Practical Knowledge Representation for Data Access to Subsurface Data: The Achievements and Potential of Optique Platform*

David B. Cameron[1]

Search and Discovery Article #42181 (2018)**
Posted January 29, 2018

*Adapted from oral presentation given at AAPG/SEG 2017 International Conference and Exhibition, London, England, October 15-18, 2017
**Datapages © 2018 Serial rights given by author. For all other rights contact author directly.

[1]SIRIUS Centre, University of Oslo, Oslo, Norway (davidbc@ifi.uio.no)

Abstract

The Optique platform is a suite of research-based software that allows effective, scalable access to data spread across multiple, often legacy, databases. It uses ontology-based data access to allow professional users, in this case exploration geologists, to pose ad-hoc queries using their own professional terminology. The platform was developed in collaboration with the exploration business unit in Statoil as part of a recently completed European Union research and innovation project.

This paper presents the Optique platform and give an end-user based perspective on how it works, using the example of the pilot installation developed with Statoil. We believe that this platform offers a novel and effective way of opening and integrating data from corporate data stores, commercial data sources, and national data repositories. To this end, we will describe remaining implementation and research challenges for OBDA and sketch possible implementation and commercialism scenarios. We also believe that best results will be obtained when OBDA is integrated in an interdisciplinary framework that also involves language processing, data science, and high-performance database systems. This is the aim of the new SIRIUS Centre.

Practical Knowledge Representation for Data Access to Subsurface Data

The Achievements and Potential of the Optique Platform

Arild Waaler, Martin Giese, Martin G Skjæveland, David B Cameron
SIRIUS Centre, University of Oslo

AAPG International Conference, 15-18th October, London

The SIRIUS Centre

- Eight years' financing from Research Council of Norway
- 11 Industrial Partners
- 3 Leading Academic Institutions
- Centre for Research-Based Innovation
- Funding for 20 Ph.D. students
- Innovation through prototypes and pilots
- 45 affiliated researchers

UiO : University of Oslo

simula

Statoil

Schlumberger

IBM

EVRY

DNV·GL

Dolphin

NUMASCALE
BIGGER DATA ANALYTICS

KADME
Knowledge and data management expertise

OSIsoft.

fluidOps

computas

NTNU
Norwegian University of Science and Technology

UNIVERSITY OF OXFORD

DEPARTMENT OF COMPUTER SCIENCE

sfi Centre for Research-based Innovation
The Research Council of Norway

Scalable Data Access

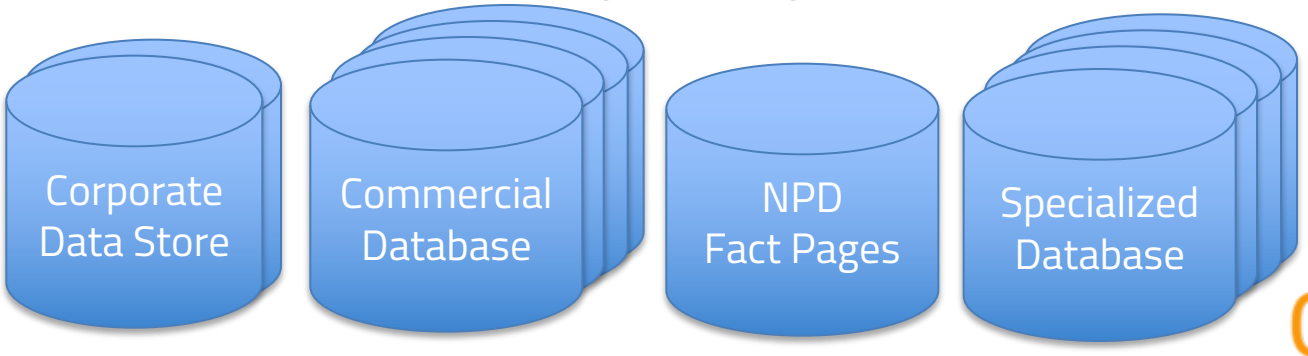
- Different formats
- Old software
- Complex, inconsistent data models
- Inefficient access methods
- Access and security
- Unstructured data
- Missing data
- Poor-quality data
- Too much data
- Manual work processes

Accessing data is a technical and organizational bottleneck.

We make poorer decisions and waste time on tedious work getting data.

