Complete reference:
Mikalsen, Marius. A Case Study of an Information Infrastructure Supporting Knowledge Work in Oil and Gas Exploration. I: COOP 2014 - Proceedings of the 11th International Conference on the Design of Cooperative Systems, Nice, France 27-30 May, 2014. Springer 2014 ISBN 978-3-319-06497-0. s. 121-136

The original publication is available from Springer: http://link.springer.com/chapter/10.1007%2F978-3-319-06498-7_8

A Case Study of an Information Infrastructure supporting Knowledge Work in Oil and Gas Exploration

Marius Mikalsen[1]

Abstract It is well rehearsed in the fields of CSCW and IS that the relationship between the social and the material is bi-directional and shaped locally. But what happens when knowledge work is stretched across space and time, and the practice of today relies on actions and reflections done elsewhere and at different times? This paper presents an on-going case study of oil and gas exploration that takes steps to shed light on this emerging issue. I argue the relevance of framing the process of generating *interpretations* in oil and gas exploration in terms of *information infrastructures*. The case is representative for other cases where practitioners´ reflections cannot immediately be confirmed by empirical observation. Through a discussion on the concepts of *coordination* and *accumulation* across the dimensions of space and time, I outline how an able information infrastructure in this domain must balance the dualism of the concepts of *naturalisation* and *historification*.

1 Introduction

In the beginning there were seeps, and oil and gas exploration was straightforward. Every major petroleum-bearing basin of the world has numerous oil seeps where oil seeps naturally to the surface. Explorers focused their search on areas near seepages, where oil bubbled up to the surface naturally (Hyne 2001). Now however, the easy finds have been done. Given still increasing demand, and high prices, oil and gas companies must explore in areas that are difficult to reach, such as several kilometres below the seabed. New discoveries are made possible by combinations of new exploration methods (human interpretations) and new technology. Still today however, drilling an exploration well is the only certain way to confirm the presence of hydrocarbons deep down in the earths crust. But drilling is

---

[1]M. Mikalsen

Norwegian University of Technology and Science (NTNU) and SINTEF

7491 Trondheim

e-mail: marius.mikalsen@sintef.no

difficult and expensive. The floating rigs are dipping on the surface, drifting kilometers, while trying to hit the reservoir five kilometers below sea level, and perhaps 2 kilometers to the side. The cost of running such a floating rig can be up to 500.000 USD per day (day rates for each rig type drawn from the RigLogix database, accessed 06/11/2013), and one exploration well can cost tens of millions of USD and upwards. This is in addition to the environmental risks involved in drilling. Drilling such a well therefore is a process O&G companies only will do when the probability is *high enough* that there is a presence of technically and economically recoverable reserves.

This paper brings forward a case of modern O&G exploration and shows how it is a sociomaterial (Parmiggiani & Mikalsen 2013) and highly information-centric process, involving a complex combination of information systems, large amounts of heterogeneous data (*the material*), and several teams cooperating across space and time (*the social*). This sociomaterial assembly can be seen as an *information infrastructure (II)*, a concept originating in science and technology studies. One of the early definitions (Star & Ruhleder 1996) explains an II consisting of these "*dimensions*"; embedded (inside of other structures such as social arrangements and other technologies), transparent (supports tasks invisibly), reach & scope (across space and time), learned as part of membership (of a community of practice), links with conventions of practice (II shapes and is shaped by practice), embodies standards (plugs into other IIs and tools using standards), installed base (built organically, not revolutionary), becomes visible on breakdown (normally invisible, but very noticeable when it breaks down). The notion of II has evolved from Star and Ruhleder´s focus on sharing learning within and across communities, through focus on how the II is shaped locally to a current focus on how there is a tension between the local and the global due to lack of global control, and is currently defined as "*a shared, open (and unbounded), heterogeneous and evolving socio-technical system (which we call installed base) consisting of a set of IT capabilities and their user, operations and design communities"* (Hanseth & Lyytinen 2010). Mayernik et al explains how the same concept of II is also increasingly used in Computer Supported Cooperative Work (CSCW); "*Coordinating technology development and scientific research is a growing theme of Computer Supported Cooperative Work (CSCW) studies, with many open questions"*, particularly in e-science, where "*large scale distributed computational, data and communication, infrastructures and middleware*" is applied to emerge new kinds of scientific practice (Mayernik et al. 2013).

This paper address the following research question: What are the characteristics of the social and the material parts of the II that cooperate and co-produce *interpretations* of the subsurface materiality across space and time?

