publishing, conducting research,

4 Atmospheric science is an interdisciplinary field that integrates aspects of physics and chemistry focused on the structure and dynamics of Earth's atmosphere. Mathematical tools such as differential equations and vector analysis, as well as computational systems, are used to study the physical and chemical relationships that describe the functioning of the atmosphere. Atmospheric sciences are traditionally divided into three thematic areas – meteorology (the study and forecasting of weather), climatology (the study of long-term atmospheric patterns and their effects), and aeronomy (the study of the physics and chemistry of the upper atmosphere). https://www.britannica.com/science/atmospheric-science

securing funding, administration). Additionally, the authors note, some disciplines lack standardized methods for data description, and they argue that researchers lack sufficient incentives to open their data (ibid.). Some studies have pointed to the impossibility of repositories fully aligning with the FAIR principles (Burrows 2011), while others have highlighted the paradox of the "selfish scientist" among researchers, namely in the life sciences where a group of scientists readily uses open data but is reluctant to share their own (Damalas et al. 2018). This study conducted at the international level showed that, e.g., early-career researchers are hesitant to share data before publication (although they are willing to do so if it would mean publication in high-ranking journals), while more experienced researchers expressed general reluctance to share data, except when sharing could lead to funding or collaboration in specific circumstances (ibid.).

The results of the EOCD Pilot, conducted by the European Commission, showed that a significant number of researchers involved in HORIZON 2020 projects withdrew from data sharing (Cabello Valdes et al. 2017: 11).

Open science has stimulated the development of "New tools, methods, and infrastructures [...] for the dissemination, processing, analysis, and preservation of research data" (Schöpfel et al. 2019: 623). It seeks to encourage researchers to share their data through various initiatives, one of which is the publication of so-called data papers. A data paper is defined as

> a searchable metadata document that describes a particular dataset or a group of datasets, published as a peer-reviewed article in a scientific journal. Unlike traditional scientific articles, which focus on research results, data papers aim to provide a detailed description of the data in order to promote their reuse, transparency, and reproducibility in scientific research. (Data Papers, Global Biodiversity Information Facility – GBIF)

Although such initiatives are not inherently problematic, it should be kept in mind, as Pampel and Dallmeier-Tiessen (2014) point out, that in a competitive research environment scientists are still primarily evaluated based on the publication of articles in journals or books. Some editors of Croatian scientific journals have made efforts to allow for more transparent presentation of research processes in articles, and, where possible, for the publication of data papers, for example by increasing the allowed word count. At the same time, they are aware that, due to the "specificities of different epistemological approaches," it is impossible to require all researchers to make their data openly accessible (Vučković Juroš 2021: 305).

Due to the frequent misuse of data obtained from marginalized and vulnerable groups, participants at a workshop in Gaborone, Botswana – led by Indigenous communities – developed a draft of the CARE Principles for Indigenous Data Governance. These principles were created by assembled Indigenous and allied academics and practitioners to ensure that Indigenous communities retain control over data governance and to support the co-creation of knowledge on an equal footing (Carroll et al. 2020). The CARE acronym stands for Collective benefit, Authority to

control, Responsibility, and Ethics. These principles apply to "data, information, and knowledge, in any format, that impact Indigenous Peoples, nations, and communities at the collective and individual levels; data about their resources and environments, data about them as Individuals, and data about them as collectives" (Carroll et al. 2021: 1). The same authors proposed a model of open data for Indigenous communities with the application of the FAIR and CARE principles (ibid.).

Figure 2. Taken from Carroll. et al. (2021: 2).

Stephanie Russo Carroll, Duarte Marisa and Max Liboiron (2024) note that the CARE principles, due to their deep roots in Indigenous philosophy and worldview, cannot simply be transferred to any other context, particularly because of the colonial and extractive relationship toward the knowledge of Indigenous communities; as such, they represent a distinct form of protection.

I believe, however, that the CARE principles can also be applied to most interlocutors or communities involved in research (Orlić 2024), regardless of whether they are indigenous or not. After all, "anthropology at home" has long been a common approach in anthropological research, and interlocutors are not always marginalized or vulnerable, since research in the field of cultural anthropology has long been conducted in all "directions" (Hannerz 2006). Regardless of who the research participants are and what their social or economic status may be, I believe that the very fact that research analyses based on such data are most often published creates an unequal relationship in the context of data governance. There is an attempt to mitigate this inequality through collaborative practices, though these are still not

the standard. Therefore, informed consent remains a crucial mechanism that allows interlocutors to retain some control over the management of the data they have provided.

If the interlocutor understands all aspects and possibilities related to data openness, then the only ethical solution is that they alone define the criteria for access to those data. After all, this is a right afforded to them by the General Data Protection Regulation (GDPR) and all the principles laid out in the Data Act[5]. According to the Data Act, adopted with the aim of creating and regulating a European data market for data obtained from smart devices and primarily concerning data collection in economic activities (although certain areas of science could also benefit from such data), it is clear that the principles of the GDPR take precedence. If this means that the data should never be shared, then both ethical principles and laws enable it. If it means, for instance, that the researcher cannot make certain data open, this decision must be considered justified. Even those that fund projects, such as the Croatian Science Foundation (HRZZ), are aware of the issue of data confidentiality and the regulations protecting personal and sensitive data. In 2022, HRZZ introduced a "mandatory Data Management Plan for all projects contracted since 2019, and as of 2022, Data Management Plans have become a required part of the tender documentation when submitting projects" (Strateški plan Hrvatske zaklade za znanost za razdoblje 2023–2027). On the HRZZ website, a sample of such a plan for the social sciences and humanities has been published, providing examples of various data management methods. One example suggests a time-limited embargo on data access, another calls for the destruction of audio recordings and the storage of only anonymized transcripts, with a certain level of anonymization, and so on. Repositories available within Croatia's open access scientific infrastructure (DABAR, PUH, CROSSDA) allow for such data to be stored in accordance with data usage rights.

