

Geoscience workflow: Tracking multiple scenarios and uncertainties

Managing asset situations throughout the E&P lifecycle, using industry-standard data formats, enables future users to revert back to a critical step, to explore alternate scenarios or revisit critical uncertainties to reduced development efforts.

■ **MARCO PIANTANIDA, CLARA ANDREOLETTI, MATILDE DALLA ROSA, BRUNO VOLPI, Eni E&P and PHILIP NERI, Energistics**

Asset teams develop subsurface models and build reservoir simulation grids to describe the rocks and fluids of a hydrocarbon system. The amount of data, both from direct measurements in boreholes and derived from indirect detection technologies, such as geophysical or logging methods, is growing exponentially and is frequently updated over the life cycle of the asset.

The need for innovative technologies to process, interpret and analyze these data results in the use of a number of different application platforms and add-on modules to work through the complete workflow. RESQML™ is a data exchange standard developed by the Energistics industry consortium to facilitate the transfer of a complete project dataset from one subsurface application to another in a rapid and reliable manner. RESQML supports the creation of a set of files that are fully application-agnostic, which can be used as a comprehensive archive of a project at any time, in addition to their data transfer role.

The authors will explore the efficient use of this standard, both in its original role facilitating data exchanges and also an innovative use of the standard, as a means of managing scenarios and uncertainties

over the lifespan of a hydrocarbon asset.

RESQML FORMAT

The RESQML standard defines the data schema and mandatory metadata that ensure that the data being transferred are properly ingested by the receiving application in terms of their completeness, the relationships between all data objects, and their correct spatial and time referencing. It also establishes the files that will be generated by the exporting application. These consist of a compressed Zip file containing all the data and metadata (i.e. the “data about the data”) in XML format, and a bulk data file based on the HDF5 standard that contains all the large data objects, such as vectors and arrays. The zip file uses an industry-specific version of the open packaging conventions (OPC) called Energistics packaging conventions (EPC).

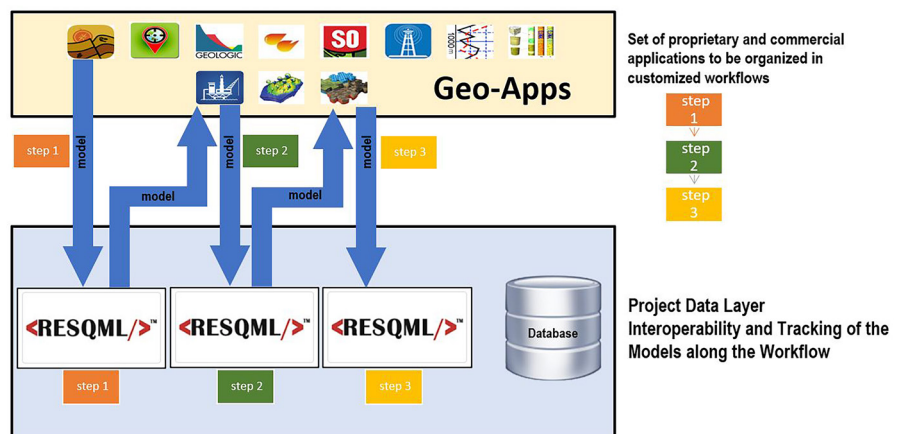
To make an application compliant with the standard, software developers need to develop an export module that will create the EPC and the HDF5 file, and an import module that can read the standard files and populate a new project or update an existing one. The standard stipulates the use of Universal Unique Identifiers (UUIDs) that are assigned to each data object. This ensures that if an

instance of a data object, e.g. a geological horizon grid, is exported to and modified within another application, when it is read back into the original application, it is not considered a new data object but appropriately an updated version of an existing object.

The standard does not mandate the creation of a data exchange file; there is also an option to directly transfer data from one application to another, using the Energistics transfer protocol (ETP). ETP, based on the WebSocket full duplex standard, establishes standard messages that allow two applications to assess their ability to exchange specific data objects and directly select and retrieve data without the use of intermediate files, which is a more efficient and faster process.

Scenarios and uncertainties. A modern dataset for a hydrocarbon asset is the result of assembling, assessing, processing, interpreting and analyzing a complex set of data. In the years that it takes the geoscientists and engineers to build and update these data, it happens frequently that the path forward is open to multiple possible solutions. This can be caused by ambiguities in the data, uncertainties in the values for certain objects or knowl-

Fig. 1. Storing packaged datasets at various steps of a workflow requires multiple geoscience applications.



edge gaps for a hydrocarbon system, especially in the early phase of discovery and development. The judgement and experience of the professionals, assisted by analytics, artificial intelligence (AI)

and machine learning (ML), will result in choosing the most likely solution, using the information and knowledge available at that time.

