

Journée Intéropérabilité et Innovation :
Le FAIR data en appui à la compréhension
de la Terre et de son environnement
Atelier thématique : DWG métiers OGC

Université Gustave Eiffel - Bâtiment Bienvenue
16 Av. Newton, 77420 Champs-sur-Marne, France
23-01-2024

Le format RESQML v2.2, le standard pour les Géosciences
100% respectueux des principes FAIR

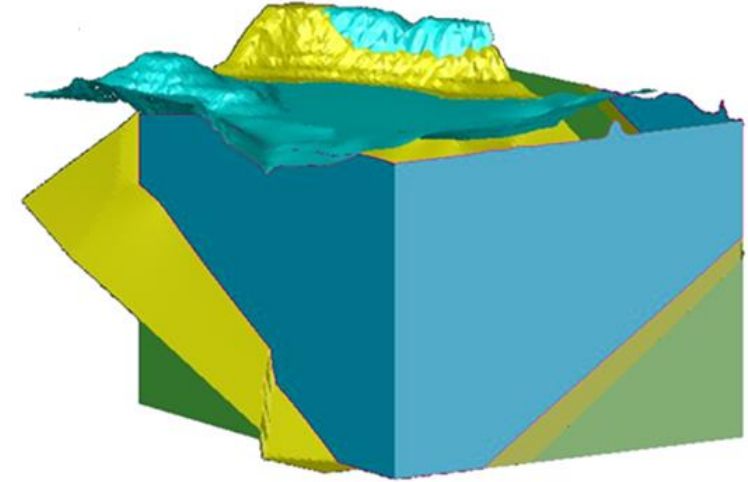
Jean-Marie LEONARD



Le format RESQML v2.2, le standard pour les Géosciences 100% respectueux des principes F.A.I.R.

1 – Principes généraux et contenus des fichiers RESQML

- Qu'est-ce que RESQML ?
- L'environnement logiciel autour de RESQML
- Structure générale et contenus des fichiers



2 - Le format RESQML et les principes F.A.I.R.

- Les principes directeurs F.A.I.R.
- Analyse F.A.I.R. du format RESQML

3 – Conclusion :