The arguments justifying the need for open data are based on the assumption that such openness is beneficial, as it potentially enables data verifiability, facilitates comparisons, and may increase the efficiency of research data, thereby advancing science and accelerating discoveries (Stojanovski 2019). This logic does indeed seem justified when it comes to quantitative data concerning natural or social phenomena (e.g., sea temperature measurements or the number of overnight tourist stays in a certain location by month), i.e. when the results of the analysis do not depend (as much!) on the researcher's interpretation. Verifiability in the context of quantified data makes sense, but what about data collected in contexts that cannot be replicated nor verified in the same way, or at all? In the case of qualitative research, personal and confidential data are collected, so their analysis and presentation are fundamentally different from those based on quantitative data.

However, ethnographic data are more than just qualitative data, in addition to being more complex. Research in the humanities, in particular, "produces and makes

5 Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

use of other kinds of evidence which are more difficult to define and categorise, and which do not fit readily into these quantitative and qualitative frameworks of the sciences and social sciences" (Burrows 2011: 176).

The requirement for data sharing has already sparked debates within certain ethnological journals. For example, in the Forum section of the journal *Ethnography*, Robert Pool (2017: 282) points out that before making demands for open data, it is necessary to clarify what constitutes ethnographic data and what is meant by verification. Pool divides ethnographic data into so-called hard and soft data, where documents and objects would be considered hard, and the researcher's recollections would fall under soft data. Ethnographic data, therefore, have a hybrid character, and these two types of data are in constant interaction. For example, soft data are already incorporated into hard data precisely through the process of collection. This is because part of the soft data is stored within the researcher themselves (ibid.: 283). The need for verification and opening of data would therefore lead to the storage and sharing of hard data (in some form), while soft data would remain beyond the reach of verification. According to Pool, the separation of soft and hard data would further delegitimize the fact that it is precisely soft data that are essential for ethnographic insight (ibid.). The logical conclusion is that such partially open data could not lead to original insight or e.g. enable verification.

Ethnology and cultural anthropology are, among other things, interpretive sciences, and the main issue with applying the argument of verifiability in their context is not just whether a given research context can be replicated, but whether there is an actual need to do so and what exactly would it accomplish. Ethnologists and cultural anthropologists sometimes return to the field themselves, repeat interviews, even with the same participants, in order to supplement or verify some of their conclusions. And they do not always receive identical responses, precisely because the research is shaped by the situation and the context. This "non-replicability of the research situation" leads to the fact, already recognized for a long time in ethnology and cultural anthropology, that a different researcher may receive entirely or somewhat different answers to the same questions within the same community (Geertz 1988: 5, according to Potkonjak 2014: 35). Victoria Reyes (2017: 8) rightly poses the question of which of two hypothetical statements given in two interviews, where the second is meant to verify the first, should be considered accurate. In the context of opening ethnographic data, the question arises whether a researcher, by analyzing transcripts from someone else's study, would conduct a different analysis. Ethnologists and cultural anthropologists believe that this is exactly what would happen, considering that in qualitative research, researchers themselves are "variables" (Potkonjak 2014: 35), a sort of "measuring instrument" (Milas 2009), or a place of storage for soft data (Pool 2017: 283). This fact has been acknowledged through the auto-reflexive turn in cultural anthropology and ethnology, and so autobiographical, reflexive elements are often incorporated into ethnographic texts.

"Ethnographic knowledge is much more than a qualitative methodology; it is an 'intellectual endeavor' that seeks to understand the observed or studied individuals

or communities, while anthropology as a discipline is realized through cultural hermeneutics, an interpretation of what has been understood" (Potkonjak 2014: 15). Reyes (2017: 3) argues that, in the process of opening data, it is more important to make decisions about transparency on a case-by-case basis, depending on the context, than to demand uniform transparency with the same rules for all.

Protection of personal and sensitive data

If we focus solely on the aspect of confidentiality of research participants' data – which, it seems, is the primary concern when it comes to demands for open data – rather than the epistemological differences between disciplines or data types, the most important distinction is that quantitative research does not reveal personal nor sensitive data, and there is generally no risk of disclosing the identity of research participants. In cases where such risk does exist, care is taken to ensure that these data are not made publicly available, as in the case of census data, which will be discussed later.

According to the Croatian Personal Data Protection Agency, personal data includes the following: an individual's address, telephone number, email address, personal photograph, personal identification number (e.g., the Croatian OIB), biometric data, educational background and qualifications, salary information, loan details, and bank account information (Celjak et al. 2020: 21). Draženko Celjak and colleagues note that personal data is defined as: "Any information relating to an identified or identifiable natural person" (ibid.). However, Article 4 of the General Data Protection Regulation provides a slightly broader definition:

> "personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. (General Data Protection Regulation)

In addition to personal data, sensitive data are also defined. These include: race or ethnicity, political beliefs, religious or other beliefs, trade union membership, health and sex life, as well as personal data relating to criminal and misdemeanor proceedings (Celjak et al. 2020: 21). It is clear that such data, just like personal data, should not be published without consent (ibid.). The same authors also cite Stanford University's recommendations on sharing sensitive data, which include both direct and indirect identifiers. Direct identifiers are any information that can directly identify an individual, and they must be excluded from datasets prior to public release. This category includes full names, initials, addresses, biometric data, email information,

web and IP addresses, phone numbers, vehicle license plate numbers, personal identification numbers, as well as photographs, audio recordings, names of relatives, and specific dates (such as weddings, etc.) (Celjak et al. 2020: 22–23).

Indirect identifiers include: the name of a person's physician or place of treatment, sex, rare diseases or treatments, risky behavior, place of birth, socioeconomic data (employment, occupation, annual income, etc.), geographic data (postal code of residence), description of the household and its members, year of birth or age, verbal responses, or interview transcripts. If a dataset "contains three or more indirect identifiers, it should be reviewed by an independent researcher or ethics committee to assess the potential risk of identification" (ibid.: 23).