In answering this research question, this paper serves three purposes. First, I take steps to "unearth", to use Mayernik et al´s (ibid.) term, the II in oil and gas exploration. Understanding the nuances in as many particular cases as possible is crucial for the understanding of II as a phenomenon, because it is so interwoven with coordinative practices that "*are not generic but domain specific"* and *"in or-*

der to develop the conceptual foundation for such technologies, in-depth studies of professional work and the concomitant coordinative practices are critically important." (Schmidt 2011). Taken together, ethnographic accounts can support generalisation (Crabtree et al. 2013) that can inform II design and evolution, see e.g. (Hanseth & Lyytinen 2010). Second, I make an argument that there is a need to not only address capabilities, applications and platforms in-situ, but also focus on accumulation of data across spatiotemporal dimensions of the II, and how this influence collaborative work and coordination. To do this, I will draw on literature from the domains of information systems (IS) and CSCW. I will discuss the O&G exploration process as an on-going naturalisation process of interpretations that is constructed collaboratively. I will show how interpretations are shaped and moulded over "*different contexts over extended periods of time*" (E. Monteiro et al. 2012), or to frame it otherwise, through "*asynchronous remote collaboration*" (Cabitza & Simone 2013), involving several geoscientists and different geoscientific disciplines as "*competent actors reflecting*" (Prilla et al. 2013) on historical interpretations, continuously refining them. An obvious challenge here is to locate relevant data in a vast amount of data, a process that is performed by people and multitude of search and database tools. Equally important however, but easier to overlook, is to "*make the invisible naturalisation visible*" (Bowker & Star 1999), that is, to support the need of the reflective practitioners to, in a sense, interpret the interpretations done elsewhere and in the past. Third and finally, I will discuss these concepts towards contributions from the domain of healthcare. The reason for so doing is that exploration for new subsurface petroleum resources, similar to that of health, is a case where collaboratively built representations, residing in information systems, are the primary sources upon which practitioners must reflect and make decisions, effectively deprived of the possibility of immediate empirical verification of their reflections and interpretations.

The rest of the paper is organised as follows. In the following method section I explain the context of this study, how data is generated, and how it is analysed. Section 3 analyses and frames oil and gas exploration case as an information infrastructure, shows what an interpretation is in this domain, the inscription devices used in the infrastructure and how they are used for accumulation and coordination. Section 4 discuss, compares and contrasts the notions of accumulation and coordination from CSCW healthcare works with this case and how the information infrastructure as a consequence need to balance the dualism between naturalisation and historificazion. Finally, in section 5, I summarise and suggest some interesting vistas for continued work on this theme.

2 Method

Study context

NorthOil (an acronym) is an international oil and gas company established in the 1970s and headquartered in Northern Europe, currently employing 30000 people with activities in 35 countries.

The unit of study is exploration. Oil and gas exploration is the initial phase in petroleum operations that includes generation of a prospect and drilling of an exploration well. The key questions NorthOil wants answered are: *i*) where are the hydrocarbons, *ii*) how much hydrocarbons are there? *iii*) Will it flow? i*v*) what is the best way to produce it (bordering to the disciplines of field assessment and production). Summarized the task at hand for exploration is to answer the ultimate question of: Can this reservoir (when identified) profitably produce hydrocarbons?

In NorthOil exploration, geoscientists like e.g. geologists and geophysicists (I will use "G&G personnel" or simply "**G&G**" as a collective term for this group, in fact, in NorthOil all geoscientist are collectively named G&G even if they are geochemists for example) are continuously interpreting the company's vast amount of subsurface data. Geologists know rocks and the formations they make in the earth crust. Geophysicists can look at physical characteristics, such as magnetics and gravitation to indicate rocks and formations exist below. Other geoscience disciplines are also supporting the interpretation process, such as geochemists, that study chemical elements in rocks and minerals and the movement of such elements into soil and water systems. Creating different kinds of interpretations is a multidisciplinary effort including different kind of geoscientific disciplines.

There are two thousand G&G in NorthOil. Two hundred **Project Data Managers (PDMs)** support the G&G with quality assuring, storing, and retrieving data. **Central Data Managers (CDM)** are responsible for maintaining the company datastores and keeping data synchronized across datastores. This is necessary to deal with the complexity of the II in terms of storing, accessing and managing all the geological and geophysical information that has accumulated over time.

G&G and PDMs are organized in units based on what geographical areas they are exploring. Currently these areas are North America, Norway, International west, and International East. The unit I am studying is Norway, and it is further divided into Licenses North Sea, Licenses North (Norwegian Sea and Barents Sea), and Access projects. The first two are more mature areas where there are already established commercial and technically viable fields and production. Access projects explore new fields and prepares for license rounds. The team I am following is working on North Sea licenses that are areas where NorthOil is presently partner in a license and need to find and assess the best places to extract hydrocarbons in the license. The team is located in an office building and sits in landscapes, where two to five people share offices. They have PCs (windows workstations), typically with more than one screen (which is handy when looking at subsurface models and geographic information systems - GIS).

Data collection and analysis

This article draws on empirical data from an ongoing study of NorthOil exploration and reports on a 6 months field study.

