Digital oil (Doil)

1. Relevance relative to the call for proposals

The lifecycle or *biography*[1] of an oil and gas well lasts several decades with enormous volume and variety of type of data used by numerous groups. Crucial data for the operation of the well comes from a vast and increasing network of *sensors, measurement- and logging devices*. A typical oil and gas well on the Norwegian continental shelf would contain about ten networked sensors and measurement devices. For the unmanned subsea wells, of which there exists more than 500 on the Norwegian continental shelf today, the data provided by sensors and measurement devices is the *only* source of data – subsea wells are run literally untouched by humans, solely based on sensors – thus motivating the academically challenging and practically relevant question of how to operate subsea wells efficiently and safely.

The project proposal Digital oil (Doil) is built around two themes that correspond to two of Verdikt's three themes. First, Verdikt's theme on the *Internet of Things (IoT)* relates to the proliferation in number and type of networked sensors in different domains. In oil and gas operations the reliance on networked sensors are absolutely necessary; for the large and rapidly growing area of unmanned subsea wells there literally is no other option. Doil asks the question: how are oil and gas operations run efficiently and safely based on the collective, ongoing assessment by engineers interpreting and managing the networked sensors?

Second, key to Verdikt's theme on *Social networks* is co-construction of content that erodes the distinction between content providers and users. In the Doil, the subsurface community's searching for relevant data about the well suffers from too strict separation between content providers and user i.e. has yet to exploit the potential of co-constructed content.

2. Aspects relating to the research project

The key question that Doil addresses: how do you obtain a sufficient *overview* of the subsea well to operate them in an efficient and safe manner. This hinges crucially on how sensor-based data is interpreted and relevant data from the biography of the well is identified. Doil is organized in five work packages, each of which contributes to the main objective of the project through their respective work package objectives:

- WP1. Identify existing *tactics* by production engineers to assess the credibility of sensor-based data in daily operations;
- WP2. Identify existing *user-driven heuristics* for generating overviews of the well lifecycle;
- WP3. Develop a *demonstrator* with associated work process for situated search of the full lifecycle of well data;
- WP4. Cultivate a *network* of subsurface community members, vendors and researchers;
- WP5. Disseminate results through NTNU Master of Management *executive education* and international publications in conferences, journals and one edited book.

[1] Recognising the long time-spans of technologies by using the analogy of a "biography" is borrowed from N. Pollock and R. Williams (2009), *Software and organisations. The biography of the enterprise-wide system or how SAP conquered the world*, Routledge.

2.1. Background and status of knowledge

Oil and gas production in the Norwegian Continental Shelf (NCS) is becoming increasingly dependent on subsea[2] installations (see figure 1). Currently, over than 50% of oil and gas production in the NCS comes from subsea wells. Globally, more than 400 subsea wells are drilled each year and the global subsea market may reach USD 50 billion in 2014, from USD 21 billion in 2009[3]. Such growth is conditioned due to ability to extract oil and gas from ultra-deep waters and under harsh environmental conditions such as in the Barents Sea and the Arctic Ocean. Successful operation of subsea installations is considered to be vital for global economy[4].