The Croatian Bureau of Statistics, an institution that conducts the population census mandatory for all residents of the Republic of Croatia which collects both personal and sensitive data, has a similar approach to indirectly identifiable data. The Instruction on the protection rules and methods for the dissemination of official statistics of the Croatian Bureau of Statistics from June 30, 2021 (official gazette "Narodne novine" 78/21) (hereinafter: the Instruction) ensures that a statistical unit cannot be identified when publishing results. A statistical unit refers to a natural person, household, but a business entity as well. Paragraph 3 of Article 4 of the Instruction states: "The public use files contain data on individual statistical units which are anonymised in such a way that the statistical unit cannot be identified either directly or indirectly, taking into account all relevant means that could reasonably be used by a third party to identify them."

Article 9 outlines protection mechanisms for households and natural persons concerning: "1) The minimum frequency rule, according to which aggregated data relating to fewer than three statistical units may not be published" and "2) The dominance rule, according to which, if applicable, the aggregated data must be checked, and if one unit contributes dominantly to the aggregated data, then such data must not be disclosed due to the possibility of identifying the dominant statistical unit".

Article 11 prescribes methods of protection, ensuring that statistical units cannot be identified directly or indirectly, while Paragraph 2 states: "Data enabling the identification of the statistical unit shall be omitted or modified". Paragraph 3 of Article 11 also anticipates the following scenario: "In the event that protected confidential data can be calculated on the basis of other data, such data should also be protected in order to ensure secondary confidentiality".

In practice, this means that confidential data, such as those relating to religion, are not shown at the settlement level in publicly available tables, but only at the level of municipalities, cities, and counties. An exception to such rules is made for scientific research, as stated in the The ordinance on conditions and terms of using confidential statistical data for scientific purposes (NN 5/23) defining the method for accessing such data. All of this documentation is outlined in the Quality Guidelines of the Croatian Bureau of Statistics and demonstrates that significant efforts are made to protect confidential information.

The Data Act, which entered into force and began to be applied on September 12, 2025, mandates the sharing of data, but it also explicitly states that it "is without prejudice to Union law on the protection of personal data and privacy, in particular Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive 2002/58/EC". Point 7 of the recitals in the Data Act clearly states: "No provision of this Regulation should be applied or interpreted in such a way as to diminish or limit the right to the protection of personal data or the right to privacy and confidentiality of communications". This aligns with Article 7 of the General Data Protection Regulation (GDPR), which states: "Individuals should have control of their own personal data".

I consider this the main guiding principle in the process of opening ethnographic data. I also believe that a transcript, even pseudoanonymized, is an indirect identifier (Celjak et al. 2020). If we apply the recommendation that the presence of three or more indirect identifiers in a dataset represents a identification risk, I would argue that the combination of published works and data available through open access poses an even greater risk. For example, if we can link a pseudoanonymized transcript to the research location (which is often named in published ethnographic texts (Reyes 2017)) and to one additional sensitive piece of information that was anonymized and published in the text (such as an interlocutor's political opinions), there is always a possibility of secondary identification of the interlocutor, especially if the research was conducted in a smaller community. Even in the case of consent for such a situation, one must bear in mind that the interlocutor may not be fully capable of foreseeing all future risks or discomforts they may experience due to the openness of the data for which they gave consent. This is why I believe the researcher should explicitly warn the interlocutor of such a possibility and clearly explain what open data entails. The potential for identifying research participants always exists (cf. ibid.), but this does not mean we should abandon efforts to protect interlocutors or avoid taking steps to ensure no harm is done to them (Code of Ethics, Croatian Ethnological Society).

In the process of obtaining informed consent, it is customary to explain to the interlocutor the nature and purpose of the research, and to clearly emphasize that they may withdraw from the interview at any time, choose not to answer a particular question, or ask the researcher to turn off the recorder at a specific moment – or even insist that the conversation not be recorded at all. It is also possible to withdraw consent at a later stage, as illustrated by the experience of a colleague collaborating on a European project:

> Before every interview, we had to provide interlocutors with a document which they would sign, thereby giving consent for their statements and photos to be published. It was up to them to decide whether we would use just their initials or full names. In our case, people were eager to speak because […] and they saw it as a positive thing that this topic was being connected with similar topics across Europe, […] So we had no problems with that. But I know that in [country], there were people who, even after giving consent, later withdrew

> their statements in the middle of the project. So even once they've signed consent, they can revoke that decision… I mostly used only first names, even for those who told me I could include their surname… So the mention of someone named X, Y, or Z – names that exist in the hundreds – was not very revealing.

In most cases, interlocutors give their consent for the interview; sometimes they ask for the recorder to be turned off, and in some cases, they do not allow it to be turned on at all. In certain contexts, interlocutors do not wish to remain anonymous, but most of them – especially when discussing difficult or sensitive topics – choose anonymity. The question remains: how will that anonymity be preserved in the future if open data were to include anonymized or pseudoanonymized transcripts that could be compared to published material?

Research on attitudes toward open data among ethnologists and cultural anthropologists in Croatia[6]

The demand for open data and the issues it entails – particularly for ethnologists and cultural anthropologists – prompted me to conduct a study on attitudes toward open access, i.e., data sharing, among active ethnologists and cultural anthropologists in Croatia (Orlić 2024). I chose an online questionnaire as the method that would allow my respondents to express their views anonymously. The questionnaire I sent out was the result of my existing knowledge, based on the literature I had read on the topic of open data, my own research experience in the field, and was partially inspired by conducted surveys (Damalas et al. 2018). I sent the questionnaire to 145 email addresses available on the websites of institutions employing ethnologists and cultural anthropologists (museums, research institutes, universities, and conservation departments). The survey was not intended for retirees or students, as the practices of retired ethnologists in Croatia have generally differed from contemporary open data requirements, and students themselves are not yet involved in the work of institutions subject to data-sharing decisions. A total of 33 participants responded to the survey (22,75%), the majority of whom were women (87,9%). I am aware that this study may not have included all active ethnologists and cultural anthropologists in Croatia, and that this survey should be repeated after certain time. The low response rate may indicate a lack of interest in the topic or insufficient familiarity with the issue.