- RESQML un bon standard pour servir de la donnée publique du sous-sol



Principes généraux et contenu des fichiers RESQML

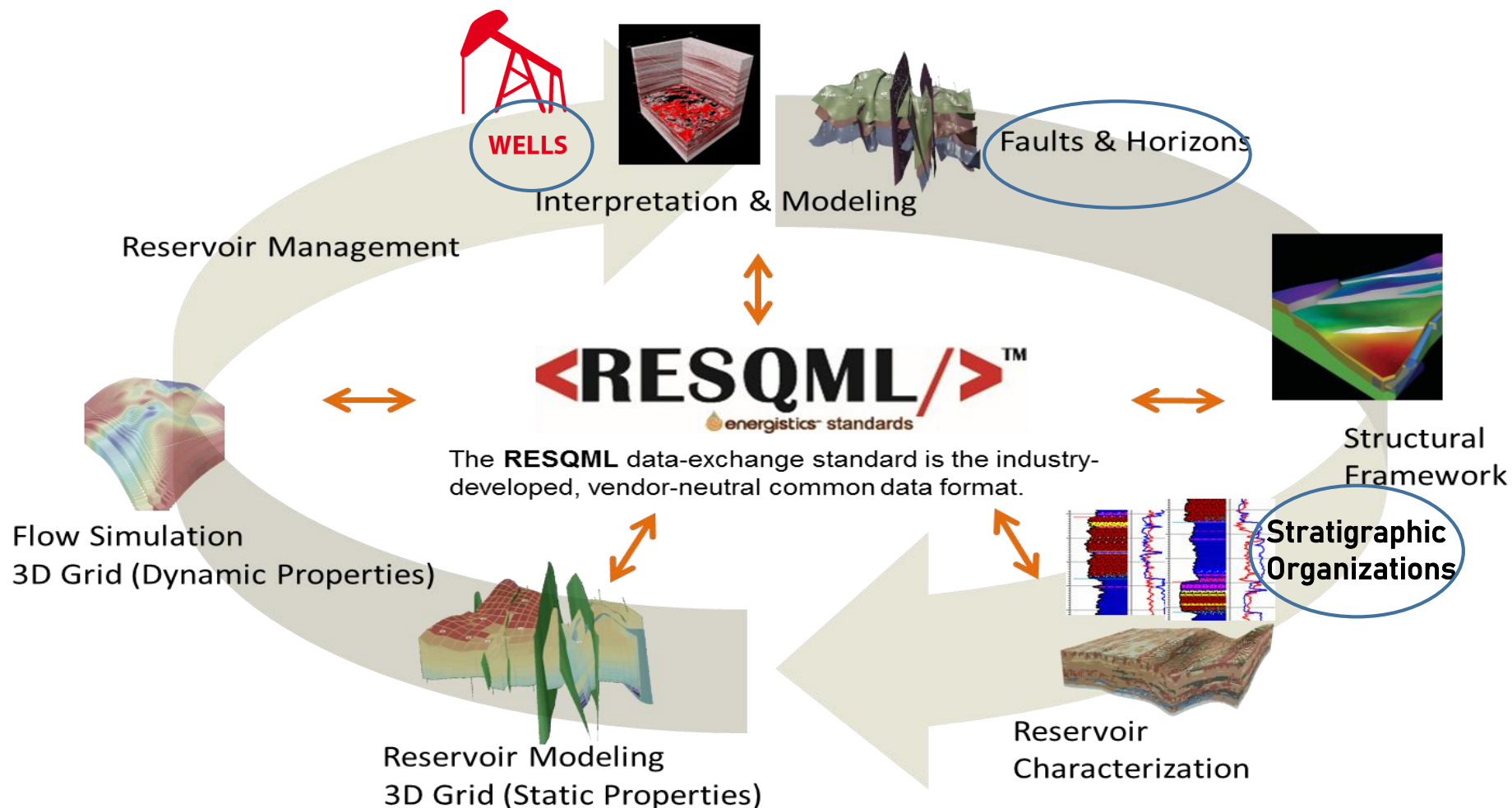


01.

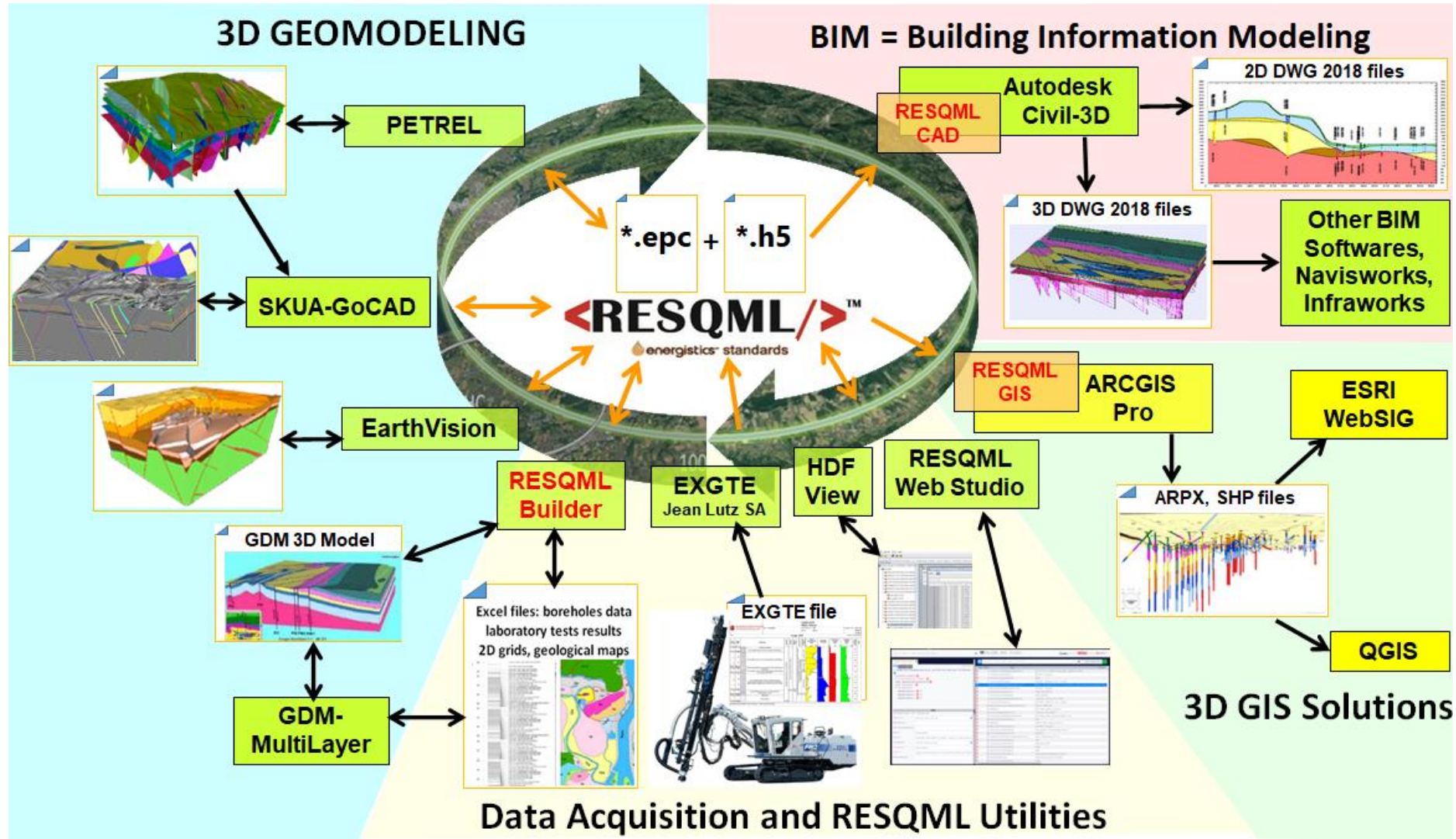
Qu'est-ce que RESQML ?

