Koraljka Kuzman Šlogar
Institute of Ethnology and Folklore Research, Zagreb
orcid.org/0000-0002-7810-0453
koraljka@ief.hr

Open Science Between Technical Feasibility and Epistemological Maturity

In recent decades, open science has been established as a set of practices through which we can recognize a shift toward a new scientific paradigm, one based on transparency, accessibility, and the possibility of reusing knowledge. In the European research context, it has in fact become a political and financial priority: the openness of data and publications is increasingly being set as a requirement for project funding, ranking of institutions, and the evaluation of academic work. This process in the field of the arts and humanities has been systematically monitored and somewhat shaped by the DARIAH-EU consortium (Digital Research Infrastructure for the Arts and Humanities), which has been gradually integrating the principles of open science into its strategic orientation and activities since 2017, particularly through the promotion of ethically grounded, context-sensitive, and infrastructurally supported openness (Edmond and Tóth-Czifra 2023). However, such a paradigm does not pose the same challenges to all disciplines. While openness in the natural sciences, for example, is often considered in terms of technical issues of standardization and reproducibility, in the humanities, especially in ethnology and anthropology, questions of relationships, trust, and the very context in which data is produced and acquires meaning come under the spotlight.

In this sense, I would like to highlight the value of Olga Orlić's text "Being Open about Open Data from the Perspective of Ethnology and Cultural Anthropology". The author clearly and precisely identifies the key dilemmas and controversies that open science brings to our discipline: the conceptual incompatibility of "soft data" with models of openness shaped within STEM frameworks; the ethical concerns regarding the sharing of records produced in intimate, trust-based relationships with interlocutors; and the insufficient readiness of researchers to cope with the infrastructural and administrative demands of new science policies. Her analysis and research carried out among ethnologists and cultural anthropologists in Croatia is an exceptionally valuable contribution because it neither trivializes the problems nor reduces them to technical obstacles, but examines them in their epistemological and ethical complexity.

As an ethnologist, I share the author's sensitivity and skepticism, yet I feel the need to further articulate the complexity of the situation in which we find ourselves. The text leads us to reflections on several connected levels, but I would above all like to emphasize the following: openness is not a simple "yes or no" decision, but a

series of small, carefully considered steps – from planning consent, through description and storage, to carefully managed access. Each step is an opportunity to weigh risks and benefits, to respect the wishes of our interlocutors, and also support the academic community. In this delicate position between sharing and withholding, between visibility and protection, we shape our responsibility toward the people, relationships, and stories that do not belong solely to us but also to our interlocutors and the communities from which they come.

Transparency and academic responsibility

My view of open science is shaped by three types of experience: ethnographic research, which included fieldwork; the years I spent working as the head of a research archive (IEF); and my experience coordinating and collaborating with the European digital research infrastructure (DARIAH-EU). The first taught me that data is formed in the fine and fragile nuances of interpersonal relationships; the second that knowledge carries value only if we preserve it and pass it on; and the third that openness can never be reduced to mere technique, because it is always also a sensitive ethical issue. This is why I approach the topic not only from a theoretical standpoint, but also from a practical position, based on everyday work in which research methodology and research ethics, archival responsibility, and digital infrastructure intersect. This threefold perspective also shapes my commentary: I do not wish to relativize ethical dilemmas nor to unilaterally advocate for open access, but rather to explore the possibilities of building data-management practices in ethnology and cultural anthropology that are ethically sustainable, technically feasible, and academically relevant. I believe that openness is not a ready-made formula, but an ongoing examination of how to act responsibly – within the frameworks of research, the technologies that shape our work, and the relationships we build with our interlocutors.