In the first part of the survey, I was interested in how respondents share their findings, how they store their data, whether they share data, do they use other people's data, and whether they even think data should be shared at all. Since the question regarding whether data should be shared was a yes/no type, the final question

6 This paper presents only part of the data (in the original language in which the research was conducted), while the complete results have already been published in English (Orlić 2024).

in this section of the survey was open-ended, inviting participants to explain their answers, should they wish to. These explanations were particularly illuminating for understanding attitudes toward open data from the perspective of ethnologists and cultural anthropologists in Croatia.

It is interesting to note that a third of respondents were against data sharing, 24% were in favor of sharing, and 15.4% supported sharing transcripts but not audio recordings or photographs. However, what I was able to conclude from the research is that, in cases of positive attitudes toward data sharing, most participants did not have in mind open access *per se*, but rather thought of sharing data within a research team – even then, individuals emphasized that this should happen "under clearly defined conditions."

> If the topic does not involve a discrepancy in power relations and cannot place the people involved in the research in a vulnerable position, and if those who will use the data are familiar with the context, then I believe that data sharing is good. This usually applies to members of a team researching the same topic, who understand the context, and sharing data saves time and provides a broader understanding of the research issue.

> I share raw data when needed with colleagues within the institution, but it is generally understood that one should not publish material collected by another researcher who intends to publish it themselves.

One respondent emphasized that they share data in accordance with institutional rules – with team members, the Research Archive of the respective institution, and with the interlocutor(s) involved in the research. The majority of respondents stated their primary concern is protecting the interlocutor, that is, ensuring that their confidential data is not compromised. Once again, the most significant obstacle to data openness proved to be the possibility of identifying members of small communities:

> Within qualitative research, everything is confidential, interlocutors can be easily recognized – academic research is not a gossip column!!!! I'm currently preparing for a study on xxx (a sensitive research topic) in Croatia. Does anyone think that such data should be public?!?!?!

> I work with very sensitive data and cannot be sure that someone will not be able to identify the person referred to in the documents and/or interview content.

> I am generally against sharing raw data – not because I oppose open science, but because I believe that in the case of cultural anthropology and ethnology, and especially in research on certain topics (e.g., taboos) or marginalized and unprotected groups, the ethical principles of the profession and the guarantee of the interlocutor's anonymity must take precedence and are a prerequisite for everything else. For example, in research involving small communities where everyone knows or can recognize each other, sharing an audio record-

> ing cannot guarantee anonymity (since a person's voice is unique and recognizable), nor can sharing other metadata captured through audio recordings or transcripts in which biographical elements or life details of the interlocutor are revealed – details that expose their identity and are not foreseen or protected by GDPR law.

One respondent emphasized the authorship aspect of interviews and the specificity of the interaction that enables data sharing between interlocutors and researchers, stressing that the use of such data by someone else would lack contextual understanding and could lead to misinterpretation.

> Interviews are an authored act between two or more participants. Another researcher may not always understand the context in which the interview was conducted, and it is possible that the data could be misinterpreted. The interviewer and interlocutor inevitably develop a certain interaction, deep or superficial, which is unique to each individual interview. The interlocutor decided to share their information and time with a particular researcher, regardless of motivation. I find it much more acceptable not to provide raw data, but only those that are published and structured. Raw data could perhaps be made subject to a time restriction, meaning they could be used only after a predetermined period has passed, with a requirement to provide all available contextual information. I believe the greatest issue with using raw data "second-hand" is an ethical one, precisely due to the risk of incorrect or superficial interpretation.

Another respondent acknowledged the value of such data for future research and reflected on the possibilities of open data with regards to sensitive archival materials:

> In practice, anthropological research deal with a wide range of topics and possibilities. If we are working with vulnerable groups on sensitive topics or with very small, easily identifiable communities, it is not easy to categorically ensure principles of data sharing. Generally, I believe it is necessary to share data, but only when doing so does not endanger our participants and when they themselves have clearly chosen to allow it. In the case of sensitive topics, I believe there should be a moratorium on raw data, transcripts, audio recordings, and other collected materials. Such a moratorium could be envisioned in a way similar to how archival science handles personal records or sensitive information such as hospital charts. These materials can become accessible for research only after a certain period of time has passed.

A third respondent described as acceptable a practice already foreseen within the FAIR principles, which refers to the openness of metadata, with institutions deciding on access to the actual data.

> I certainly believe that data should be shared; the only question is with whom and in what format. A possible interim solution would be to make public the existence of specific data within a database, along with the suggestion that interested parties contact the institution holding the raw data.

A fourth respondent proposed a way to tailor transcripts to specific purposes, while also expressing a general opinion that data sharing is a counterproductive and dangerous process, precisely because the context of a research conversation is unrepeatable.

> Ethnographic research is complex and often intimate in nature, biographical, and inseparable from the context of its creation. Given this, a transcript which may exist in one of three versions – can be prepared for consultation and sharing or may serve as a reflection of the audio recording. Since this is conversational relationship which does not register repeatable, fully comparable data that can be extrapolated and used out of context, I believe data sharing is counterproductive and dangerous.

In the second part of the survey, I explored respondents' willingness to share data under specific conditions. This part of the survey consisted of a series of statements for which the participants had to grade their agreement/disagreement (Likert scale). Some of these statements were adapted from a questionnaire used by Damalas et al. in 2018, as I found some of the hypothetical scenarios and self-assessments the authors explored to be important for analysis.

I asked respondents whether they would be willing to share data before publication, whether they would agree to share data if it enabled research funding, whether they would share data in support of open science goals, and whether they would be willing to share data if it would secure publication in high-ranking journals or collaboration with renowned researchers. I was also interested in their stance on supporting the goals of open science in general. The second-to-last question (in a yes/no format) addressed attitudes toward sharing raw data under specific conditions. In the final question, participants were invited to explain their answer, if they wished to do so.