I have so far been present there in 5 different periods for 14 days total (periods ranging from one to four days), combining data collection activities such as participatory observation, informal chats, and interviews. The case study is ongoing.

A summary of the data sources foreseen is given in the table below.

Data source	**Examples**	**Collection**
Semi structured interviews	Interviews with exploration geologists, geophysicists and geochemists, exploration PDMs, CDMs	8 audio recorded.
Participant observation	Group meetings, project meetings, informal meetings, chatting during lunch.	20 pages of machine written notes.
Document analysis	Official portal, internal documentation, email communication, plans, strategies, training material.	Stored digitally.

Table 1 Data sources.

It is a large dataset and it is unstructured. Langley outline a number of different strategies for making sense of such process data (Langley 1999).We need not to choose one, she explains, but we must be aware of the strengths and weaknesses of each strategy, and the potential they have to build theory that is "*accurate, parsimonious, general, and useful*" (ibid, p 691). In the analysis I work close to the empirical material, the texts that has been generated from notes, interviews, presentations and documents.

3 The Oil and Gas Exploration Information Infrastructure

The overall process of exploration has the following steps. First there is a global basin screening and ranking to determine what basins to work with. Following this, basin and prospect evaluation is done. Here the goal is to identify prospects for drilling. The prospects are quality controlled, approved and ranked. Exploration drilling is done and the discovery is evaluated. Finally, the feasibility of the prospect is appraised. The process can take several years from beginning to end, for instance it can take years from a successful prospect evaluation until exploration drilling is actually conducted, due to constantly changing priorities. This means also, that the fact check of the interpretations done in preparation of a prospect, will come in a final drill report, potentially several years after they were originally done.

In all of these phases different kinds of interpretations are done (the following section detail the notion of interpretations). All of these phases are highly data intensive. Typically the G&G will need an overview of all relevant data in an exploration area defined within quadrants and blocks. Quadrants are 1 degree by 1 degree, and blocks of 15 minutes of latitude by 20 minutes of longitude (12 blocks in a quadrant). A request for well data could be formulated as "*within quad 35, find me all wells which have total depth of more than 3000 meters"* (from interview with access project geologist).

There is a lot of data available. In the corporate data store alone there is in excess of 110TB (2011 numbers). Data is brought to G&G primarily from two main sources, Diskos DB and the corporate data store. The Diskos DB is a national database where all oil and gas companies operating on the Norwegian continental shelf are required by law to store all of their seismic, well and production data. Searching for public data (such as production licenses, exploration wellbores, discoveries, fields, development wellbores and business areas) is available through a Public Data Portal[2]. The corporate data store holds NorthOil´s proprietary interpreted data, such as seismic interpretations, well data, production logs, and maps. Business information such as license areas, infrastructure and business associates are also included. The database has thousands of tables and attributes

Northoil's Exploration II involves a lot of different people, roles, processes and technology. There is an obvious need of NorthOil to define proper processes and work flows for this. Just considering the complexity of the systems alone, a PDM says when asked for an overview of the systems they operate: "*I will try to find a list, but it is a challenge we have, there are too many lists"* (from field notes). There are so many systems that it has resulted in too many list trying to give overview of the systems. A CDM confirms the concern of complexity "*I wonder how much complexity an organisation like NorthOil can handle*" (CDM informant, from field notes).

Interpretations

Geologists and the geophysicists work primarily on two categories of data, seismic data and well data respectively. From seismic data, a structural model is

[2] Available here http://www.pdp.diskos.com (accessed 09/10/2013)

interpreted (mostly done by geophysicists). From the well data, the physical properties of the rocks are interpreted (mostly done by geologists). By combining the structural model with the physical properties, a higher order model is created, a more accurate interpretation if you will. The seismic data is rather coarse grained and covers large areas that cannot otherwise be measured in detail. Properties that can be found in seismic data are; horizons, faults, structure, salt and other bodies, amplitude anomalies, fluid presence, traps and rock properties. For the well data there are lots of precise measurements, but rather sparse areal coverage. From the well you can get several kinds of petrophysical and geo-mechanical data. This data is then interpreted geophysical and geological.

Geophysical interpretations: seismic interpretation and well correlation

In geophysics, *seismic interpretation* is the analysis of seismic data to generate reasonable models and predictions about the properties and structures of the subsurface. Interpretation of such seismic data is the primary concern of geophysicists.

Well correlation matches rock layers between wells "*When a well is drilled a record of the rock layers in the well is made on a well log. The rock layers between well logs are correlated to make a cross section. The correlation is started with a marker bed or key horizon. The marker bed is a distinctive rock layer that is easy to identify. Volcanic ash layers; thin beds of coal, limestone, or sandstone; and fossil zones are good marker beds. A key horizon is the top or bottom of a thick, distinctive rock layer. After correlating the marker bed or key horizon, the rock layers above and below the marker beds can then be correlated on physical similarity and their position in the sequence of layers.*" (Hyne 2001).