Ce standard a été conçu par l'industrie pétrolière pour stocker et échanger les données et les interprétations utiles à la modélisation 3D de réservoirs pétroliers

RESQML Supported Workflows



La plupart des objets et données manipulés pour l'industrie pétrolière constitue un ensemble commun d'objets et de données utilisées pour les différentes disciplines des géosciences. Pour les besoins spécifiques à une discipline, par exemple la géotechnique, il sera facile de compléter ce qui manque



❑ RESQML v 2.0.1 (Energistics)

→ Standard implémenté dans Petrel, GoCAD, EarthVision, etc

❑ EGIS a créé des outils basés sur RESQML 2.2-beta

→ Transfert des données et des modèles 3D vers Civil-3D et ArcGIS Pro

❑ RESQML v2.2 - version officielle : 16 Mars 2022

→ Futures mise à jour de Petrel, GoCAD, EarthVision, etc

❑ Les standards RESQML, WITTSM et PRODML du consortium ENERGISTICS sont maintenant repris par l'OPEN GROUP

→ Intégration de RESQML dans la plateforme **OSDU™** (Open source, cloud-native, subsurface reference architecture).

Organisation des outils pour le projet FCC (Future Circular Collider) du CERN à Genève (Suisse)

❑ Separation of data and metadata in two files

❑ The EPC file :



EPC = Energistics Packaging Conventions :

it is a variant of the OPC (Open Packaging Conventions) file, initially created by Microsoft.

Archive (as a ZIP file) containing two kinds of XML files (Text files in markup language readable by a human and a computer) :

- **TopLevelElement files : they contain the description metadata** of any element of type Feature, Interpretation or Representation plus the associated properties. A topLevelElement is identified with an UUID (Universal Unique Identifiant).
- **Relation files :** they describe the relation between two TopLevelElements.

❑ The HDF5 file :