In qualitative disciplines, as Orlić rightly notes, open data cannot be based solely on the FAIR principles; it is necessary to integrate ethically oriented standards such as the CARE model and to build infrastructures that not only preserve data but also acknowledge the complexity of the relationships, trust, and contexts from which they emerge. Instead of allowing open science to be imposed as a bureaucratic imperative, by incorporating this ethical dimension we can transform it into a space of sustainable academic practice conducted with solidarity. In this commentary, I reflect on the conditions under which, in my view, open data in the cultural-anthropological research space becomes both feasible and well-founded: transparently documented research processes, reliable institutional storage, and a fair and well-regulated access framework.

In ethnology and cultural anthropology, transparency certainly does not mean replicating results, since every fieldwork experience is undoubtedly unique and unrepeatable. Transparency here rests on the documentation and verifiability of the process: from planning and data collection to processing, storage, and access, all supported by metadata and notes that allow us to understand how the material was

shaped by a specific relationship and situation. Such transparency is a prerequisite for sound subsequent interpretation. If we know who shaped the material, how did they do so, and under what circumstances, we can more accurately assess its limitations and scope. These elements are considered in advance in the Data Management Plan (DMP): levels and types of metadata, types of licenses, moratoriums, access procedures, and models of reuse.

But how do we ensure such transparency? Who will leave these traces, document the data, prepare it for storage, shape metadata, write permissions and licenses? In many European countries this work is carried out by so-called data stewards, but in Croatia this role is not yet sufficiently recognized nor institutionalized, and the responsibility often falls on researchers themselves. This is a form of "invisible labor" without which data cannot truly function within open science, yet our system of academic evaluation still does not acknowledge it as a scholarly contribution.[5]

Archiving and opening research data

I fully support the author's clear position that attempting to universally apply open science principles – especially regarding the sharing of so-called raw data – is conceptually, methodologically, and ethically problematic in qualitative disciplines. In ethnology and cultural anthropology, data is not a neutral record, but the result of situated, corporeal, and affective encounters that cannot be reduced to technical formats or fitted into universal standards. In this sense, in ethnology there is, strictly speaking, no such thing as "raw" data – every record is already shaped through relationship, method, context, and trust. Therefore, archiving is not the mere technical storage of files, but an interpretive practice with ethical consequences. This raises a legitimate question: how can emotionally charged and relationally structured records be archived without losing their meaning – and under what conditions does such transmission remain trustworthy?

Instead of universal standardization, the DARIAH community has emphasized the need for a contextual approach to metadata, one that preserves information about the context of creation, provenance, and subsequent interventions made in the material (Tóth-Czifra 2019). Data access is regulated through access levels of availability, embargo periods, and agreed-upon rules, while ethical aspects are not added retroactively but integrated into the infrastructure and operational protocols from the very beginning.

However, when this type of data is standardized for repository or open-system environments, there is a risk of reducing their context and meaning. Erzsébet Tóth-Czifra (2019) rightly warns that "FAIR-ification" (if implemented carelessly) may

5 This is why activities within DARIAH are increasingly directed toward reforming the research assessment system. The position paper The Role of Research Infrastructures in the Research Assessment Reform (Tasovac et al. 2023: 3) emphasizes that "work on data must be recognized and systematically evaluated in hiring" and promotion procedures, on an equal footing with traditional publications – including datasets, software, and other non-narrative research outputs. Otherwise, open science becomes a burden rather than an opportunity, especially in disciplines such as ethnology and cultural anthropology.

inadvertently mold humanities materials according to STEM norms. A particular challenge is posed by the so-called soft data discussed by Orlić: silences, affective gestures, glances, discomfort – what is left between words. Such records contain knowledge embodied in the researcher's experience, situated in a specific place, time, and set of circumstances, and produced through the interactions among participants. This is why they cannot be separated from the context that shaped them. However, I do not think that the solution is to reject standardization altogether, but rather to carefully apply it, in ways that take into account the specific conditions of material creation and its sensitivity. For this reason, FAIR, which provides crucial technical prerequisites for sustainable data management (see Wilkinson et al. 2016), needs to be complemented by the CARE principles, which emphasize collective benefit, authority over data, responsibility, and (relational) ethics (Carroll et al. 2020: 1–2, 5). Combined with ethics "embedded in the process" and contextual metadata, the result is an operational framework for openness that aligns with the sensitivity of the material and the obligations toward participants. Thus, openness ceases to be merely declarative and becomes operational – measurable, accountable, and feasible in practice.