Geological interpretations: time-depth conversion, picks and gridding

Since seismic data are recorded in seconds (time domain), and well log is recorded in meters or feet (depth domain). This makes the vertical scales on each different and cannot be directly compared, which is what you want to do in the correlation as described above. If you know the seismic velocity through each layer, you can do a *time-depth conversion* on the seismic data to make it compatible with well-log data (Hyne 2001). This process is iterative, in that it that begins with seismic processing (as described above), seismic velocity analysis and study of well data to refine the conversion. In order to improve the conversion, acoustic logs, check-shot surveys and vertical seismic profiles can improve correlation of well logs and drilling data with surface seismic data.

The typical seismic refers shows structure of the rocks and their characteristic layering, but it does not show individual sedimentary rock layers or the rock type (Hyne 2001). To increase the value of the seismic record, G&G need to identify the individual sedimentary rock layers, so that potential reservoir rock (hydrocarbon producing) and seals (trap rock) can be found. This is done by running the seismic lines through wells that has been drilled, using well log and physical samples as basis for identifying subsurface rock layers in the seismic record.

The well log is in itself also interpreted by mapping the physical measurements and samples to rock formations found in different sources. *Picks* are done, which is the interpretation of data, such as creating seismic section based on marker beds or geologic picks, such as formation tops interpreted from well logs to improve interpretations.

Gridding is the process of determining values for grid elements on a map. The grid element values are chosen from nearby data points ("picks"). Gridding is usually applied to one characteristic per map, such as structure, thickness, porosity, permeability or saturation.

Inscription for coordination and accumulation in the II

Two applications are key in the information infrastructure are used by the G&G to create the interpretations we have seen in the above section. This is Petrel (from Schlumberger) and Openworks (from Landmark). These are so called "*interpretation tools"* that is, the tools where G&G access the data (loaded from the datastores by the PDMs), do their interpretations, and store them. There are historical and political reasons for why there are two different tools originating in a company merger some time ago. Petrel and Openworks are different in many regards, for instance Openworks runs on Linux and is accessed as a virtual machine on G&Gs PCs. Petrel runs native on the PC. A key difference is also that at the time of writing, Petrel (at least the version used in NorthOil) does not include its own database, while openworks runs on an Oracle database. This has interesting implications for the accumulation and coordination, as we will discuss in the next section.

Exploration begins by creating a new exploration project in Openworks by defining the geographical borders of the area to be explored (defining the *blocks* and *quadrants* to be included). Ideally, and in accordance with the defined workflow, a new project should be defined, but typically G&G will extend a bordering project, because they then know that they get all the data that is in that project and reduce the danger of missing some key information. The G&G collaborates with the PDMs that help them get the data they need. For instance they would like to have or "rock core permeability and porosity filtered by location, stratigraphic unit and depth" (source; G&G presentation), or; "*in quad 35, find me all wells which have total depth of more than 3000 meters or a certain pressure of 200 bars or based on types of stratigraphic units or sand*" or "*find all wells in block X that took more than 30 days to drill, and had problems that resulted in blowouts*". For seismic data, a request can be "*I want all the surveys in quadrant X, block Y, shot from 1990 by company Z*" (source: G&G informant).

Also, previous interpretations are interesting, and it is relevant to be able to find all interpretations done by a person in one area in one time period because you know that person and trust what he did in that period in that area. This can typically be data that has not been made "STAT" (meaning the official, quality controlled version) meaning it is not in the corporate data store, but rather in Openworks projects as "unofficial" versions. Interestingly, also the corporate data store did not earlier register whose interpretation that was the official “STAT” one. Now this has been changed and an official interpretation will always be linked to whom it was that did the interpretation, so that G&G always can see who did it.

The G&G ask the PDMs for all the seismic in an area, and one G&G report that the results are presented as A4 pages printouts of all the surveys (with outlines) found in the area that may be relevant, and they go through it together to determine what is to be loaded into the project. The process takes 2-3 days. Ideally, G&G would want to have all data that is found and relevant gridded up in a map, in their interpretation tool.

The corporate data store, the Diskos DB and other Openworks projects are searched to find data. The amount of information is the same regardless if you are working on access projects (exploring new areas) or mature fields already in production. This is because in access fields you are exploring larger geographical areas, making rougher more coarse grained interpretations, and therefore include for example 300 wells, while in the North Sea you could include 300 wells exploring a much smaller geographic area.

Some data that is of interest to the G&G are old enough to not having been digitalized, and need to be access in physical copies. Other information is so new, like really fresh drill reports or interpretations, that they are not in the official data stores yet. This information then is found e.g. in team sites. All content in the team sites are indexed and searchable, but access control regimes limits what you can access and not. This is considered a “*huge problem*” (source: PDM informant).