Some reiterated previously mentioned arguments that data sharing within a research team is justifiable, while others emphasized that its justification should be evaluated even within such collaborations. One respondent noted that it is necessary to clearly explain to the interlocutor before the interview what raw data and open access actually mean.

> With the previous explanation in mind, I would like to emphasize that partial sharing of raw data with researchers and colleagues involved in the same project is justified, as it contributes to the overall research of a particular topic or area, as well as the analysis and interpretation of data aimed at ensuring research excellence. However, I believe it is important that the informed consent form explicitly states this to the interlocutor, who should be made aware not only of the goals of the research, but also of who will have access to the data.

> The participant or interlocutor must be informed about the concept of "raw data" and how those data will be further distributed within the framework of open access. This is a precondition for consent to the interview. Raw data should not be shared afterwards if the interlocutor was not previously

> informed. It is common practice to inform the interlocutor about the purpose of the research, which is usually limited to a specific topic, and collecting "raw data" may seem too general and unclear to interlocutors.

The need to protect the interlocutor (as well as the researcher) is a frequent argument, and the most numerous responses provided different justifications for protecting the interlocutor and preserving the trust gained in the research process. Some statements clearly emphasize that the decision on openness should be made based on the type of data collected and the potential harm data sharing might cause to participants.

> It depends on the type of research and the data collected. It's easier to share data when it concerns public knowledge, but much harder and more problematic when it involves private experiences.

> Data should be shared with the consent of those involved (and sometimes not just with their consent, but with their active wish/intent), but not during the project, only after the project has concluded and the research coming from the "raw data" is published. That said, the answer heavily depends on the nature of the raw information. In principle, it's a "yes," but in many specific cases, it could be a "no" (e.g., if it is deemed that publishing might harm the participants, even if they are willing to share the data; or if the raw data contains elements that have not yet been thematized in published work, etc.).

> It might depend on the topic of the research, but every answer is personal and may reveal something the interlocutor does not want shared.

One respondent pointed out that data sharing could also contribute to the production of digital waste.

> The fact is that we, as researchers, despite our initial obligation and desire to objectively address a selected topic, are prone to subjective interpretations that support our original assumptions. In this sense, some data may remain unused but could be useful to another researcher or user. However, it is also a fact that some of the material collected is often "unusable" due to the quality of … [data provided by] … the interlocutor, which raises the question of whether making such material public only creates a sort of digital waste.

Others, however, emphasized the issue that many raw data remain in private archives of researchers, and also noted that we should start thinking of ourselves as creators of archival materials that will remain for future research.

> From my own professional experience, I know that many raw data from research in Croatia (not only by ethnologists and cultural anthropologists) are stored on researchers' personal computers, external hard drives, printed and shelved at home, in offices, etc. There is a need to raise awareness among researchers about more careful and at the same time more open storage of such information, which could serve other researchers as well.

Of course, this also means that the wishes of the research participants must be respected, especially when creating archival materials intended for future researchers. One explanation was related to the way in which contemporary researchers use archival materials themselves.

> Raw data should be shared in any case. Just as we have access to data in archives, so too should we enable others to access our data. Of course, this should involve resolving ethical concerns, handling sensitive information, and separating interview transcripts from field notes.

Most responses showed a willingness to share data when working as part of a team, but even in these cases, concerns about protecting the interlocutors prevail. Participants in this survey demonstrated exceptional caution regarding the sharing of raw data, even with all the mechanisms for interlocutors' protection in place, primarily because research is most often conducted in small communities and frequently involves sensitive or difficult topics.

Some explanations and remarks point to the belief that open data in the context of ethnographic research is a completely misguided demand. Others express concern that, given the potential for open data in the future, it may become increasingly difficult to motivate participants to speak about sensitive topics:

> As stated in my previous response, I consider data sharing in ethnography to be a misinterpretation of ethnography as an epistemological strategy.

> I would just like to emphasize that my views stem from the conviction that, in our research, the primary focus should be on protecting interlocutors and (to a degree) maintaining a balanced relationship. If all raw data were made universally accessible, there is a legitimate concern that it would become harder to find interlocutors when researching "sensitive" topics. I also wonder how misuse of such data would be prevented. Finally, in the context of cultural anthropology and ethnology, raw data generated in the field are the result of an interaction between two people (the researcher and the interlocutor), and are therefore inherently contextual. Another researcher might receive different answers to the same questions from the same interlocutor. For that reason, I do not see the justification for the use of raw data by other researchers, especially not in different projects or analyses. In such cases, the use and analysis of raw data would only make sense if the research is concerned with the anthropologist/ethnologist in question and their approach to fieldwork.

The results indicate that the potential for identifying research interlocutors emerges as a key issue and obstacle to open data, but that most survey participants frame this concern as the need to protect interlocutors, in line with the ethical principles of the discipline.

Given that transcripts are one of the primary forms of qualitative data storage, and are also, in some viewpoints, considered indirect identifiers, it becomes highly

problematic to simultaneously share data such as transcripts and protect personal and sensitive information if this is stipulated by informed consent agreements. Anonymization or pseudonymization of transcripts, which is often proposed as a technical solution (Celjak et al. 2020: 23), may work in a technical sense, but the already discussed problem of the usability of such materials in the context of ethnographic research remains unresolved. The possibility of interlocutor identification increases when data are made public, while the usefulness of such decontextualized, anonymized or pseudonymized transcripts remains questionable.

In conclusion

As previously mentioned, it is a well-known fact within ethnology and cultural anthropology that a specific feature of ethnographic research is that the researcher themselves is also a variable in the research. I have already emphasized that data openness requirements should take into account the epistemological differences between disciplines, as well as the ways in which knowledge is produced within those disciplines. A smaller number of participants in the study conducted among Croatian ethnologists and cultural anthropologists reflected on sharing fieldnotes, considering them something that should not be shared. Some scholars have noted that ethnologists and cultural anthropologists tend to keep their fieldnotes until death (Zeitlyn and Lyon 2012), while others have warned of numerous potential issues that introducing the practice of sharing fieldnotes could inadvertently cause (despite its potential benefits) such as the standardization of notes, loss of creativity, and other unintended consequences (Reyes 2017). I believe that open data should also imply that the researcher retains the right to determine access to the data they themselves have created, such as notes and diaries.