As time goes by and new data are ac-

quired, the “best case” scenario adopted at different steps in the workflow and the life of the field may prove to be flawed. This will require that the team reverts to the step in the workflow at which that decision was made, and execute the subsequent steps with an alternate scenario supported by the new data or updated concepts. In traditional application usage patterns, this would be at best difficult and often impossible. Application-specific archives may no longer be readable, due to version changes, or it may be difficult to pinpoint when the scenario option was exercised, because past team members—who could recall the timing of project work execution—may no longer be part of the business unit.

A solution: Store RESQML project files at each major workflow step. Eni, the global Italian independent energy company, has developed a process to manage efficiently the scenario and uncertainty challenges encountered while operating multiple geoscience applications (Geo-Apps). Illustrated in Fig. 1, the users create a standard RESQML set of files at each major step of the project. The illustration shows these files being created and stored as data for exchanges between different Geo-Apps, but it would be equally possible to create such files at important steps within a large application suite. The application-agnostic nature of RESQML files ensures that these packages are ever-green and can be put to use at any time in the future by a standards-compliant application or data management system.

The e-RESQML project versioning repository system. Based on the RESQML standard, the actual storage of datasets takes place in a system named e-RESQML, which consists of three major components, Fig. 2:

- A database that references dataset snapshots stored on a central server, including the binary content of the RESQML file and specific metadata that identify the step in the workflow and other information that will facilitate identification and retrieval at a later date.
- A web application, allowing the user to upload/download RESQML files or to directly connect to interpretation applications, using the ETP protocol.
- Eni’s data platform, i.e. the company

Fig. 2. Schematic of the components of e-RESQML and their interface to Geo-Apps.

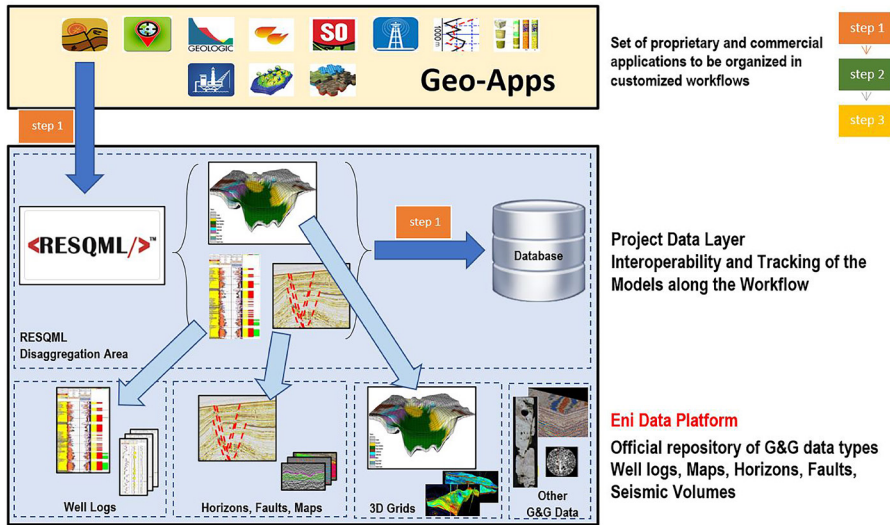


Fig. 3. Choosing the most appropriate RESQML elements to disaggregate and save in Eni’s data platform.

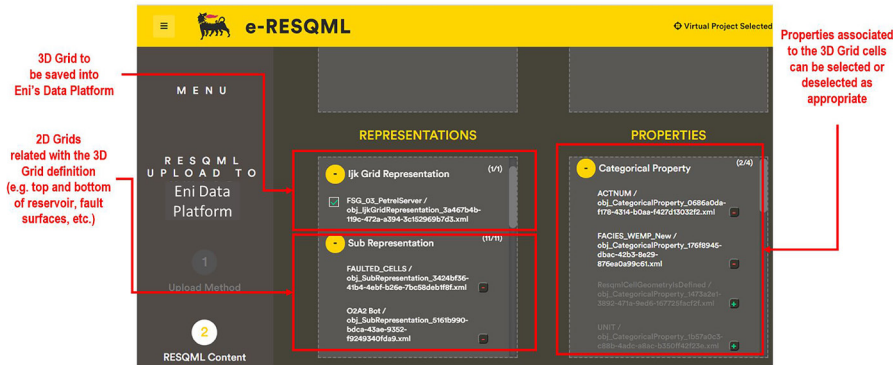
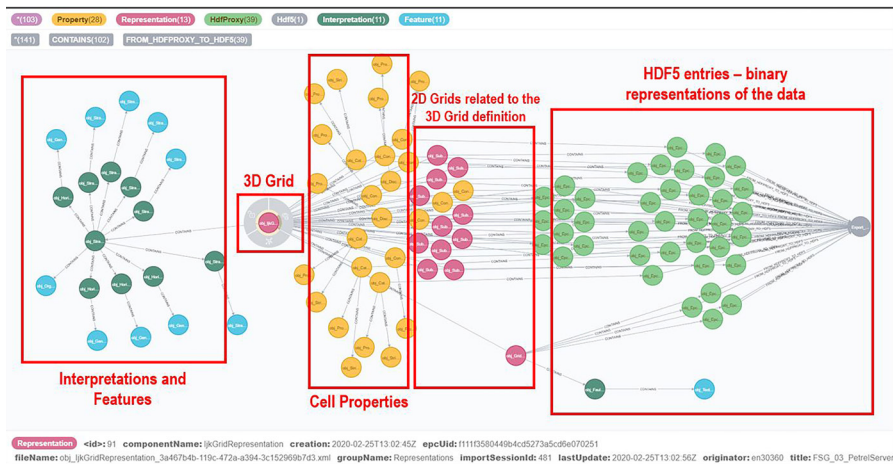