When the G&G has done an interpretation, e.g. interpreting a pick from a well log to determine a formation, they need to name it. There are standardized names and lists for picks, so that a G&G should choose from a “pick list” of formations in order to categorize their pick. This is not always done however, and one enters another name using free text (“formation Rogaland” instead of “Rogaland formation” for example). Using the same standardized category allows correlation between wells. One wish to follow standards (particularly that others do it), but it is often not done. One potential reason is that a lot of lithographic details are finer grained than the official categories allow.

As stated above a key difference between Petrel and Openworks is that Openworks includes its own database. Projects are created and stored in Openworks while many (if not most) work in Petrel. PDMs then need to keep an exact mirror of the data available in Petrel, and make sure to copy back to the Openworks database the interpretations that the G&G feels should me made available for others, that is not official (STAT) as in the corporate data store but rather, something that is good enough to share. This is a key difference from working in Openworks itself. While working in Openworks, everything stored directly in the DB, with the effect that it was available for searches by the PDM afterwards so that it can be retrieved, if needed for some reason or another, in some future exploration. Now, for the G&G working in Petrel, they have to make explicit decisions to upload data to the Openworks DB. This is often not done. One reason is the time constraints, one always needs to run off to the next area to explore, and there is no time to tidy up (uploading) the data. Second, to do a deliberate upload make the interpretation seem more "official" than the G&G feels it is, and they are therefore more hesitant. An unfortunate side effect is that potentially relevant interpretations done in Petrel, but not uploaded, are lost.

4 Spatiotemporal Accumulations and Coordination in Interpretation Work

The above empirical account shows how collaboration is achieved in a case that is clearly not restricted to the local, but rather where coordinating activities and shared understanding must be performed in an "*unbounded*" II (Hanseth & Lyytinen 2010), using data that is generated across space and time. In the discussion, I will draw on CSCW studies from the healthcare domain and point out some similarities but also some notable differences in the domain of O&G exploration.

Berg argues that we need to seek empirical accounts of what it is "reading and writing artifacts" do within medical practice (Berg 1999). Berg suggests giving a "*minimalist, empirical depiction of what it is information technologies can be occasioned to do without falling back on the essentialist, non-relational accounts we want to avoid*?" (ibid.). In his analysis he finds coordination and accumulation to be the two central capacities that reading and writing artifacts can be occasioned to perform in work practices. The reading and writing artifacts that Berg studies are arguably less complex than the II perspective on inscription devices that I have described above (Berg´s order form and fluid balance sheet in one medical record application versus Openworks and Petrel that has many features, such as helping in construct interpretations, algorithms and tools for 3D modeling for instance). The key capacity is the same, however, and hence the unit of analysis here; the *accumulation* of data in the O&G exploration II and the *coordination* it entails.

Coordination

In a more recent work, Bansler et al explains how CSCW studies over the last decade has shown how the medical record is "*complex and variegated*" and that studies has focused on "*the coordinative practices of clinical staff with special emphasis on the role of the medical record in these practices*", and the medical record is best viewed as "*an ecology of artifacts*" and a "*heterogeneous assembly of specialized representational and coordinative artifacts*" (Bansler et al. 2013). These characteristics are similar to those of II. In their continued analysis, Bansler et al narrows their focus down to one artifact, the progress notes and explains how it is "*is constructed in an ongoing process of aggregation and arrangement of test results and observations, of offering hypotheses and suggestions, of deduction and* *allusion, of explicit reference and tacit omission*." and that the notes coordinate; "*They function as a cognitive artifact that facilitates memory and recall and they enable collaborative sense-making and coordination of actions in a highly complex, distributed work practice*." (ibid.). On a similar note, Berg explains how information technologies "*afford the increased distribution of work practices over a greater number of entities, and over more times and spaces*" (Berg 1999). Fitzpatrick and Ellingsen also shows how CSCW over the last 25 years has evolved from artifact mediated healthcare work, through locating healthcare in space and time, to expanding contexts of healthcare work with "*large scale implementations - integration and standardization challenges*" (Fitzpatrick & Ellingsen 2013). This coordination over time and space in the domain of healthcare is similar to what we see in the empirical account above. Data is entered into the different data stores from many places and many sources, of many kinds and about many places. At certain times, for different reasons, the area becomes interesting, and the II is searched for all data belonging to that area. The G&G want an overview of all the data, for then to be able to select what data is relevant to them. To aid their work, and select what to use and not, they need to know meta-data, such as when is the data from, who made it, etc. They buy and trade data if the company does not already have what they need, a case where the metadata becomes even more relevant. The data, and data about the data (metadata) entered yesterday, guides in many ways the work performed today. The hasty performance when creating the yesteryears final drill report for example, constrains and shapes the work performed today, e.g. by forcing the G&G of today to investigate more to fill in the gaps that exist in the report. They question the report and fill in the gaps, making new interpretations. Questioning the interpretations of others is something the G&G do, and something that is key for the II to support, as we shall see when we discuss *accumulation* below.