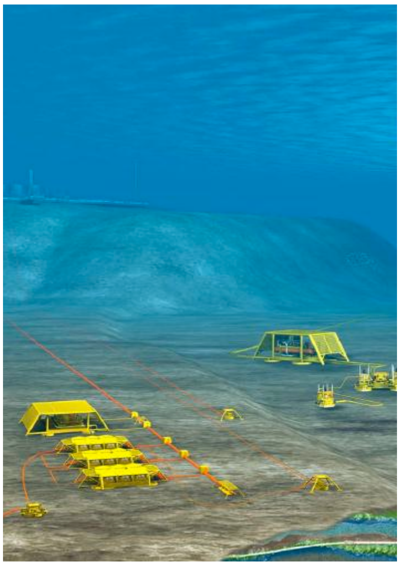

Figure 1: Subsea installations that are equipped with multiple sensors

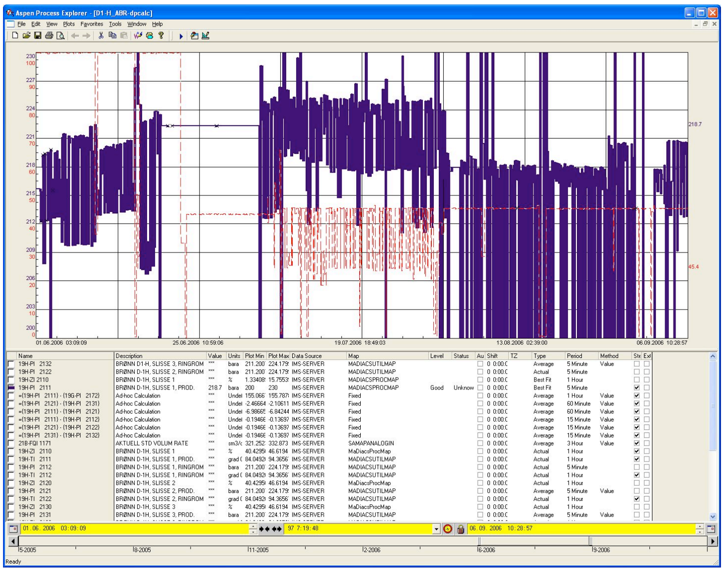

Figure 2: An example of sensor-based data representation that illustrates pressure changes over time

The daily operation of an oil and gas well rests on a truly massive amount[5] of data generated over the full lifecycle[6] of the well, generated by number of different disciplines within the subsurface community[7] engaged with distinct purposes and real-time and historical data from extensive networks of sensors. Sensors are mounted at fixed locations within a well to measure properties such as pressure, fluid flow rate, temperature, vibration, composition, fluid flow regime, and fluid holdup. Sensors are crucial to the subsea specialist to develop an understanding of what (hydrocarbons, water, sand) and at with what rate a well is producing as well as whether production is under safe conditions (see figure 2 for an example of sensor-based data representation).

Sensors, however, are of varying accuracy, they are sometimes faulty (but costly to repair) and differ significantly i.e. you need to know "your" sensors to make sense of their

---

2 In contrast to topside wells that are drilled and maintained from platforms, subsea wells are completed on the sea floor and considered to be "invisible", as they are remotely assembled, operated and maintained.

3 http://www.regjeringen.no/nb/dep/oed/aktuelt/taler_artikler/politisk_ledelse/taler-og-artikler-av-statssekretar-per-r/2010/Subsea-Technology--Norwegian-and-International-Perspectives.html?id=620417

4 http://events.sut.org.uk/past_events/2006/060518_london/SUT18May06SubseaDomain.pdf

5 To illustrate, during drilling phase alone more than one thousand documents can be produced for a single well.

6 The lifecycle of a well includes exploration, well planning, drilling, production, well interventions, altering of the production regimes and of well status (e.g. production to injection), and termination.

7 Subsurface community consists of professionals from a variety of disciplines, e.g. geophysicists, geologists, reservoir engineers, well engineers, production engineers and process engineers.

readings[8]. The lifetime of a sensor is limited, especially for high pressure/ temperature wells. An instant increase of sand production may damage sensors within hours or days. As a result, the central concern for subsea engineers is to assess the reliability of sensor-based data including comparing measurements across sensors in different zones or different timescales and relating sensor-based data to other data sources created my multiple subsurface communities during the life cycle of the well[9].

Clearly it is impossible to grasp all data about a well; simplifications are absolutely necessary. The difficult questions are *which* simplifications are to be trusted, where and *how* to access fuller versions of the data and *who* to consult when making sense of it. We map out what we take to be the principal research issues related to Doil.

Theme A: Trusting sensor-based information

Oil and gas companies, nuclear and chemical plants as well as many other organizations in high-risk industries increasingly rely on the effective management of sensor-based data. Employing advanced ICT's sensor-based data is transformed into multiple representations to be displayed in control rooms or made available to specialists for additional/custom analysis. A number of detailed studies illustrated that managing such data pose significant challenges and occasionally lead to disastrous accidents. Specifically, scholars have clearly established that management of such data can only by partially explained by normative decision-making models where actions and interactions between the involved actors are predefined in advanced[10]. For instance, a study of The London Underground control rooms[11] vividly illustrates that sensor-based data is only one and partial source of data that has to be complemented with other data sources, discussed with colleagues and related to previous experiences.

Production engineers in oil and gas companies are dependent on sensor-based data since they cannot physically investigate a well. In order to understand well performance, production engineers investigate real-time sensor-based data as well as historical data that is produced during the lifecycle of a well. Data related to a particular well is continually accumulating, resulting in truly massive amount of data. Multiple subsurface disciplines work with specific aspects of a well and during distinct phases of well's lifecycle. Each discipline has specialised IS and produce data with *specific purposes* in mind but may later on be used by other disciplines. In short, data about a single well is highly heterogeneous and fragmented[12]. The reuse of data outside its initial and intended context of use is often problematic: there are tacit assumptions about how to make sense of the data (is a blank a zero or missing data? what kind of equipment was used for this measure?) that quickly get left out in subsequent reuse[13]. It is neither practical nor possible to make explicit all tacit aspects of the context of the data.

