

# RDA Adoption story

## Your Organisation

---

### Organisation/Entity Profile:

*5-10 lines short profile of your organisation*

#### **Platform for Experimental, Collaborative Ethnography (PECE).**

The Platform for Experimental, Collaborative Ethnography (PECE) is a Free and Open Source (Drupal-based) digital platform that supports multi-sited, cross-scale ethnographic and historical research. The platform links humanities and qualitative social science researchers in new ways, enables new kinds of analyses and data visualizations, and activates researchers' engagement with public problems and diverse audiences. While supporting a number of specific projects in the "empirical digital humanities," PECE is also a research project in itself, exploring how digital infrastructure can be designed to support open-ended, collaborative hermeneutics. PECE has been built and is governed by an interdisciplinary design group centered first in the Science and Technology Studies Department at Rensselaer Polytechnic Institute (Troy, New York, USA), and now in the Department of Anthropology at the University of California, Irvine.

### Contact reference

*Photo, Name, surname, qualification, affiliation, country, email and phone of the contact person reporting the adoption story*

**Contact reference: Lindsay Poirier, Department of Science and Technology Studies, Rensselaer Polytechnic Institute, Troy NY 12180 USA; [poiril@rpi.edu](mailto:poiril@rpi.edu)**

**Primary author of this document and technical consultant for adoption described here: Dr. Luis Felipe R. Murillo; anthropologist; Institut Francilien Recherche Innovation Société (IFRIS), Paris, France; [lfmurillo@cnam.fr](mailto:lfmurillo@cnam.fr)  
Phone: +33 0769442344**

### Organisation Logo

*Note: please send a thumbnail high resolution logo (best size 500x500)*



PECE

Platform for Experimental and Collaborative Ethnography

## Adoption story brief overview

---

*Please provide a short description of your adoption story (100-150 words)*

In February 2015, the PECE design group started to implement the recommendations of the Research Data Alliance's (RDA) "Practical Policies" Working Group (WG-PP) on data management into a digital platform designed to support collaborative ethnographic research projects. Our goal was to implement the RDA best practices in data management to 1) protect and anonymize (often sensitive human subjects data), 2) enable, enrich, and incentivize data sharing in our research community, which has traditionally not shared primary data, and 3) to ensure that the data stored in our systems can be contextualized with metadata about its creation and complex chains of custody. In total, 11 practical policies for data management were adapted and adopted by the PECE project and incorporated into our digital platform.

## The challenge

### What was the challenge that you addressed?

*Please describe the scientific, research or technological challenge/s that you and your team are trying to solve with the help of RDA  
(100-150 words)*

PECE was designed to address the need for collaborative ethnographic research and interpretative data analysis. RDA helped our project by providing the technical guidelines for implementing a set of data management policies which speak to the specific needs of ethnographers in the context of "digital humanities." In particular, we tackled three interrelated problems with support from the RDA community: 1) development of a web platform supporting interdisciplinary collaborations (with proper data management capabilities built-in) meant to address the issue of scarcity of adequate institutional repositories for the humanities and social sciences; 2) specification of a flexible, basic data model for ethnographic projects which allows for large-scale data exchange (with proper metadata description) across ethnographic collections; and 3) translation of open standards and best practices from Free and Open Source communities as practical policies for our digital platform, including provisions for open formats, open licenses, multiple authorship, and open interfaces for automated data management.

### Which RDA recommendation or output did you adopt?

*Please indicate which output you chose to adopt and why  
(100-150 words)*

The PECE design group has been working in collaboration with many interest and working groups, but it has only adopted the "Practical Policies" Working Group (WG-PP) output on "Data Management" thus far, given this is the most basic output for our purposes. The WG-PP surveyed policy sets, required by various institutions, that control the execution of data management - policy sets that maintain data collection properties such as authenticity and chains of custody, and that automatically execute actions to control data retention, back-up, and integrity. The working group's recommendation detailed how 11 practical policies could be organized into an administrative workflow within research projects and designed templates that documented how the policies could be automatically triggered in digital platforms. The PECE design team adopted the output so data management would be automatically built into the digital platform we have designed to support collaborative, ethnographic research.

### **What are the potential consequences faced without a solution?**

*Please explain what consequences would ensue if a solution (RDA or any other) is not found (100-150 words)*

Many web platforms are designed and implemented without proper data management provisions. This leads to future difficulties, when third party solutions have to be integrated to perform data management tasks, complexifying the workflow, introducing potential points of failure, and expanding the attack surface for security breaches. Further, humanities data poses considerable specific data management challenges. Data in the humanities is often particularly sensitive, including interviews and observation of “human subjects.” Access control settings need to be particularly strong and often complex to maintain confidentiality of research subjects when consent to publicly share the data is not granted. In hermeneutic traditions like the one in which we work, data is often interpreted and then re-interpreted by a variety of different researchers; data ownership and chains of custody are thus particularly important and challenging to track. The PECE design group was fortunate to find the RDA output on “practical policies for data management” early on in its design and implementation phase, which allowed for integrating data management provisions as integral part of the platform.

### **How did you identify RDA as a potential supporter?**

*Please state how you got in touch with RDA and identified its recommendations and outputs and how did you decide to go into the testing phase. (100-150 words)*

The founding members of the PECE design group have been part of the RDA community for many years: two of them, Mike Fortun and Kim Fortun are the original proponents of the “Digital Practices in History and Ethnography” (DPHE) interest group (with Jason Baird Jackson), for example. From their early experiences with RDA, we found the opportunity to become an early adopter of the WG-PP for data management. This is when Luis Felipe Murillo became involved as a technical consultant, helping to translate Free and Open Source development approaches into the design and implementation of a web platform for collaborative and experimental ethnographic research. The space of exchange of the DPHE has been very important during the design, implementation, and testing phases of PECE. Another group that has been important to this work, given the crucial need of incorporating security provisions into the platform due to the sensitive nature of ethnographic data, is the interest group on “Ethics and Social Aspects of Data (ESAD).”