***

Discussions on open data often overlook the question of storage and responsibility: where is the material kept, and who takes care of it? When I speak of storage, I do not refer only to the technical act of storing files, but rather to a long-term, professionally guided, policy-driven process. It is important to distinguish general storage – for example on personal computers, USB drives, or cloud services – from archiving, which entails institutional, systematic, and well-documented care for the material. Archiving includes metadata, access protocols, ethical evaluation, and long-term availability, and the responsibility for preserving knowledge shifts from the individual to the institution. Only through such archiving do records acquire context and scientific value – not so much in terms of verifiability, but in terms of their usefulness to future generations of researchers.[6]

While working in a research archive, I noticed that since the early 2000s – with the advent of digital recorders, smartphones, and easily accessible personal equipment for fieldwork – the amount of systematically preserved material within the institution has significantly decreased. Although digitization has opened many possibilities for preservation and accessibility, it has paradoxically also contributed to the fragmentation of research documentation. Materials are increasingly stored on personal computers, private servers, or in the "cloud," without long-term strategies and outside institutional frameworks, leaving open the question of who is responsible for the long-term care of the data and their accessibility.

6 Ultimately, this material is not used exclusively by researchers. For an overview of user typologies and examples of the use of archived records, see Kuzman Šlogar, Koraljka, Anamarija Žugić Borić i Antonia Hladilo Duspara. 2023. "Arhivi i korisnici. Primjer odjela dokumentacije Instituta za etnologiju i folkloristiku u Zagrebu". *Arhivski vjesnik* 66/1: 65–95.

Part of the reason may lie in the lack of infrastructure, institutional support, and clear procedures, but another factor is certainly the growing sensitivity of research topics and ethical dilemmas (as also indicated by the responses in Orlić's survey). The fact is that research topics have changed. The "collecting" ethnology of the past predominantly documented songs, customs, dances, and other forms of collective practice, whereas contemporary ethnology and cultural anthropology examine more intimate, personal, and socially sensitive topics such as experiences of migration, marginalization, poverty, or sexuality. Therefore, storage must be designed in a way that protects sensitive data through clearly defined access authorizations, oversight committees, and ethical protocols. When depositing material, it is necessary to carefully consider its sensitivity and the potential consequences for interlocutors, communities, and researchers, and determine access levels and any restrictions based on these criteria. Systemic archiving of data, in any case, should not be an exception but an integral part of academic responsibility – the responsibility to ensure that the knowledge we produce is accessible and useful – both now and in the future.[7]

Open science as a responsibility, not an imperative

I fully support Orlić's position that making metadata openly available is a necessary precondition for academic transparency. Metadata provides basic visibility of the material: information about its existence, location, and access conditions; even when the content itself is not open. However, I believe that we can and should take a step further. Transparency does not end with visibility; it also requires a responsible storage infrastructure. If records are kept on personal computers, outside institutional frameworks, transparency may be formally satisfied, but this makes them academically ineffective. Who, other than the researcher alone, will be able to access such records? Moreover, it is not uncommon for researchers to lose access to their own data and work: they may accidentally delete them, misplace them, or lose them due to a damaged hard drive. Alongside visible metadata, I therefore consider it essential to ensure institutional storage of originals, with clear access rules. This achieves what private storage cannot: long-term preservation of materials, sustainable format migration, stable identification through permanent identifiers, and transparent and responsible management of user access.