For data to be useful in action-taking with significant consequences, the data needs to be rendered *credible* and trustworthy. Data is not credible in itself as there regularly are

---

[8] A number of scholars have pointed out the difficulties of reusing data e.g. Berg, M. and E. Goorman (1999). The contextual nature of medical information, *Int. Journal of Medical Informatics* (56:1-3), pp. 51-60.

[9] To a single well there will be a whole range of different types of drilling and completion data, production data from the platform and down hole gauges, log data, resistivity measurements, documentation of the completion and cement, well tests, documentation of production problems and well interventions etc. Most wells during its production phase will be subject to studies and evaluations, that to a varying degree will be documented.

[10] Weick, K. E. (2001). *Making Sense of the Organization*, Blackwell.

[11] Heath, C. and P. Luff (1992). Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms, *Computer Supported Cooperative Work* (1:1), pp. 69-94.

[12] Hepsø, V., E. Monteiro and K. H. Rolland (2009). Ecologies of e-Infrastructures, *Journal of the Association for Information Systems* (10:5), pp. 430-446.

[13] (Goorman and Berg (1999) (ibid.)

errors, inaccuracies or essential tacit assumptions not satisfied. The question is *what do users do* to transform mere data into credible, action-relevant data? A considerable amount of validation, double-checking and sense-making of data is necessary. These skills involve the *communication of uncertainty* i.e assumptions about the data. The working knowledge is a combination of the various sources and knowledge representations that in particular involve assessing others' point of view[14]. In short, knowledge production, in complex organisations, is not a product of individual effort, but rather is highly collective endeavour. Tasks that even appear individual are interactionally organised. In oil and gas companies production engineers are responsible for safe and efficient oil and gas production optimisations, yet they rarely make decisions on their own. In collaboration with other disciplines they spend significant amount of time in pre-planned and ad-hoc meetings where they make sense of historical and real-time (sensor-based) data and collectively decide future plans. In their own words, production engineers spend more time in meetings than in front of their PC screens. As a result, decisions are rarely predefined in advanced, they are situation dependent and collectively made on tacit assumptions.

The significant effort required to make data credible and trustworthy has implications for automation in high-risk industries. Regarding oil and gas organisations, the question of how as well as to what extent oil and gas production should be automated (i.e. predefined decision making based on sensor-based data) becomes central.

The notion of high reliability organisations (HRO)[15] has been coined to denote socio-technically complex organisations, e.g. nuclear reactors and industrial plants, with the potential for catastrophic events that appear remarkably good at avoiding disasters. The question, then, is what HRO do and what others can do to learn from them? A consistent theme in discussions around HRO is the pivotal role of *redundancies*[16]: no robust, fault-tolerant distributed and collective operation involving technology exists without forms of redundancies. It is not a question whether redundancies are productive but where and what forms to establish[17], a result underscored also in studies of resilient, knowledge-based organisations[18].

The study of how production engineers assess the credibility of sensor-based data in daily operations (WP1) will develop an in-depth understanding of theoretical and actual risks related to sensor-based data management and will contribute to discussions on automation in oil and gas industry as well as HRO.

Theme B: Situated search for information

Doil is based on the assumption that information systems should be viewed as *collections* of systems rather than stand-alone applications. These collections are constantly evolving, mutating, integrating, with episodical disruptions far beyond the image of systematically organised portfolios[19] of information systems. They consist of numerous, historically layered information systems where new components partly extend, partly substitute and partly superimpose existing ones[20]. The functionality of subsurface IT tools is not the sum of the

[14] Cf. Boland, R.J and Tenkasi R.V. (1995) Perspective Making and Perspective Taking in Communities of Knowing. *Organization Science,* Vol. 6, No. 4 (Jul. - Aug.), pp. 350-372

[15] La Porte, Todd R. and Paula Consolini (1991) "Working in practice but not in theory: Theoretical challenges of High-Reliability Organizations". Journal of Public Administration Research and Theory Vol. 1 pp.19–47

[16] See Perrow, C. (1999). *Normal accidents: Living with high-risk technologies,* Princeton, N.J.,University Press.

[17] Hutchins, E. (1995). *Cognition in the wild, MIT Press*.

[18] Hollnagel, E. and Woods, D. (2006) *Resilience engineering: concepts and precepts*, Ashgate Publishing

[19] M Jeffery and I Leliveld. (2004)Best practices in IT portifolio management,*MIT Sloan mgmt. review*,pp.41-49.