I also agree with Reyes' view that it is less about sharing for its own sake (depending on the context), and more about making one's research process more transparent, while still adhering to the ethical principles of the profession. Insisting on full openness of raw data could, in certain contexts, conflict with the ethical standards of the discipline as well as with the General Data Protection Regulation. Access to raw data obtained through academic research must comply with data protection laws (Mondschein and Monda 2019), which is why it is essential to properly design appropriate consent, i.e. draft good informed consent forms which clearly and unambiguously defines how the research data will be managed.

In Croatia, informed consent is still sometimes given verbally, although calls for written consent can be increasingly heard, both for professional standards and out of decency (Potkonjak 2014: 45). The Croatian Ethnological Society has published two examples of informed consent forms[7] on its website, alongside the Code of

7 These informed consents are called the Letter of Collaborative Consent (Pismo suradničke suglasnosti) and the Extended Letter of Collaborative Consent (Prošireno pismo suradničke suglasnosti). https://hrvatskoetnoloskodrustvo.hr/o-nama/eticki-kodeks/, accessed 20 May 2025).

Ethics, as early as 2013. These forms could serve as templates or guidelines for those drafting their own consent forms. In her book "Fieldwork for Beginner Ethnologists" (*Teren za etnologe početnike*), Sanja Potkonjak published a draft proposal for an informed consent form, primarily intended for students. The draft includes the possibility of changing the context of use of research data and therefore contains the following important clause: "Any subsequent use of the research results outside the approval granted by this letter requires the researcher to obtain consent again" (Potkonjak 2014: 45–46).

In the informed consent form created for the Croatian Science Foundation project Solidarity Economy in Croatia: Anthropological Perspective (SOLIDARan), the following statement is included: "Your personal data collected in this research interview will be used exclusively for academic purposes. Data from your interview may be used (in the form of quotes or selected statements) during the dissemination of research results by any project collaborator: in academic papers, books, or during conference presentations, orally or via posters." It is also clearly states that "the transcripts of the research interview will be made available only to the collaborators of this research, with the aim of conducting academic analyses." This means that for any other use, the researchers would be required to contact the interlocutor again and obtain their explicit consent (Orlić 2024).

The primary repository for long-term storage in line with both FAIR and CARE principles should continue to be well-written ethnographies, which describe the nature of the research process that led to published results in a transparent manner, as has been the case up to now, or even more transparent. Data used for publication that were obtained from research participants should be used and stored in accordance with the informed consent given by the interlocutors.

An important question is how open data will affect the quality of future ethnographic research in the long term, even when the collected data are fully opened with the interlocutor's consent. It seems to me that a key change will occur precisely due to the separation of soft and hard data, as only the hard data will remain in repositories, while everything else will remain within the realm of trust created between the interlocutor and the researcher – stored within each of them. The opening of data for project administration purposes and to meet the requirements of open science, especially if data are opened merely because openness becomes a standard or an element of researcher evaluation (!), without assessing whether such requirements are meaningful, fundamentally alters a core element of ethnographic and cultural anthropological research: insight from the interaction between interlocutor and researcher. Awareness of open data in the near or distant future could significantly change the nature of our interlocutors' narratives. Some researchers believe that the character of such narratives will change considerably, and not for the better:

> I believe we will get highly filtered, highly refined statements… As ethnologists, we often achieve a certain intimate atmosphere in our interviews, and interlocutors may not even be aware of how much they reveal during these conversations. However, the moment they become aware that everything they

> say might be made publicly accessible somewhere, that will change the relationship between ethnologists and their interlocutors in the field, even when it comes to seemingly benign topics… I think that is not good at all.

For all the reasons mentioned above, I believe that opening metadata would provide an optimal level of openness in the case of ethnology and cultural anthropology, precisely because it simultaneously aligns with professional ethical standards and the General Data Protection Regulation. The key to data management remains the informed consent.

BIBLIOGRAPHY

Brewster, Christopher, Barry Nouwt, Stephen Raaijmakers and Jack Verhoosel. 2020. "Ontology-Based Access Control for FAIR Data". *Data Intelli*gence 2/1–2: 66–77. https://doi.org/10.1162/dint_a_00029

Burrows, Toby. 2011. "Sharing Humanities Data for E-Research. Conceptual and Technical Issues". In *Sustainable Data from Digital Research*. Nick Thieberger, Linda Barwick, Rosey Billington and Jill Vaughan, eds. Melbourne: PARADISEC, 177–192.

Cabello Valdes, Cecilia Bernard Rentier, Eeva Kaunismaa et al., eds. 2017. *Evaluation of Research Careers Fully Acknowledging Open Science Practices. Rewards, Incentives and/or Recognition for Researchers Practicing Open Science*. Bruxelles: European Commission Directorate-General for Research and Innovation. https://doi.org/10.2777/75255

Celjak, Draženko, Ivana Dorotić Malič, Marta Matijević, Ljiljana Poljak, Kristina Posavec and Ivana Turk. 2020. *Istraživački podaci – što s njima? Priručnik o upravljanju istraživačkim podacima*. Zagreb: Sveučilište u Zagrebu, Sveučilišni računski centar.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez et al. 2020. "The CARE Principles for Indigenous Data Governance". *Data Science Journal* 19/1: 1–12. https://doi.org/10.5334/dsj-2020-043

Carroll, Stephanie Russo, Edit Herczog, Maui Hudson, Keith Russell and Shelley Stall. 2021. "Operationalizing the CARE and FAIR Principles for Indigenous Data Futures". *Scientific Data* 8: 108. https://doi.org/10.1038/s41597-021-00892-0

Carroll, Stephanie Russo, Duarte Marisa and Liboiron Max. 2024. "Indigenous Data Sovereignty". In *Keywords of the Datafied State*. Jenna Burrell, Ranjit Singh and Patrick Davison, eds. New York: Data & Society Research Institute, 207–223.