Accumulation

Accumulation in reading and writing artifacts Berg says; "*reorganizes individual inscriptions into aggregates – through its spatial layout, or through computational operations*" (Berg 1999). The notion of inscriptions here draws on Latour´s study of scientific laboratory work, where he finds that the production of scientific "*fact*" is a process of "*literary inscription*" in scientific papers. Latour explains; "*A text or statement can thus be read as "containing" or "being about a fact" when readers are sufficiently convinced that there is no debate about it and the processes of literary inscription are forgotten*" (Latour & Woolgar 1986). Latour points to how all the debate and all the "messy" process of generating the "fact" is at certain points forgotten, and it is accepted in the community as facts. Berg explain from the medical domain; "*Rather than having to check all the individual entries, nurses and doctors can work from the aggregated fluid balance, or wait* *for the monitor to beep when e.g. a patient´s blood pressure drops below a certain point*" (Berg 1999). The fact that you "can" does of course not mean that you do, or should. In the exploration case described above, they certainly do not accept the accumulations, or layered inscriptions, simply as facts, but rather indications that needs to be questioned. Remember they want to have an overview of all the relevant (given some criteria) data in an area. From that, they select the data they want to work with to create their interpretation. So, different from scientists that sometimes accept things as facts by referring to e.g. a theory in a paper, a paper where the link from the data to the theory is absent, the G&G need to be able to always get to the data that supports the current interpretation. Accumulation therefore is not a tool for simplification, but rather tools for questioning and reflection.

Bansler et al, attributing questioning practice also to scientists, draws similarities between the clinicians progress notes in the domain of health, to the practice of research: "*In a way that is similar to a scientific community's evolving repertoire of papers (apart from the imperative to act that is defining of clinical work), some entries serve to present bits of fact (similar to research notes), other entries serve to outline treatment plans or strategies (research problems and hypotheses), while other entries again serve to review what has been learned so far. Written over time by several clinicians, often from different specialisms, in a highly distributed process, the progress notes serve to reflect ongoing external developments, to select and counterpoise bits of data, to formulate hypotheses as to causation, to suggest lines of action, etc*" (Bansler et al. 2013). The argument that data is "counterpoised", at certain points in the patient trajectory (or biography), resonates well with how the G&G gradually build interpretations from the historical archive.

The information infrastructure should naturalize and historicize to support cooperation

Different kinds of interpretations (e.g. geological and geophysical picks, horizons) are generated at certain times at certain places, but becomes part of an installed base in the information infrastructure, resides there, prepared to enable generation of gradually new interpretations, in concurrent cycles. Gradually new and improved inscriptions (interpretations) are made on the inscription devices, inscription upon inscription. But when is the inscription finished? When is it a fact? When is a "stabilized representation" ready? The answer seems to be never. It is an ongoing naturalization process. Bowker and Leigh-Star discuss how objects naturalize, objects being defines as "*stuff and things, tools, artifacts and techniques, and ideas, stories and memories – objects that are treated as consequential by community members*" and thereby also relevant to data such as in our case, explains: "*Naturalization means stripping away the contingencies of an object´s creation and its situated nature. A naturalized object has lost its anthropo-*

logical strangeness. It is in that narrow sense desituated – members have forgotten the local nature of the object´s meaning or the actions that go into maintaining and recreating its meaning." (Bowker & Star 1999). But it is not fixed. Objects become naturalized in certain communities at certain times, along "*trajectories of naturalization*" and it is not known a-priori whether an object will become naturalised, or how long it will be so (ibid.). An interpretation in the II will move in and out of naturalization trajectories throughout time and space, and an interpretation is constructed upon combinations of earlier and new interpretations, being stable and natural at certain points of time, (e.g. a prospect or a final well report), before being moved away from the focus, stored, becoming a part of the accumulated archive in the II. G&Gs insist on having flexibility in choosing their own naming schemes (categories in Bowker´s terms), creating "unfinished" drill reports, and how concerned they are with the meta-data (who created the data, what equipment, when was it created). The II is consequently need to exercise flexibility – e.g. as when the PDMs accept jobs, how they clean the data, the data that need to migrate between Petrel and Openworks, being archived in different datastores and accessed in a variety of ways and forms. For Bowker studying biodiversity infrastructures (Bowker 2000), this is a key insight; the II must reflect the diversity of the work at hand, but equally important, keep a historical record of it. When working with the complex issues of describing nature (be that the case of a human and their disease as in the cases of Berg and Bansler et al, rock formations as in my case, or biodiversity as in Bowker´s case), practitioners face two emerging issues. First, that the way we store information and categorize it is *performative*, in that it shapes how we view the world. If all we store is "Rogaland formation" although it is considered too broad a category, we, over time, construct a reality based on the "Rogaland formation". Second, as a consequence of the first, we will have irreversibility. If all we store is "Rogaland formation", we cannot go back (in a DBs consisting of 110Tb and counting), to recreate the lost categories. We have lost the reasons why "formation Rogaland" was relevant.

