Wide Field Ethnography and Exploratory Analysis of Large Ethnographic Datasets

David Socha  
University of Washington Bothell  
Bothell, WA 98011, USA  
socha@uw.edu

Alfredo Jornet  
University of Oslo  
Postboks 1092 Blindern  
0317 Oslo, Norway  
a.g.jornet@iped.uio.no

Robin Adams  
Purdue University  
West Lafayette, IN 47907, USA  
rsadams@purdue.edu

Abstract  
With inexpensive access to recording technologies, CSCW researchers are collecting increasingly rich and complex sets of unstructured and multi-modal ethnographic data. Increasingly often, those materials are collected in highly cross-disciplinary contexts that require diverse backgrounds and levels of expertise for their analysis. How can we take advantage of our increased capacity for collecting rich ethnographic materials? What are the affordances and limitations of such large and complex datasets? How can we enable teams of researchers from diverse disciplines to jointly navigate, recover, and analyze the emergent social phenomena constitutive of the settings recorded? These are questions we are exploring through Wide Field Ethnography (WFE), an emergent approach based on intensive data collection and cooperative inquiry. In this position paper we present its principles, some of the questions it raises, and some of the lessons learned.

Author Keywords  
Cooperative inquiry; ethnography; multimodal.

ACM Classification Keywords  
K.4.3 Computing Milieux: Organizational Impacts.
**Introduction**

Research in the field of CSCW has been increasingly concerned with collecting rich ethnographic datasets that often include video and other multimodal materials \[3\]. With the possibilities that inexpensive video and other technologies offer, these ethnographic datasets will grow both in the amount and types of materials collected. This poses challenges to the majority of research on large datasets, which focuses on automatic and quantitative data analysis, usually of a single type of data (e.g., text-based messages), with the goal of being able to analyze the entire corpus.

In this position paper, we identify topics for a discussion on an alternative and complementary approach that aims to capture and preserve the richness and detail of people's collaboration in natural settings, while making it possible for cross-disciplinary teams of researchers and practitioners to gather, analyze, and explore the resulting large and complex ethnographic datasets. We use the term *Wide-Field Ethnography* (WFE) to refer to the set of cooperative practices we are developing for this. WFE uses multiple video cameras, microphones, screen captures, data loggers, and other sensors to intensively collect ethnographic materials that provide a perspective as wide as possible on events of cooperation.

WFE's analytic approach contrasts to other automatic and/or quantitative approaches that reduce complexity by focusing on already known or well established categories (lower-right corner, Figure 1). Instead, WFE's analyses are designed to draw from the contextually rich particulars of the settings to enable multiple types of research. But, can an ever-increasing level of detail in the data collection bring us closer to the actual practices of the settings, which are complex and emergent (upper left and right corners, Figure 1)? Or does this level of detail lead to (analytical) chaos (lower-left corner, Figure 1)? For cross-disciplinary research projects, does cooperative inquiry help or does it make chaos even more likely?

**WFE Cooperative Inquiry**

To better understand the nature and potential of the WFE approach, we recently started a cross-disciplinary and international cooperation to jointly explore the *BeamCoffer dataset* \[5\]. This dataset was collected over an 11-day period from a software development organization in the Seattle area by using nine GoPro cameras, six high-quality Zoom H2n audio recorders, and screen capture software. The six-terabyte dataset includes 380 hours of video of software developers pairing at developer stations, doing standup meetings, and working at the whiteboard (see Figure 2).

The large amount of video, audio, time-lapse, and other data in the BeamCoffer dataset offer a range of possibilities for both individual and cooperative inquiry. At the same time, however, the size and complexity of this dataset makes it difficult to navigate, browse, filter, annotate, analyze, and otherwise use to identify moments of theoretical and practical interest. This difficulty is exacerbated since the researchers inquiring into BeamCoffer come from a variety of disciplinary backgrounds, and have different backgrounds with respect to software development.