Fig. 4. Analyzing the relationships across elements to ensure consistency of the RESQML model.



platform developed to govern, host and share subsurface data among all the subsurface applications, including well logs, maps, horizons, faults, seismic volumes, etc.

The RESQML model provides the following benefits: 1) Models are saved as a whole, with the packaged content of 3D grids, horizon/fault surfaces and well data that were extracted from the interpretation application at the end of the interpretation step. This packaged content can be consumed directly by applications used in the following steps of the interpretation process, therefore allowing an efficient way of transferring full models across applications; and 2) data are also disaggregated into the basic components, so that each can be saved as a versioned element of a 3D grid, horizon, map or well log data type within Eni's data platform.

This feature is extremely important, because future interpretation efforts might start from the basic data components (e.g. from the latest version of the interpreted horizons and fault surfaces, or from the latest CPI logs), rather than from the packaged RESQML content. It is therefore necessary to allow applications to browse across the Eni data platform and retrieve the latest updated data element at the minimum level of granularity. This would not be guaranteed, if such elements were kept only in the packaged RESQML content.

Therefore, when uploading a RESQML model, users are provided the possibility of previewing the content of the RESQML model and selecting the components to be disaggregated and explicitly saved into Eni's data platform, **Fig. 3**. Upon selection of a data element, such as a 3D grid, the user is shown: 1) All the "sub-representations" that are linked to the grid, including the horizon and fault surfaces that were used for the definition of the grid; and 2) the properties associated with the cells of the grid (e.g. porosity, permeability, water saturation), with the possibility of selecting/deselecting the properties to save within the data platform.

To save a consistent set of objects and preserve all the relationships across the data elements, the e-RESQML application uses a graph database utility (**Fig. 4**) to identify:

- The selected 3D grid element, shown in the center-left section of the figure.
- All the attached properties, shown

Fig. 5. Metadata input panel for each version of the model.

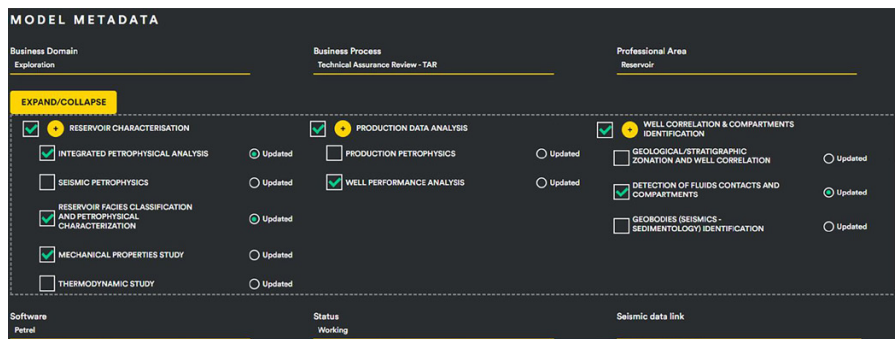
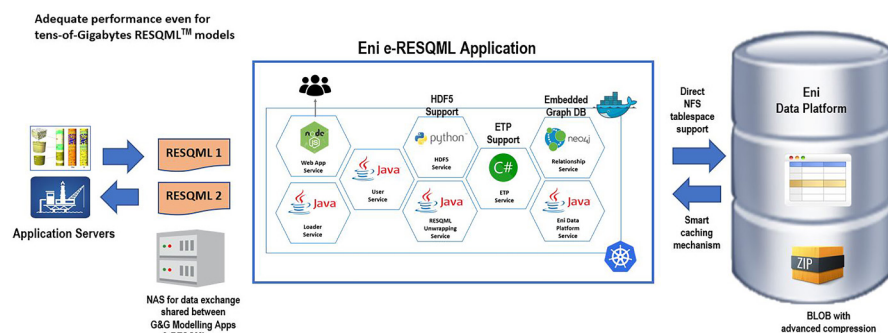


Fig. 6. Architecture of the e-RESQML application and technological choices.



as yellow circles in the figure.