As researchers, we are often unaware that the value of the materials we collect may far exceed their initial, original purpose. When properly and institutionally archived, the ethnographic records that once served us for analysis, interpretation, and the publication of research results can, in time, become the basis for entirely new

7 The Research archive of the Institute of Ethnology and Folklore Research is an example of good practice, as the collected material is systematically stored, classified, catalogued, and used according to clearly defined rules. Access is available almost exclusively for academic purposes (and for other purposes only with explicit author approval), and decisions regarding use are made by considering the sensitivity of the content, the level of consent, the methodological context, and the relationship between the researcher, the institution, and the community. This model demonstrates that it is possible to protect interlocutors, respect context, and ensure that fieldwork data continue to have a life within the academic domain – not as a "betrayal of trust," but as responsibly preserved research and cultural heritage.

research questions: from linguistic change and discursive patterns to social norms, gender roles, daily life, material culture, or migration dynamics. And not only for ethnologists, as a single record can serve a range of social sciences and humanities disciplines as a solid foundation for comparative and longitudinal studies. As time passes, the same record can become a reference point for new research and new questions. In this sense, the unrepeatability of the ethnographic encounter is not an obstacle, but another argument for the systematic preservation of these materials.

I would also like to highlight one aspect that is often overlooked in discussions on archiving, but which Orlić rightly identifies: the researchers' own reluctance to deposit their records institutionally. Interview recordings do not capture only other people's stories, they also capture our own voices: mistakes, digressions, jokes; moments that arise as we try to establish trust and connect with interlocutors. This self-exposure often produces discomfort, even resistance to archiving, because it confronts us with the question: What will future researchers hear about us? Here, however, I must recall an experience many of us share: as researchers who consult archival materials, we know how recordings from the past can serve as both orientation and inspiration. In listening to the field recordings of Maja Bošković-Stulli, Jerko Bezić, and others, we hear not only their interlocutors but also the researchers themselves: their ways of thinking, reacting, and shaping relationships through conversation. Such recordings are not mere documents, they offer insight into research practice. They assist us in preparing for fieldwork, in conducting comparative readings, and in developing new interpretations. Withholding such knowledge for ourselves would be both unethical and counterproductive. I believe we should consciously temper that aspect of professional vanity: the long-term value of the material, I would argue, outweighs our understandable discomforts. Whether someone hears our voice on a recording decades from now is far less important than the fact that we are leaving behind unique empirical traces as a legacy. Ultimately, such documents do not need to be archived nor published indiscriminately: they can be edited, protected, and classified beforehand, preserving key information while ensuring the dignity and integrity of all participants – both interlocutors and researchers.

One possible solution is the storage of transcripts – although often incomplete and devoid of the contextual richness provided by recordings, they are still better than having no records at all. Yet even this is, for many, not an acceptable option, and the reason is not always a lack of time, but often a lack of willingness. A not entirely justified fear of "data theft" frequently emerges, and this fear is one of the reasons why photographs, videos, and other forms of documentation are also left unarchived. As researchers, it is understandable that we are deeply attached to our materials as much effort, knowledge, methodological decisions, and personal engagement has been invested in them. For this reason, the idea that someone else might use them can lead to discomfort. However, it is important to note that today there are numerous mechanisms that protect research work. Data can be deposited in an archive or repository with a moratorium period or usage restrictions, and the ethical framework for citation and attribution is increasingly well-defined. Just like

academic publications, research data is citable – clearly attributed to their authors and assigned permanent signatures. Responsible data use is founded on rules of source citation, not on silent appropriation. In other words, when someone uses our data, it does not mean we have lost control over them; on the contrary, it can lead to greater visibility, recognition, and acknowledgment of our work.