In the sub-sections below, we identify three areas that we are exploring through our collaboration to address the affordances and challenges described above, and which could serve as starting points for contributions.
Cooperatively building indexical tools

The goal of WFE datasets is to preserve the contextually rich particulars of the settings investigated - talk, gestures, gaze, body orientations, inscriptions, interactions with materials and computers, and all those other features that people routinely make public to each other to carry out work and maintain social order and cooperation. These features exist as unitary wholes for the participants, who experience them seamlessly and in relation to their history of interactions and knowledge of the setting. However, the data collection process fractures the system being observed into streams of information gathered by disparate sets of data collectors.

One of the aspects we are working with involves our need for tools to re-assemble the dataset into structured fields that enable sufficient coherence for us to share a focus of inquiry, while supporting sufficient diversity to allow for flexibly responding to different kinds of research questions. Existing analytical tools such as Nvivo, Transana, ELAN, or Atlas afford useful cataloguing and coding features, but are limited in their capacity to flexibly respond to situations in which emergent flows of cooperation need to be re-structured from disjointed data materials. To remediate this, we are developing cooperative ways of indexing the multiple layers of contextually-rich information, enhancing the dataset by constructing new visualizations and notations using existing and newly created software tools. These include flexible shared spreadsheets that are jointly created and expanded as we identify common threads and/or gaps in the materials, identifying key time-stamps for synchronizing across media, and linking otherwise diverse events into coherent trajectories. We are also building rich visualizations that combine spatial, temporal, and biographical clues to further enhance our possibilities to explore those identified trajectories.

Human-centered design of visual analytics tools

Another area we are exploring concerns the use of enhanced techniques to analyze pictures and video. We are learning that visual analytics will be central to WFE practice, and are taking a human-centered design approach to develop a set of WFE principles, practices, and tools that allow researchers to navigate, filter, annotate, analyze, and share results. We are using the Theory of Constraints design cycle [2] to iteratively identify, develop, deploy, and evaluate our WFE ecosystem: 1) gather a Collective Analysis WFE dataset of us collectively analyzing BeamCoffer; 2) observe this dataset to document workflow and identify important bottlenecks; 3) create an intervention to reduce or eliminate a bottleneck; and 4) deploy the intervention, and return to step 1 to evaluate the intervention and uncover the next bottleneck to address.

Based upon Walter’s [8] experience applying the above approach for creating the Attenex eDiscovery platform, we are exploring how to design tools to enable cooperative decision-making, making visible salient aspects of the system being analyzed. We consider the need of automated techniques that can continually troll, uncover, and represent a dynamic set of connecting concepts or topics to the researcher.
Reflexively Studying Cooperation

Finally, we are studying our own cooperative practices as a means to understand BeamCoffer’s cooperative practices. This approach grows from theory and our experience in WFE. Theoretically, we draw from ethnomethodology’s insight that the possibility of making accountable phenomena of social order observed in our database presupposes our own competence to draw from and produce those very phenomena [1]. In our cooperative analysis, we have several times found ourselves reproducing some of the phenomena that we were observing in the BeamCoffer dataset. These experiences are allowing us to explore aspects of human cooperation that would not have become apparent if we had been working individually. Accordingly, we are recording our co-located and online meetings and shared data sessions, turning them into further analytical resources, and helping us to jointly pose and test hypotheses about cooperation.

An ongoing journey

In this paper we describe WFE as a set of cooperative inquiry practices for the study of computer supported cooperative work that retain the richness of qualitative techniques in data science. The WFE tool ecosystem is continually evolving – supporting the dynamic and flexible use of whatever tools become available for the emerging purposes of the cooperative inquiry. Our brief experience suggests that WFE affords researchers to follow work across space, time, and modality [7]. They can employ a wide range of units of analysis to follow the dynamic nature of the work [6]. Researchers can more easily find data that disproves their hypothesis or change their research question in response to emergent constraints or opportunities in the data [7]. As we continue our nascent explorations of WFE’s affordances and limitations, we wonder how WFE might differ from participatory ethnography, and whether and how cooperative inquiry’s results are substantially more than the sum of the individuals involved.

References

2. Eliyahu M. Goldratt. 1990. What is this thing called Theory of Constraints and how should it be implemented? North River Press, Great Barrington, MA.