It is interesting to note in the empirical account how the two inscription devices of openworks and Petrel affect the building of such historical accounts. In Openworks, with the DB structure, interpretations were saved behind the scenes, and fixed by the PDMs. In Petrel, the process is made more explicit, and a barrier is made towards saving. This creates a challenge. Bowker (Bowker 2000) would argue towards the Openworks approach as a strategy out of the "*irreversibility*" bind. You may not save everything, but you should aim for "*deep historicization of our datasets*" (*ibid*.). The goal is to categorize and formalize historical perceptions of data, so as to enable the practitioner in one location of today to understand the data generated by practitioners elsewhere, from the past. The same principle is found when discussing the importance of context and how the design of the system influence and constrain to what degree the context can be captured and shared; *"…any computer system that affords representation and awareness of human activity necessarily involves a degree of reduction and objectification, due to the formal representational schemes of programs and databases, and finite capac-*

ities for storage, communication and calculation." (Chalmers 2004). The II constrains and enables to what degree historificazion is feasible. The awareness of the history needs to be coupled to the work and tools practitioners use, not separated from them. To achieve this, we need "*to make coordination and collaborative functionalities an aspect of the collaborative artifact rather than of a collaborative application*" (Cabitza & Simone 2013). The empirical case here suggests that an II view would move the focus away from a single artifact of cooperation and seek more holistic explanations with potentially more power. How both the PDM plays a pivotal role in storing and accessing the correct data as well and the possibility of the datastores to support queries of contextualized historical data (who created the data? who was the interpreter? Etc.). This is the sociomaterial working of an II. Schmidt, in explaining the fragmentation of CSCW, shows how the field have moved from computer mediated communication, through office automation, to CSCW where the key problem *"…is not "communication*" or "*resource sharing*", but "*the cooperating actors´ control of their interaction and, by implication, of the computational regulation of their interaction*" (Schmidt 2009). These issues of coordination are not restricted to the local either "*Indeed, coordinating interdependent activities across space is one of the problems faced by actors engaged in cooperative work "in the wild*."" (Ibid. p 237), and perhaps we could add, across time, implying asynchronous coordination, as noted by Cabitza and Simone: "*Too often the fact that actors actively monitor and proactively display awareness information is disregarded in favor of undifferentiated mechanisms of notification. In doing so, the fact that the proactive part of the phenomenon remains unsupported, especially in asynchronous remote collaboration, is weighted against the fact that the resulting technology might seem easier for the user to appropriate and surely simpler for the designer to construct*." (Cabitza & Simone 2013)

5 Concluding Remarks and Future Work

In the discussion above I have discussed accounts from the CSCW healthcare literature with observations in oil and gas exploration, using an II approach. While both Berg and Bansler et al consider specific artifacts to be coordinative, we see from the description above the sociomaterial components (social; G&G, PDMs; material; Openworks and Petrel) that go into an II to enable it to store, process, and retrieve historical data to support the reflection of today. I have started unearthing an II in exploration with the focus on data in the form of interpretations, and *what* is standardized *when* in the biography of the interpretations it is standardized and for *whom* it applies (E. Monteiro et al. 2012). Williams and Pollock, although focusing on systems rather than data, also argue for the notion of studying IIs from a biography point of view based on a growing dissatisfaction with the "*single site implementation study*" (Williams & Pollock 2012). Systems and data need, in an II perspective, to be placed in broader perspective, to understand the "locales and actors" that play a role in shaping the performance of the II, to; "*develop better temporal understanding of ERP implementations that include not simply the immediate response by actors but also the multiple and often longer-term temporal conceptions that might surround deployment and appropriation*" (*ibid.*).

In this work I have taken the first steps of unearthing the II of oil and gas exploration in terms of coordination, accumulation, historificazion and naturalization. I have empirically shown the relevance of "*deep historicization"* of datasets in an II, giving practitioners the ability to properly reflect and make proper interpretations. This is inline with recent work from Haavik (Haavik 2014), who frames this sociomaterial process as "*sensework"* (as separated from sensemaking) being characterized as sociomaterial (different distributed teams in hi-tech environment), cognition and meaning is indistinguishable from work itself, there is no final right solution, nothing is final, only worked on, and sense connotes sensors (no direct empirical confirmation is possible). Future research here should provide insight into the data and models of an II that shape and are being shaped by spatiotemporal independent reflective practice. Finally, it is relevant in this setting, given the role of nature, to extend the notion of materiality in the socio-materiality debate from the interplay between the social organization and their information systems, towards also including additional levels, such as the physical materiality of nature and how it is dealt with (Østerlie et al. 2012). In so doing, interesting challenges emerge, both for the CSCW and IS field, in how we methodologically and analytically address the bi-directional impact between the material nature, the materiality of the II representing nature, and the social organization.