Damalas, Dimitrios, Gregoria Kalyvioti, Evelina Sabatella and Konstantinos I. Stergiou. 2018. "Open Data in the Life Sciences. The 'Selfish Scientist Paradox'". *Ethics in Science and Environmental Politics* 18/1: 27–36. https://doi.org/10.3354/esep00182

Hannerz, Ulf. 2006. "Studying Down, Up, Sideways, Through, Backward, Forwards, Away and at Home. Reflections on the Field Worries of an Expansive Discipline". In *Locating the Field. Space, Place and Context in Anthropology*. Simeon Coleman and Peter Collins, eds. Oxford: Berg, 23–42. https://doi.org/10.4324/9781003085904-2

Ivanjko, Tomislav, Goran Zlodi and Zoran Horvat. 2024. *Građanska znanost*. Zagreb: Sveučilište u Zagrebu, FF-press. https://doi.org/10.17234/9789533791333

Jacobsen, Annika, Ricardo de Miranda Azevedo, Nick Juty et al. 2020. "FAIR Principles. Interpretations and Implementation Considerations". *Data Intelligence* 2/1–2: 10–29. https://doi.org/10.1162/dint_r_00024

Kolbe, Vera. 2022. "Open Science versus Data Protection. Challenges and Solutions in Sign Language Acquisition Studies". *Hrvatska revija za rehabilitacijska istraživanja* 58/Special Issue: 109–120. https://doi.org/10.31299/hrri.58.si.6

Kriegeskorte, Nikolaus. 2012. "Open Evaluation. A Vision for Entirely Transparent Post-Publication Peer Review and Rating for Science". *Frontiers in Computational Neuroscience* 6: 79. https://doi.org/10.3389/fncom.2012.00079

Landi, Annalisa, Mark Thompson, Viviana Giannuzzi et al. 2020. "The 'A' of FAIR. As Open as Possible, as Closed as Necessary". *Data Intelligence* 2/1–2: 47–55. https://doi.org/10.1162/dint_a_00027

Marić, Ivana, Kosjenka Dumančić and Marija Šimunović. 2023. "Akademska zajednica kao akter zagovaranja otvorene znanosti u visokom obrazovanju". *Vjesnik bibliotekara Hrvatske* 66/3: 119–149. https://doi.org/10.30754/vbh.66.3.1122

Milas, Goran. 2009. *Istraživačke metode u psihologiji i drugim društvenim znanostima*. Jastrebarsko: Naklada Slap.

Miller, Daniel. 2012. "Open Access, Scholarship, and Digital Anthropology". *HAU. Journal of Ethnographic Theory* 2/1: 385–411. https://doi.org/10.14318/hau2.1.016

Mondschein, Christopher F. and Cosimo Monda. 2019. "The EU's General Data Protection Regulation (GDPR) in a Research Context". In *Fundamentals of Clinical Data Science*. Pieter Kubben, Michel Dumontier and Andre Dekker, eds. Cham: Springer, 55–71. https://doi.org/10.1007/978-3-319-99713-1_5

Mons, Barend, Cameron Neylon, Jan Velterop, Michel Dumontier, Luiz Olavo Bonino da Silva Santos and Mark D. Wilkinson. 2017. "Cloudy, Increasingly FAIR. Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud". *Information Services and Use* 37/1: 49–56. https://doi.org/10.3233/ISU-170824

Mons, Barend, Erik Schultes, Fenghong Liu and Annika Jacobsen. 2020. "The FAIR Principles. First Generation Implementation Choices and Challenges". *Data Intelligence* 2/1–2: 1–9. https://doi.org/10.1162/dint_e_00023

Orlić, Olga. 2024. "Is This the End of Anthropology as We Know It? Some Implication of FAIR Principles on Tales in Ethnological and Anthropological Qualitative Research". *Publications* 12/4: 42. https://doi.org/10.3390/publications12040042

Pampel, Heinz and Sünje Dallmeier-Tiessen. 2014. "Open Research Data. From Vision to Practice". In *Opening Science. The Evolving Guide on How the Internet Is Changing Research*. Bartling Sönke and Sascha Friesike, eds. Cham: Springer, 213–224. https://doi.org/10.1007/978-3-319-00026-8_14

Pool, Robert. 2017. "The Verification of Ethnographic Data". *Ethnography* 18/3: 281–286. https://doi.org/10.1177/1466138117723936

Potkonjak, Sanja. 2014. *Teren za etnologe početnike*. Zagreb: Hrvatsko etnološko društvo, Odsjek za etnologiju i kulturnu antropologiju Filozofskog fakulteta Sveučilišta u Zagrebu.

Priem, Jason, Dario Taraborelli, Paul Groth and Cameron Neylon. 2010. *Altmetrics. A Manifesto*. https://doi.org/10.5281/zenodo.12684248

Reyes, Victoria. 2017. "Three Models of Transparency in Ethnographic Research. Naming Places, Naming People, and Sharing Data". *Ethnography* 19/2: 204–226. https://doi.org/10.1177/1466138117733754

Stojanovski, Jadranka. 2019. "Promicanje ideja otvorene znanosti u Hrvatskoj". *Socijalne teme* 1/6: 166–170. https://hrcak.srce.hr/233286 (accessed 18 August 2025).