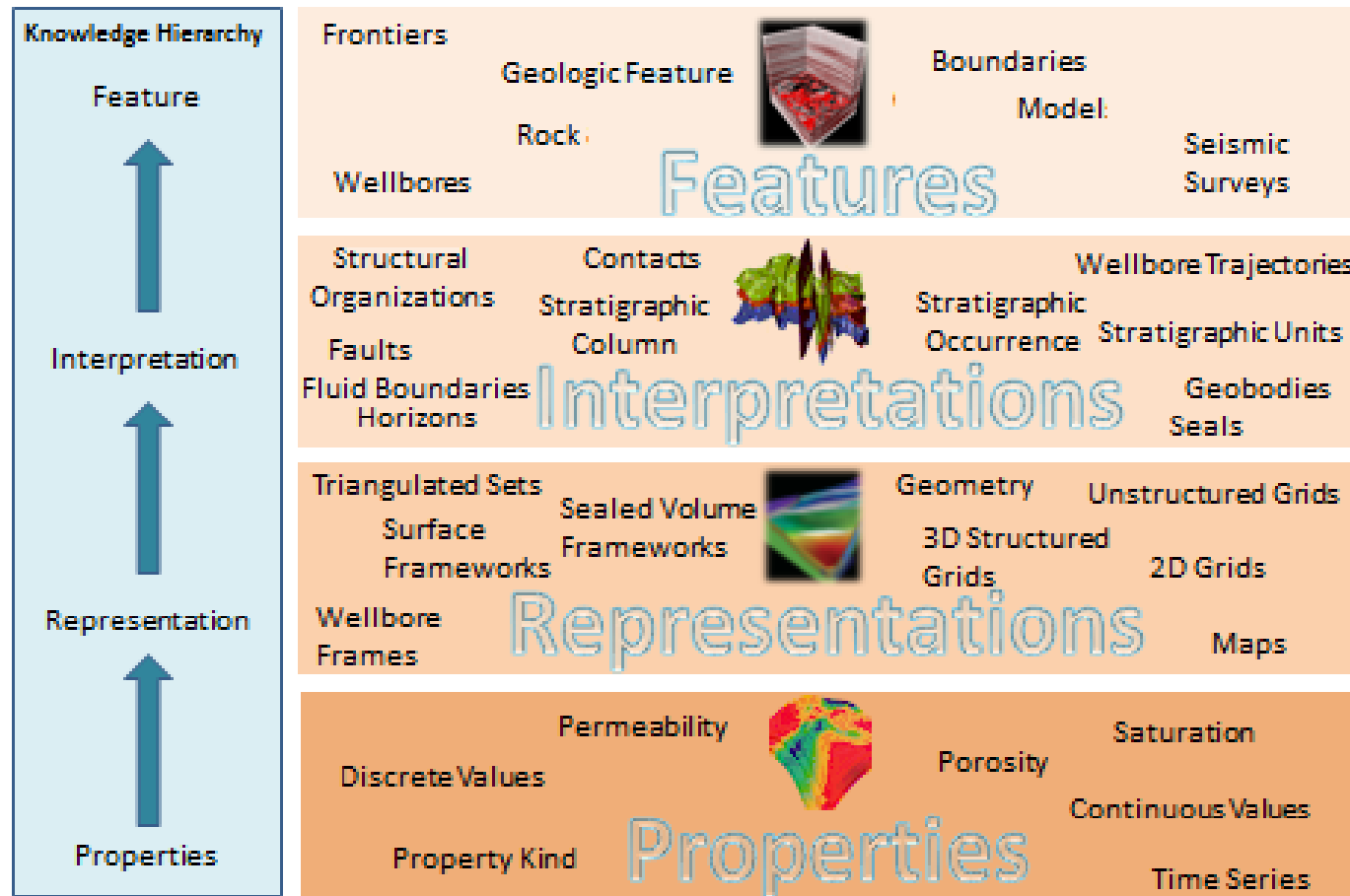
HDF5 = Hierarchical data Format, version 5

developed at the NCSA (National Center for Supercomputing Applications)

Structured binary file allowing to store the numerical data such as the coordinates (X,Y,Z) of the summits and the edges of the triangles of a TriangulatedSetSurface (a TIN) or the values of pressuremeter data. A system of reference allows us to find the data associated to a TopLevelElement..