- All the attached sub-representations, shown as red circles.
- All the attached "interpretation" and "features", i.e. the RESQML elements that add semantics to the 3D grid object. These are shown in green and light-blue colors in the left part of the figure.
- All the links to the appropriate section of the HDF5 file, where the binary content is stored. Such links are shown in light green in the right part of the figure.

After the right set of elements to be disaggregated into Eni's data platform has been identified, the key to this system's success is the metadata describing the RESQML model content (**Fig. 5**), including two important sets of metadata: 1) In the upper part of **Fig. 5**, information about the workflow and the interpretation step, where the model was generated is collected; and 2) in the lower part of the figure, information about the semantic content of the RESQML model is collected. For example, the user can specify that, within this step of the workflow, a new mechanical property model and well performance analysis model were created, while the integrated petrophysical analysis and the facies classification

models, which were already present in the previous step of the workflow, have been updated. Such metadata allow the user to easily identify the correct starting point from which future reservoir studies can restart. For example, if new well data became available and the reservoir facies distribution must be recomputed, the reservoir specialist can start from the RESQML file with the most recent "reservoir facies classification and petrophysical characterization."

From the technological point of view, the e-RESQML application has been designed carefully to work, even with tens-of-gigabytes RESQML models, **Fig. 6**. Users run the interpretation applications from an "application farm," where dedicated NAS areas are made available for RESQML file exchange. The NAS areas are also made visible to the application server, where the e-RESQML application runs, and even to the back-end database server. This allows the system to shift the data from the application server to the database server in a highly efficient manner, moving them within the same storage volume, without time-consuming copy operations of the data.

Moreover, the RESQML files are uploaded into the database binary storage in the background, while keeping a

cached local copy of the file on the file system, which will be retained for a specified amount of time. This guarantees the user a very quick interaction with the system: upload and download operations are perceived as immediate, because the cached copies are moved quickly from one storage volume to another, to make them visible again to the application servers. Moreover, the usage of BLOB storage with advanced compression within the database provides an average 15-to-1 compression ratio, optimizing the storage space and making the upload/download operations even quicker.

As shown in **Fig. 6**, the e-RESQML application is designed as a microservice containerized application. It exploits a graph database technology for the tracking of the relationships across RESQML data elements, HDF5 file support, and exposing both ETP server and ETP client functionalities to directly interact with interpretation applications.

CONCLUSION

The e-RESQML system has demonstrated the ability to store, in a struc-

ture environment, successive versions of a subsurface and reservoir dataset for a hydrocarbon asset with ample metadata. The solution enables future users to identify each dataset version and choose to revert to a critical step in a workflow to explore alternate scenarios or revisit critical uncertainties. The use of the existing RESQML industry standard reduced the development effort and ensures that the solution is future-proof and compatible with a large number of products available within the Geo-App market. [WVO](#)

MARCO PIANTANIDA is technical advisor for technical/scientific applications and processes at Eni SpA, Milan, Italy. Mr. Piantanida joined Eni in 1992, and has been integrally involved with developing multiple, technical IT projects. He is taking part in the digital transformation program at Eni, specifically dealing with the company's data platform, in addition to machine learning applications and cognitive technologies.

CLARA ANDREOLETTI joined Eni in 2002 as an R&D geophysicist, working on development of the company's proprietary seismic imaging platform. In 2011, she moved to the exploration department and has held various management positions since then. She currently leads the subsurface data governance and management group. Ms. Andreoletti graduated with a

degree in telecommunications engineering from Politecnico di Milano in 2001.

MATILDE DALLA ROSA is the head of the geophysical data management unit within Eni's exploration department. She is project manager for the development of Eni's subsurface data platform within the digital transformation program. She has co-authored multiple technical publications on geological modeling and geostatistics. Ms. Dalla Rosa joined Eni, working in the geoscience division, after graduating from Politecnico di Milan.

BRUNO VOLPI is senior advisor for reservoir study and innovation at Eni SpA in Milan, Italy. Previously, he was department manager for reservoir characterization and petrophysics for eight years before taking on his current role as focal point for reservoir R&D activities within Eni, cooperating with the Eni data platform project. Mr. Volpi joined Eni in 1988, working mainly in the reservoir and reservoir management area, with involvement in multiple reservoir studies and technology innovation projects.

PHILIP NERI is marketing director for Energistics in Houston, where he drives the adoption of digital data exchange standards. Mr. Neri started his career as an explorer for Shell in Europe and Asia. He also worked on design, development and commercialization of subsurface software and data management solutions for software providers in Europe and the U.S., including Schlumberger and Emerson. He holds a BS degree in geology and an MS degree in geophysics and computer science.