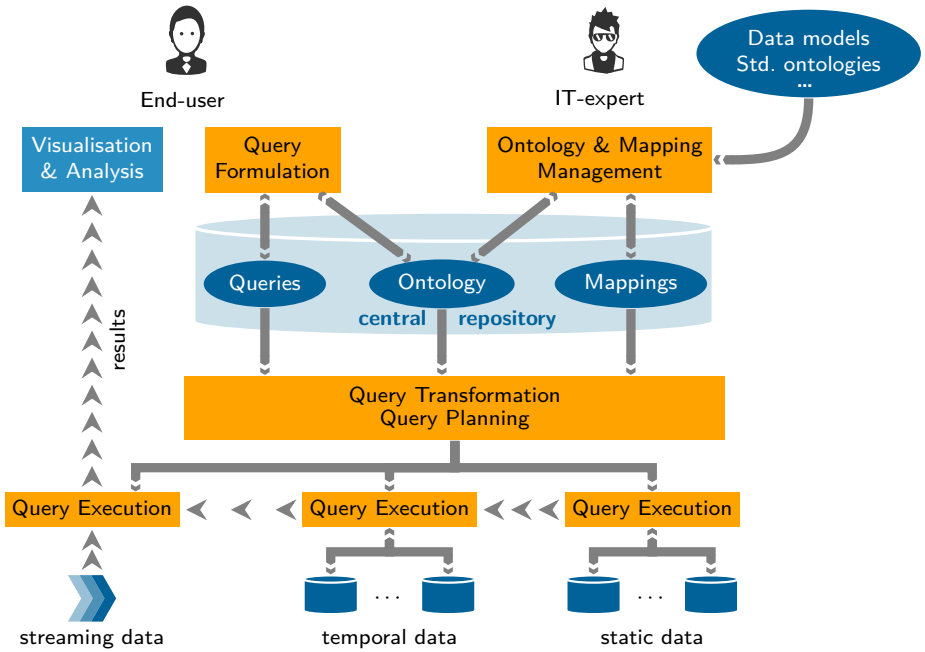
Scalable Data Access in Exploration

- 30-70% of exploration engineers' time spent on data gathering
- Dependent on IT experts to deliver reports and data warehouses
- The professional wants to:
  - Use own terminology in technical software
  - Make quick and easy ad hoc requests. Not a project!
- **Speed-up from weeks to minutes in data access.**
- **Project met all EU agreed targets after 3 of 4 years.**

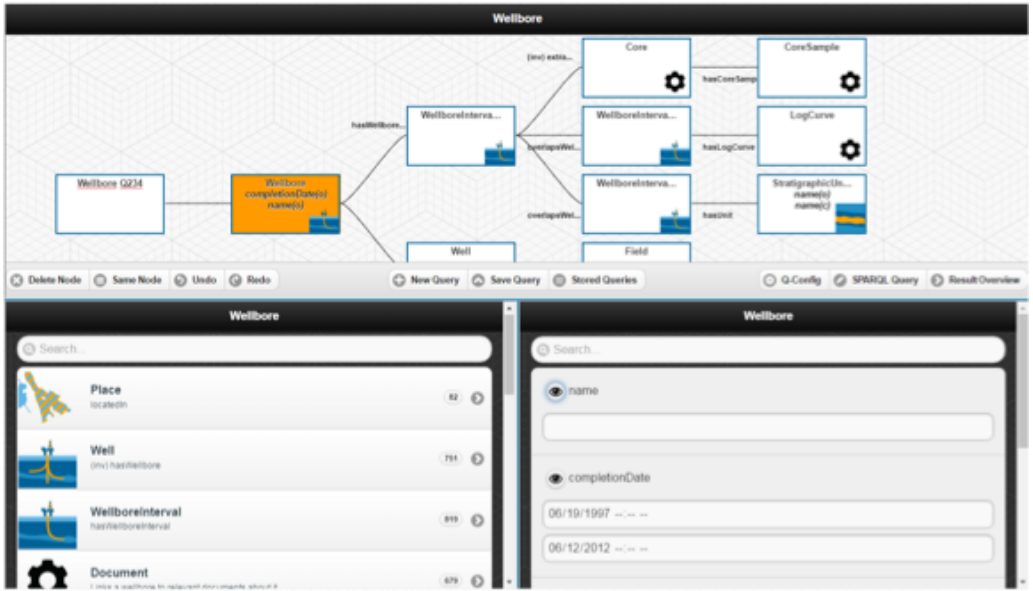


Optique

Ontology-based Data Access: How the Optique platform works



Accessing the data using a visual query system and geoscience concepts



Public example: SLEGGE database schema

Dag Hovland, Martin Skjæveland, Arild Waaler, University of Oslo
R. Kontchakov, M. Zakharyaschev, Birkbeck College, University of London

- Open package containing
  - Subsurface Exploration OWL Ontology
  - SLEGGE schema
  - Example information needs and queries
  - Mappings to the SLEGGE database
  - http://purl.org/slegge
- SLEGGE
  - Based on Epicentre / POSC data model
  - Statoil operational database for subsurface data
  - Oracle with 700 GB of data
  - 1700 views
  - 1700 tables

Can we use OBDA to activate the data in National Repositories?

Dag Hovland, Michael Heeremans, Arild Waaler

- National data repositories – DISKOS and CDA – are a mine of useful data
- Old models are based on archiving
- How will the support digitalisation?
- Will the 2015 model still be in place in 2025?
- We are working to demonstrate potential for opening up repositories to support federated and ontology-based queries.
- We are starting by taking our own medicine: the databases and DISKOS access at the University of Oslo
- Please lobby your NDR!

OTTR: Building maintainable ontologies: ontology templates

http://www.ottr.xyz

Martin G. Skæveland

- OWL can be very verbose and repetitive
- Easy to make errors
- “Needs a Ph.D. to drive it”
- Templates: re-usable pieces that can be built into an ontology
  - Common constructions
  - Good practice embodied in patterns
  - Reusable items
- Proven with an industrial project in engineering data management in Norway.

Conclusions

- Digitalisation and data science requires order in data
- Must be able to find and use data
- Pragmatically applied semantic technologies are necessary