❑ THE FIRP : Feature / Interpretation/ Representation/ Properties

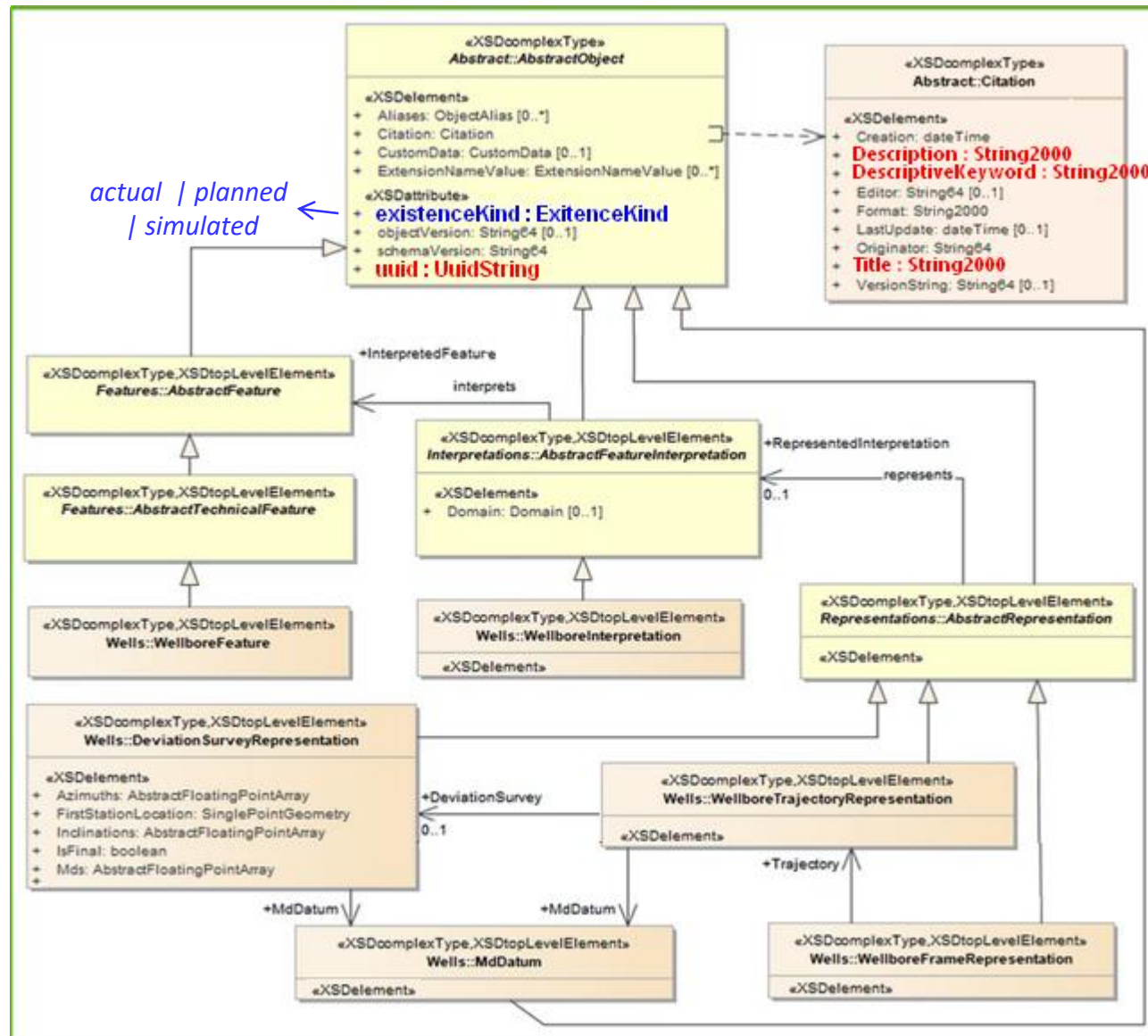
RESQML Knowledge Hierarchy and Data Relationships



- Reference (Individuals and Models)
- Interpretation Meta Information
- Topology & Geometry
- Properties (attached to topology)

Structure générale et contenus des fichiers RESQML

- The WellboreFeature, the WellboreInterpretation and the Wellbore Representation Elements



WellboreFeature

Descriptive keywords : type of the borehole on 2 characters : PM = trial pit, SC= core borehole, SD = destructive borehole, SPT = standard penetration test, CPT = cone penetration test, ...

Title : Borehole name

WellboreInterpretation

Description : log_litho | log_struct | log_strati | ... | log_Pressio...

Descriptive keywords : [Lithol] | [Struct] | [Strati] | ... | [Pressio]...

Title : Borehole name + Description

WellboreFrameRepresentation

- supports a sampling (a list of mdValues or a lattice of values)

Description, descriptive keywords and Title : same data as for the WellboreInterpretation

WellboreTrajectoryRepresentation

- between the startMd = 0.00 meter and the finishMd = total depth of the borehole

Description : borehole name

Descrip. keywords : VERTICAL | HORIZONTAL | INCLINED | DEVIATED

Title : Borehole name + Description

MdDatum

- Represents the head point of the borehole using a SinglePointGeometry associated a LocalDepth3dCrS element

DeviatedSurveyRepresentation

- If the borehole is non vertical and not drilled downward
- Defines 1 couple of azimuth and inclination measures at the startMd = 0 (HORIZONTAL, INCLINED, upward VERTICAL)
- Defines n couples of azimuth and inclination measures associated md values of stations (DEVIATED)

Le format RESQML et les principes F.A.I.R.