[20] This has led scholars to conceptualise such collections of information systems as "ecologies" or information "infrastructures", see e.g. C. Ciborra et al. (eds.) (2000), *From control to drift. The dynamics of corporate infrastructures*, Oxford Univ. Press.

functionality of its constituting components: users' needs and demands may be served by multiple functions and one function may serve multiple needs. The implication is that users adopt and "workaround" intended use of functionality and constraints[21].

In addition to data fragmentation and heterogeneity, data management policies and practices are changing over the time. Such changes are invoked due to increasingly stricter external laws and regulations with witch oil and gas companies have to comply as well as possibilities afforded by new oil and gas and ICT technologies. As a result, what documentation is available about each well varies, notably older wells are significantly less documented. More specifically, it implies that the same type of data is documented in varying detailed, can have different labels is stored in multiple formats in different IS or in case of old wells in paper based format. Given the vast and growing amounts of data attempts to produce a complete mapping and develop search engines based on unified terminologies were largely unsuccessful in such contexts.

In order to find the data that is fragmented across multiple and heterogeneous sources, *users develop extensive repertoires of socio-technical search tactics relying on heuristics, social networks and tool-based search capabilities*[22]. Such perspective has consequences for standardization and ontology development. The reliance on top-down structured and planned meta-data structures works quite well in small contexts and during short periods of time. Such approaches, however, are only partially suitable in large-organisations that aim to increase collaboration and knowledge sharing across multiple disciplines and sustain data retention over long periods of time.[23]

An alternative to top-down approaches is illustrated by folksonomies. Folksonomies are user-generated, therefore inexpensive to implement and can potentially develop into an emergent business taxonomy. Folksonomies build on institutionally supported taxonomies or controlled vocabularies. They also acknowledge the distribution of data classification to those professionals that are actually doing the work. Recent explosion of tagging, RSS and weblogs on the Web illustrate the possibilities of decentralised data production and sharing[24]. Such approaches are particularly relevant to large-scale organisations as they acknowledge different user communities with distinct data management needs. Recent research shows that such approaches can be successfully employed in organisations, yet it requires significant effort to empower users and ensure that they are continually engaged in the process of classification development[25].

Over the years, oil and gas companies, have been primarily focusing on implementing controlled vocabularies, yet limited success generate increasing interest in the application of folksonomies[26]. How can folksonomies be implemented and managed in large-scale organisations is to a large extent unresearched topic. To this extent, it is important to mention that folksonomies have weaknesses such as ambiguity of tags, lack of structure, limited lexical control as well as cross-context challenges[27]. Our working hypothesis is that

---

[21] Gasser, L. (1986). The Integration of Computing and Routine Work, *ACM Transactions on Office Information Systems* (4:3), pp. 205-225.

[22] Jarulaitis, G. and E. Monteiro (2010). Unity in Multiplicity: Towards Working Enterprise Systems. *Proceedings of the 18th European Conference on Information Systems Pretoria,* South Africa.

[23] Karasti, H., K. S. Baker and F. Millerand (2010). Infrastructure Time: Long-Term Matters in Collaborative Development, *Computer Supported Cooperative Work* (19, pp. 377-415.

[24] Boast, R., M. Bravo and R. Srinivasan (2007). Return to Babel: Emergent Diversity, Digital Resources, and Local Knowledge, *The Information Society* (23:5), pp. 395-403.

[25] Ribes, D. and G. C. Bowker (2009). Between meaning and machine: Learning to represent the knowledge of communities, *Information and Organization* (19:4), pp. 199-217.

[26] Jarulaitis, G. (2010). The Uneven Diffusion of Collaborative Technology in a Large Organisation. *IFIP WG 8.2 + 8.6 Joint International Working Conference.* Perth, Australia.

[27] Ellingsen, E. , Monteiro, E. and Munkvold, G. (2007) Standardised work: co-constructive practice, *The Information Society*, vol. 23, no. 5, 2007, pp. 309-326

folksonomies should be combined with controlled vocabularies that would result in a hybrid classification of knowledge referred as collabulary[28].

Our primary aim (WP2, WP3) would be to provide incremental improvements to existing search technologies, enriching both indexing and searching with knowledge rich structures.

[28] Collabulary is a word combining “collaborative” and “vocabulary”. It refers to a theoretical method of labelling and organising data by collaborative tagging. It combines strengths of controlled vocabulary, ontology and folksonomy, see here: http://maisonbisson.com/blog/post/11196/collabulary/