Naturally, today data cannot be used nor shared without the consent of the interlocutors. Their permission, specifying what they wish to share, in what form, and under which conditions, is a fundamental and unquestionable ethical (and legal) requirement. Without such a framework, discussions about openness in humanities research are meaningless. Consent, however, cannot be reduced to a single signature at the beginning of fieldwork. It is a process that spans the entire lifecycle of the data: from collection and transcription, to storage and sharing, and finally subsequent analyses and interpretations. Ethnographic records are not neutral "data," but traces of relationships, thus their use entails responsibility toward the people and contexts from which they originate. Within this framework, consent entails the possibility of revision and withdrawal, transparent documentation of changes (versions of consent), clearly defined access regimes, and monitoring of reuse. Only when conceived and documented in this way does consent cease to be a mere formality and become a tool of trust and shared responsibility. However, such a practice requires professional support to be feasible: clearly written forms and guidelines, consultation with data protection officers (GDPR) and/or data stewards, as well as basic training for researchers in ethical data management.[8]

Finally, before signing consent forms, I consider it important to clearly communicate to interlocutors that they are co-bearers of knowledge and co-creators of academic heritage. Recognizing their role in this way often makes them more willing to approve the deposit of materials, even when it includes precisely defined and restrictive conditions of future use. This certainly requires additional effort from the researcher in the phase of explanation and conversation, but it is the best way to reduce discomfort and fear related to granting consent.

In conclusion

I am well aware of how complex the path toward responsible openness is in our discipline: too little time, too much administration, uneven infrastructure, and constant doubts surrounding rights and ethics. Moreover, working on data is rarely truly recognized, even though we cannot fully contribute to open science without it. Finally, ethnographic research carries an affective dimension: no matter how much we try

8 In line with the idea of shared responsibility, the need for an infrastructural solution that supports ethical research practice has emerged. As the head of the Research archive, I often observed uncertainty among colleagues when preparing consent forms. Within DARIAH's working group ELDAH (Ethics and Legality in the Digital Arts and Humanities), we therefore initiated the development of the Consent Form Wizard (CFW): a digital guide for the consent process. The tool helps researchers design and document consent, adapt it to different contexts, and align it with legal standards (GDPR), and it also allows for further upgrades and the development of new scenarios (https://consent.dariah.eu/).

to act according to professional standards, trust, discomfort, and responsibilities toward our interlocutors shape our decisions about sharing and limiting access. This is why data sharing in ethnology and cultural anthropology is not a mere technical task, but a complex ethical and epistemological undertaking.

If applied without an understanding of the discipline's specificities, open science can easily slip into a form of administrative coercion rather than an incentive for the responsible dissemination of knowledge. But if we understand it as a call to collective care for knowledge – not only the knowledge we produce in the form of presentations, articles, and books, but also the knowledge we store, share, and leave to future generations in the form of data – then it becomes an act of scholarly solidarity rather than something imposed upon us.

The key question is therefore not "for or against," but *how* to be open: under what conditions and according to which values. That answer cannot come solely from "above" from academic policy, but from our own community: from our self-reflection, established practices, ethical criteria, and actual infrastructural capacities. Ethics and openness are not opposites. "As open as possible, as closed as necessary" in practice means planning, contextualization, and clear mechanisms of access management (access levels, time limitations, licensing). Thoughtful and sustainable openness can be built on these considerations, with the recognition that the work of preparing and publishing data is, in itself, a valuable academic contribution.

From this perspective, open science is not only a matter of access but also of care: Who actually preserves the data, prepares it for future use, documents it, and protects it? As long as these questions are not embedded into institutional systems (assessment, promotion, and funding criteria), openness remains a privilege of those with resources and a burden for everyone else. This is why infrastructures such as DARIAH are important: not only as technological platforms but also as shared spaces of collective reflection – places where values are articulated, models of collaboration are shaped, and ethically sustainable practices of sharing knowledge are developed.

Open science in the humanities must be situated, reflexive, and iterative, just like the ethnographic process itself. In such a framework, open data becomes a lasting research asset under responsible stewardship, and we can say with a clear conscience that we have not only produced knowledge, but have also made provisions for its durability, integrity, and meaningful future use.

In this comment, I have deliberately refrained from commenting on the recently published Croatian Plan for Open Science, as well as the questions raised by the rapid expansion of artificial intelligence; their complexity requires a paper of its own. I can thus make a concluding remark from my own standpoint: open data in the humanities make sense only if they stem from care for our interlocutors and from systematic, responsible storage.