02.

Box 2 | The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Les données contenues dans les fichiers au format RESQML sont-elles facilement :

☐ Trouvables ,

☐ Accessibles ,

☐ Intéropérables ,

☐ Réutilisables ,

- par les logiciels

- et par les humains ?

Wilkinson et al. (2016) The FAIR Guiding Principles for Scientific data management and stewardships. Scientific data. March 2016.

- ❑ **F1. Les (méta)données se voient attribuer un identifiant unique et persistant au monde.**
 - Les Top Level Elements ont un identifiant unique composé par leur UUID (Universal Unique Identifier), la définition de leur type et une indication de version,
 - Les **UUID sont générés selon la norme RFC4221 (ISO / OGC)**
Attention : Il faut cependant gérer les éléments réemployés à partir de précédents fichiers pour qu'ils réutilisent les UUID
- ❑ **F2. Les données sont décrites avec des métadonnées riches (définies par R1 ci-dessous)**
 - Les données numériques sont contenues dans les fichiers .HDF5 mais elles sont rattachées aux différents éléments de type Représentation
 - Chaque type de Représentation est définie par un ensemble de métadonnées, de Propriétés et a son système d'indexation.
 - Ensuite ces éléments de type Représentation caractérisent des Objets métiers qui contiennent, eux, des informations décrites de manière sémantique
- ❑ **F3. Les métadonnées incluent clairement et explicitement l'identifiant des données qu'elles décrivent**
 - Chaque Type de Représentation contient un identifiant unique « Type d'élément, UUID, Version » qui permet d'identifier le fichier XML ainsi qu'une adresse dans un fichier hdf5 (de type URI du fichier h5, UUID, Type de Valeur) qui permet de retrouver les valeurs numériques (x,y,z ou propriétés).
 - Avec RESQML v2.2, il est possible d'utiliser plusieurs fichiers HDF5 différents contenus sur un PC ou dans les clouds en utilisant le serveur HSDS du HDF Group)
- ❑ **F4. (méta)données sont enregistrées ou indexées dans une ressource consultable**
 - Les métadonnées sont enregistrées dans des fichiers XML séparés correspondant à un seul Top Level Element. Ces fichiers .xml sont nommés à partir du nom du type d'élément et son UUID (Universal Unique Identifier) et réunis dans une archive .EPC
 - Chaque composant possède donc une URI



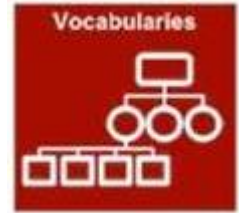
Réponse : Oui

- ❑ **A1. Les (méta)données sont récupérables par leur identifiant à l'aide d'un protocole de communication standardisé :**
 - Le standard EPC permet de regrouper les fichiers XML et de les transférer avec le fichier hdf5 par le protocole f
 - Le standard de communication ETP (Energistics Transfer Protocol » permet de transférer des composants XML et Hdf5 un par un ou en groupe en utilisant les Web sockets développés pour la plateforme OSDU.
- ❑ **A1.1 le protocole est ouvert, gratuit et universellement mis en œuvre :**
 - La version finale du format RESQML v2.2 est décrite sur le site du consortium « The Open Group ».
 - L'utilisation du format est totalement gratuite et libre de droit,
 - Des API en différents langages sont disponibles gratuitement pour faciliter le travail des développeurs.
- ❑ **A1.2 le protocole permet une procédure d'authentification et d'autorisation, si nécessaire :**
 - Le fichier .EPC est une archive de type .zip qui peut être sécurisée par mot de passe
 - Le système FTP de transfert peut être protégé par des mots de passe aussi bien coté EPC que hdf5
 - Le protocole ETP satisfait aux exigences de sécurité actuels
- ❑ **A2 Les métadonnées A2 sont accessibles, même lorsque les données ne sont plus disponibles :**
 - Oui on peut lire un fichier EPC qui n'a pas son fichier associé HDF5. C'est ce que font les logiciels RESQML Editor et RESQML WebStudio
 - Chaque fichier XML contenue dans le fichier EPC est compréhensible par un humain qui connaît le schéma conceptuel de données.