### **What is your stakeholder community?**

*Please provide a short description of the community (scientific but not only) that will benefit from your work, how big it is and where it is localised? (100-150 words)*

PECE supports experimental ethnographic projects that are very attentive to the dynamics of language, history, and power in the study of contemporary socio-technical phenomena. It grew out of a collaboration involving students and researchers in various disciplines whose primary interest was to study science, technology, and governance from an ethnographic standpoint at Rensselaer Polytechnic Institute (Troy, New York, USA). The PECE stakeholder community includes geographically-dispersed historians, anthropologists, science and technology researchers and students, and their research co-participants. Researchers in our community both collect and produce

RDA Adoption Story template – July 207

data (such as interviews, field notes, and archival documents) and then interpret the data according to diverse genealogies of humanities and social science theory. To address emerging, complex techno-scientific problems, our research community has only recently begun to advocate for data sharing - not necessarily so that researchers can *reproduce* analysis, but instead so that data can be interpreted from diverse perspectives. This is the primary target group of PECE users and developers, but the platform can be easily adapted to work for different research groups and historical ethnographic archives, given its flexible nature.

### *Only for completed adoption processes*

Please leave the fields below empty if your adoption case hasn't been completed yet and no impact can be measured so far.

## The RDA output adopted

---

### **Benefits of adoption and impact**

*Please describe how the community that you are addressing will benefit from the adoption of the selected output/recommendation and which is the foreseen impact of this activity (100-150 words)*

After the adoption process, we launched a first public version of the platform for download on Github ([github.com/pece-project](https://github.com/pece-project)). In addition to the public release, we also released a complete documentation of all the aspects of the platform with respect to its data management capabilities ([pece.readthedocs.io/en/latest](https://pece.readthedocs.io/en/latest)). This release has the potential of impacting collaborative research in the humanities and social sciences, animating different forms of data sharing and analysis: we designed and implemented various features which facilitate the work of curating ethnographic data in a collaborative mode (such as providing basic metadata descriptions, automated and encrypted backups, public API for data harvesting or data sharing across PECE instances, persistent storage of URLs with Perma.cc addressing, among several other capabilities for collaborative work among ethnographers).

### **The adoption process**

*Please describe the adoption process that was put into place by your organisation. A detailed description is preferred, including timing, resources involved, methodologies adopted, etc... (250-300 words)*

The adoption process involved an initial assessment of the existing state of the platform (its existing documentation and implementation) and available expertise in the group (to distribute the load of technical tasks). This assessment was accompanied by the evaluation of Free Software implementations we could rely upon, the description of functional and technical requirements as elaborated by the design group (with respect to the "design logics" which guided the software implementation), and creation of a timeline for implementation and testing.

During the design phase, we evaluated the templates detailed in the output in relation to our own community's commitments and needs. We aimed for a balance between the necessity of preserving privacy and anonymity and the need for creating conditions for data sharing and collaborative

analysis among ethnographers, historians, and co-participants of our projects. We designed our own data management policies by discerning the relevant components in the working group's templates and adapting them 1) to meet the specific needs of our research community and 2) to work with the Drupal framework.

In practical terms, we relied on many extended features of the Drupal framework in order to implement the data management policies. To aid the process of implementation, the PECE design group hired two Drupal software companies, Taller and Agaric. The hiring process was only finalized after several estimates were collected with specialized companies. The project was divided, then, in four phases: the first assessment phase took three months, and the subsequent three phases were carried out in the span of one academic year, culminating in the public release of the platform on time for the RDA "Seventh Plenary Meeting" in Tokyo on March 1<sup>st</sup>, 2016. For the software development project, the agile methodology was used. As of August 2017, all of the policies have been documented in our data management plan and all but one of the policies has been implemented in our digital platform.

### **Which technical challenges did the output solve?**

*Please explain what technological needs or other requirements the selected output helped to address. (100-150 words)*

The "WG-PP for Data Management" provided a complete description of the basic needs for our platform: it summarized the state-of-the-art for large-scale data management solutions, which allowed us to make careful decisions about the needs of our community. One of the biggest challenges was to design around the fact that our disciplines in the humanities and social sciences often do not have the available technical expertise to take the tasks of data management into their own hands. The output of the "WG-PP" provided us with necessary information to automate and simplify the process of data management for members of our community.

## **What advice would you give?**

---

### **What lessons did you learn from the adoption process?**

*Please provide your feedback on how this process could be supported, what could have been done to facilitate more the effort, etc. (100-150 words)*

Adopters would benefit greatly from more substantial feedback on their implementations early on: following-up with adopters regularly might help speed-up the process of adoption (as well as early identification of pitfalls and difficulties in the implementation phase). Further, this follow-up could serve to highlight areas where the recommendation may need to be made more specific or more flexible to meet the specific needs of certain domain groups. We would recommend that RDA create mechanisms for following-up with adopters in a more regular, systematic fashion.

### **Based on your direct experience would you recommend the same adoption?**

*Please provide insights that you feel other adopters would benefit from know in advance of adoption (100-150 words)*



[www.rd-alliance.org](http://www.rd-alliance.org)

At RDA, adopters will find the necessary guidance from experienced data professionals but the capacity to operationalize their guidelines will depend upon your local capacity (available expertise + resources). Having a good working knowledge of your environment will certainly expedite the adoption process, making it less demanding for designers, developers, and project managers overall.