Stojanovski, Jadranka and Danijel Mofardin. 2025. "Diamond Open Access Landscape in Croatia. DIAMAS Survey Results". *Publications* 13/1: 13. https://doi.org/10.3390/publications13010013

Tenopir, Carol, Suzie Allard, Kimberly Douglass et al. 2011. "Data Sharing by Scientists. Practices and Perceptions". *PLoS ONE* 6/6. https://doi.org/10.1371/journal.pone.0021101

Ukić, Šime. 2022. "Mišljenja i komentari. (Is)korak prema reformi sustava vrednovanja istraživačkog rada u EU-u". *Kemija u industriji* 71/1-2: 77–77. https://hrcak.srce.hr/clanak/393196 (pristup 18. 8. 2025.).

van Rooyen, Susan, Fiona Godlee, Stephen Evans, Nick Black and Richard Smith. 1999. "Effect of Open Peer Review on Quality of Reviews and on Reviewers' Recommendations. A Randomized Trial". *British Medical Journal* 318: 23–27. https://doi.org/10.1136/bmj.318.7175.23

Vlašiček, Denis and Ivan Flis. 2021. "Otvorena znanost. Kratak pregled pokreta i metodološkog značaja". *Revija za sociologiju* 51/3: 507–516. https://hrcak.srce.hr/269814 (accessed 18 August 2025).

Vučković Juroš, Tanja. 2021. "Pogled prema naprijed. Prethodna priopćenja u ubrzanom postupku i politika otvorene znanosti". *Revija za sociologiju* 51/3: 301–307. https://hrcak.srce.hr/269681 (accessed 18 August 2025).

Walther, Alexander and Jasper J. F. van den Bosch. 2012. "FOSE. A Framework for Open Science Evaluation". *Frontiers in Computational Neuroscience* 6: 32. https://doi.org/10.3389/fncom.2012.00032

Wilkinson, Mark.D., Michael Dumontier, IJsbrand Jan Aalbersberg et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship". *Scientific Data* 15: 3. https://doi.org/10.1038/sdata.2016.18

Zeitlyn, David and Stephen M. Lyon. 2012. "Varieties of Openness and Types of Digital Anthropology. Avoiding Confusion in Discussing Danny Miller". *Durham Anthropology Journal* 18/2: 97–110.

SOURCES

Budapest Open Access Initiative. 2002. https://www.budapestopenaccessinitiative.org/read/ (accessed 15 October 2025).

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. 2003. https://openaccess.mpg.de/Berlin-Declaration (accessed 15 October 2025).

Direktiva 2002/58/EZ Europskog parlamenta i vijeća od 12. srpnja 2002. o obradi osobnih podataka i zaštiti privatnosti u području elektroničkih komunikacija (Direktiva o privatnosti i elektroničkim komunikacijama). 2002. *Službeni list Europske unije* 13/52: 111–121. (accessed 30 September 2025).

DZS (Državni zavod za statistiku / Croatian Bureau of Statistics). 2025. *Smjernice za kvalitetu u hrvatskome statističkom sustavu*. https://dzs.gov.hr/UserDocsImages/dokumenti/Smjernice%20za%20kvalitetu%20u%20HRSS_.pdf (accessed 3 June 2025).

EC (European Commission). 2017. *H2020 Programme. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020.* https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf (accessed 30 September 2025).

GBIF (Global Biodiversity Information Facility). 2025. *Data Papers. Getting Scholarly Recognition for Your Datasets*. https://www.gbif.org/data-papers (accessed 10 December 2025).

HED (Hrvatsko etnološko društvo / Croatian Ethnological Society). n.d. "O hed-biblioteci". https://hrvatskoetnoloskodrustvo.hr/o-hed-biblioteci/ (accessed 15 October 2025).

HED (Hrvatsko etnološko društvo / Croatian Ethnological Society). 2013. *Etički kodeks Hrvatskog etnološkog društva*. https://hrvatskoetnoloskodrustvo.hr/o-nama/eticki-kodeks/ (accessed 27 May 2025).

HRZZ (Hrvatska zaklada za znanost / Croatian Science Foundation). 2022. *Plan upravljanja istraživačkim podacima*. https://hrzz.hr/plan-upravljanja-istrazivackim-podacima-za-projekte-hrvatske-zaklade-za-znanost (accessed 1 June 2025).

Naputak o načinu zaštite povjerljivih statističkih podataka u državnom zavodu za statistiku. 2021. *Narodne novine 78/21*. https://narodnenovine.nn.hr/clanci/sluzbeni/2021_07_78_1463.html (pristup 30. 9. 2025.).

Naputak o pravilima i metodama zaštite pri diseminaciji službenih statističkih podataka Državnog zavoda za statistiku. 2021. *Narodne novine* 78/21. https://narodne-novine.nn.hr/clanci/sluzbeni/2021_07_78_1464.html (pristup 30. 9. 2025.).

Pravilnik o pristupu povjerljivim podacima Državnog zavoda za statistiku unutar sustava službene statistike. 2021. *Narodne novine* 59/21. https://narodne-novine.nn.hr/clanci/sluzbeni/2021_05_59_1140.html (pristup 30. 9. 2025.).

Strateški plan Hrvatske zaklade za znanost za razdoblje 2023. – 2027. 2023. *Narodne novine* 121/23. https://narodne-novine.nn.hr/clanci/sluzbeni/full/2023_10_121_1679.html (pristup 3. 6. 2025.).

UNESCO. 2021. *UNESCO Recommendation on Open Science*. https://unesdoc.unesco.org/ark:/48223/pf0000379949?posInSet=3&queryId=4e9e5e25-46d0-428c-86ae-4191da67426c (pristup 12. 10. 2025.).

Uredba (EU) 2023/2854 Europskog parlamenta i Vijeća od 13. prosinca 2023. o usklađenim pravilima za pravedan pristup podacima i njihovu uporabu i o izmjeni Uredbe (EU) 2017/2394 i Direktive (EU) 2020/1828 (Akt o podacima). 2023. https://www.zakon.hr/c/eu/61393/uredba-%28eu%29-2023-2854-europskog-parlamenta-i-vijeca-od-13.-prosinca-2023.-o-uskladenim-pravilima-za-pravedan-pristup-podacima (pristup 3. 6. 2025.).

Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka te o stavljanju izvan snage Direktive 95/46/EZ (Opća uredba o zaštiti podataka – GDPR). 2016. https://www.zakon.hr/z/3112/opca-uredba-o-zastiti-podataka---uredba-%28eu%29-2016-679- (pristup 3. 6. 2025.).