Réponse : Oui

- ❑ I1. Les (méta)données utilisent un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances :
 - Le modèle conceptuel de données utilise le langage UML 2.0 pour décrire tous les composants.
 - Le fichier EPC est un container classique de type archive ZIP basé sur le modèle OPC de Microsoft. Les fichiers qu'ils contiennent sont en langage XML et font référence au modèle de données.
 - Le fichier binaire contenant les données est au format HDF5 dont la structure est normalisée par le HDF Group
- ❑ I2. Les (méta)données utilisent des vocabulaires qui suivent les principes FAIR :
 - Le vocabulaire utilisé pour désigner les composants et les attributs est formé de mots de la langue anglaise dont le sens est assez implicite. Il est choisi par des spécialistes de la géologie et qui ont une bonne connaissance des modèles de données.
 - Le guide technique de RESQML permet aux utilisateurs de bien comprendre le vocabulaire technique à l'aide des schémas normalisés UML 2.0 .
 - Le vocabulaire s'appuie sur des dictionnaires officiels (PWLS, GeoSciML etc..)
- ❑ I3. Les (méta)données incluent des références qualifiées à d'autres (méta)données:
 - Les relations entre les différents composants du modèle utilisent des noms explicites pour qualifier les relations conformément aux exigences de la norme UML 2.0



Réponse : Oui

- ❑ R1. les méta(données) sont richement décrites avec une pluralité d'attributs précis et pertinents :
 - Tous les éléments du modèle de données sont décrits avec des attributs dont le typage et la signification est précisé dans la documentation technique.
 - Certains attributs peuvent être contraint par le choix d'une valeur comprise dans une liste énumérée. C'est ce qui permet d'avoir une description précise et de pérenniser et de partager l'information..
- ❑ R1.1. Les (méta)données sont publiées avec une licence d'utilisation des données claire et accessible :
 - Les données publiées sont sous licence Apache 2.0
- ❑ R1.2. Les (méta)données sont associées à une provenance détaillée :
 - La provenance est indiqué par un objet Citation qui est systématiquement attaché à chaque TopLevelElément du modèle de données.
 - On y trouve des informations sur la date de création, la date de modification, le nom de l'auteur et le nom du logiciel qui a généré l'élément contenant la métadonnée et la donnée associée.
 - Ces informations sont basées sur la norme Dublin Core qui fait l'objet d'une spécification ISO à l'internationale depuis 2003.
- ❑ R1.3. Les (méta)données répondent aux normes communautaires pertinentes pour le domaine :
 - RESQML est basé sur de nombreux standards comme GML (Geography Markup Language) et le géo référencement systématique des données selon les systèmes de coordonnées référencés par l' EPSG (European Petroleum Survey group), la norme Dublin Core, un modèle conceptuel de données basé sur UML 2.0, les formats XML et HDF5



Réponse : Oui

**En guise de conclusion :
RESQML un bon standard pour —
servir de la donnée publique !**

RESQML v2.2 : standard de stockage et de distribution de la donnée publique du sous-sol

Journée Intéropérabilité et Innovation
Champs-sur-Marne (77) - 23-01-2024

- ❑ Le format RESQML v2.2 est relativement complet, il peut stocker :
 - les données de sondages et leurs interprétations (description sondeur, interprétations géologiques, géotechniques, hydrogéologiques, etc.
 - les modèles 3D du sous-sol avec différentes interprétations et différentes représentations
 - ❑ Le format RESQML respecte tous les principes F.A.I.R.
 - ❑ Les données et métadonnées du format RESQML peuvent être distribuée par un serveur de données
 - ❑ Il est utilisé par des logiciels du commerce et il est possible de développer des outils qui importe ou exporte des données et des modèles au format RESQML v2.2
 - ❑ Le format RESQML peut être utilisé conformément aux recommandations européennes INSPIRE
- ➔ Toutes ces conditions sont favorables pour qu'un bureau géologique puisse utiliser ce format pour :
- recueillir des données de sondages, d'essais laboratoire et des modèles 3D et leurs diverses interprétations
 - les vérifier, les interpréter ou les réinterpréter
 - distribuer les données brutes et les différentes interprétations des sondages et des modèles
- ➔ Il peut donc être créer un système dynamique d'échanges entre un bureau géologique, les différents types d'acteurs et d'utilisateurs (chercheurs, écoles, entreprises de génie civil, collectivités territoriales, personnes privées, etc) .