References

Bansler, J. et al., 2013. Physicians' Progress Notes. In O. W. Bertelsen et al., eds. *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*. Springer London, pp. 123–142. Available at: http://link.springer.com/chapter/10.1007/978-1-4471-5346-7_7 [Accessed November 7, 2013].

Berg, M., 1999. Accumulating and Coordinating: Occasions for Information Technologies in Medical Work. *Computer Supported Cooperative Work (CSCW)*, 8(4), pp.373–401.

Bowker, G.C., 2000. Biodiversity Datadiversity. *Social Studies of Science*, 30(5), pp.643–683.

Bowker, G.C. & Star, S.L., 1999. *Sorting Things Out: Classification and Its Consequences*, MIT Press.

Cabitza, F. & Simone, C., 2013. Computational Coordination Mechanisms: A tale of a struggle for flexibility. *Computer Supported Cooperative Work (CSCW)*, 22(4-6), pp.475–529.

Chalmers, M., 2004. A Historical View of Context. *Comput. Supported Coop. Work*, 13(3-4), pp.223–247.

Crabtree, A., Tolmie, P. & Rouncefield, M., 2013. "How Many Bloody Examples Do You Want?" Fieldwork and Generalisation. In O. W. Bertelsen et al., eds. *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*. London: Springer London, pp. 1–20. Available at:

http://dblp.l3s.de/d2r/page/publications/conf/ecscw/CrabtreeTR13 [Accessed November 7, 2013].

Fitzpatrick, G. & Ellingsen, G., 2013. A Review of 25 Years of CSCW Research in Healthcare: Contributions, Challenges and Future Agendas. *Computer Supported Cooperative Work (CSCW)*, 22(4-6), pp.609–665.

Haavik, T.K., 2014. Sensework. *Computer Supported Cooperative Work (CSCW)*, pp.1–30. 10.1007/s10606-014-9199-9. Springer Netherlands.

Hanseth, O. & Lyytinen, K., 2010. Design theory for dynamic complexity in information infrastructures: the case of building internet. *Journal of Information Technology*, 25(1), pp.1–19.

Hyne, N.J., 2001. *Nontechnical Guide to Petroleum Geology, Exploration, Drilling, and Production*, PennWell Books.

Langley, A., 1999. Strategies for Theorizing from Process Data. *The Academy of Management Review*, 24(4), p.691.

Latour, B. & Woolgar, S., 1986. *Laboratory Life: The Construction of Scientific Facts*, Princeton University Press.

Mayernik, M.S., Wallis, J.C. & Borgman, C.L., 2013. Unearthing the Infrastructure: Humans and Sensors in Field-Based Scientific Research. *Computer Supported Cooperative Work (CSCW)*, 22(1), pp.65–101.

Monteiro, E. et al., 2012. From Artefacts to Infrastructures. *Computer supported cooperative work (CSCW)*.

Østerlie, T., Almklov, P.G. & Hepsø, V., 2012. Dual materiality and knowing in petroleum production. *Information and Organization*, 22(2), pp.85–105.

Parmiggiani, E. & Mikalsen, M., 2013. The Facets of Sociomateriality: A Systematic Mapping of Emerging Concepts and Definitions. In M. Aanestad & T. Bratteteig, eds. *Nordic Contributions in IS Research*. Lecture Notes in Business Information Processing. Springer Berlin Heidelberg, pp. 87–103. Available at: http://link.springer.com/chapter/10.1007/978-3-642-39832-2_6 [Accessed November 7, 2013].

Prilla, M., Pammer, V. & Krogstie, B., 2013. Fostering Collaborative Redesign of Work Practice: Challenges for Tools Supporting Reflection at Work. In O. W. Bertelsen et al., eds. *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*. Springer London, pp. 249–268. Available at: http://link.springer.com/chapter/10.1007/978-1-4471-5346-7_13 [Accessed November 7, 2013].

Schmidt, K., 2009. Divided by a common acronym: On the fragmentation of CSCW. In I. Wagner et al., eds. *ECSCW 2009*. Springer London, pp. 223–242. Available at: http://link.springer.com/chapter/10.1007/978-1-84882-854-4_14 [Accessed November 7, 2013].

Schmidt, K., 2011. The Concept of `Work' in CSCW. *Comput. Supported Coop. Work*, 20(4-5), pp.341–401.

Star, S.L. & Ruhleder, K., 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, 7, pp.111–134.

Williams, R. & Pollock, N., 2012. Research Commentary—Moving Beyond the Single Site Implementation Study: How (and Why) We Should Study the Biography of Packaged Enterprise Solutions. *Info. Sys. Research*, 23(1), pp.1–22.