Suppléments :

04.

Energistics (2024). WITSML v2.1, RESQML v2.2 and PRODML v2.2 bundle.

<https://publications.opengroup.org/downloadable/customer/products/>

Léonard J.M., Labourg P., Gauthier V. & Rainaud J.F. (2023). Update on the development of tools for transferring boreholes data and 3D models using the RESQML v2.2 format.

6th European Meeting on 3D geological modelling – Copenhagen, Denmark, 23 - 26 May 2023.

Léonard J.M., Gauthier V. & Rainaud J.F. (2022). Geotechnics and subsurface information : ontologies and data model of the RESQML v2.2 format with examples of interoperable transfers between geomodelling, BIM and GIS softwares.

Infra BIM Open 21, 10-12th of January 2022 - Lyon, France.

Gauthier V., Labourg P., Léonard J.M., Rainaud J.F. (2021). Operational Use Cases Using RESQML Standard to communicate Geotechnics and Subsurface Information to BIM. – *EAGE Near Surface Geoscience 21, 2nd Conference on Geophysics for Infrastructure Planning, Monitoring and BIM - Bordeaux, France – August 31st 2021.*

Grellet S., Liu Y., Boisvert E., Simons B., Rainaud J.F., Lorenz H., Haener R., Beaufile M., Hollingsworth J., Lieberman J. et al. (2020). OGC Borehole Interoperability Experiment Engineering Report. September 2020 OGC Meeting - Virtual, Open Geospatial Consortium, Sep 2020, Orléans (visioconférence), France. hal-03943388

Léonard J.M. (2020). Interoperable transfer of 3D Geomodels from GDM-Multilayer to Autodesk BIM-CAD softwares using the RESQML v2.2 format: the EGIS solution – *RING Meeting 2020, 9th of September 2020, ENSG, Vandoeuvre-les-Nancy, France.* [\(Lien\)](#)

Léonard J.M. & Le Hello B. (2019). From BIM to Industry 4.0 – Egis BIM Strategy for Metro. Part 3: interoperability and BIM *Autodesk University LONDON – 18th of June 2019, London, UK.*

Léonard J.M., Zhu-Colas B. & Rainaud J.F. (2019). Interest of accurate 3D Geomodels for Geotechnical Projects and Interoperability between Geology and BIM - Part 2 : Interoperability with RESQML Builder and RESQML-CAD *5th European Meeting of 3D geomodelling, 21-24th of May 2019, Bern, Switzerland*

UTILISER L'EDITEUR DE FICHER RESQML V2.2

EPC Resqml V2.2 Editor - E:\GLP\GPA-Guadeloupe\RESQMLv2.2\GPA-GUADELOUPE-POINTE-JARRY_Sondages_2021-07-21.epc

File Edit Navigate Show View Resqmlv2 Editor Validate Help

CategoricalProperty_f8806c56-5e56-487a-a451-29d3f14895f3.xml

Resource Set

- file:/E:/GLP/GPA-Guadeloupe/RESQMLv2.2/GPA-GUADELOUPE-POINTE-JARRY_Sondages_2021-07-21/CategoricalProperty_
- Document Root
 - Categorical Property actual
 - Citation SP170 code_litho
 - Data Object Reference propertyKind - PropertyKind : [Lithol] code_litho
 - Data Object Reference supportingRepresentation - WellboreFrameRepresentation : SP170 log_litho
 - Patch Of Values 0
 - Integer External Array -1
 - External Dataset
 - External Dataset Part 1
 - Data Object Reference lookup - StringTableLookup : code_litho

(a)

Properties Log View

| Property | Value |
|-----------------------|---|
| Count | 1 |
| Path In External File | /RESQML/c48fdec2-108b-4462-922a-75797fd81051/code_litho |
| Start Index | 0 |

UTILISER HDF VIEW POUR CONTRÔLER LES DONNÉES DES FICHIERS HDF (**.H5)

Recent Files E:\GLP\GPA-Guadeloupe\RESQMLv2.2\GPA-GUADELOUPE-POINTE-JARRY_Sondages_2021-07-21.h5

(b)

code_litho

desc_litho

mdValues

nom_litho

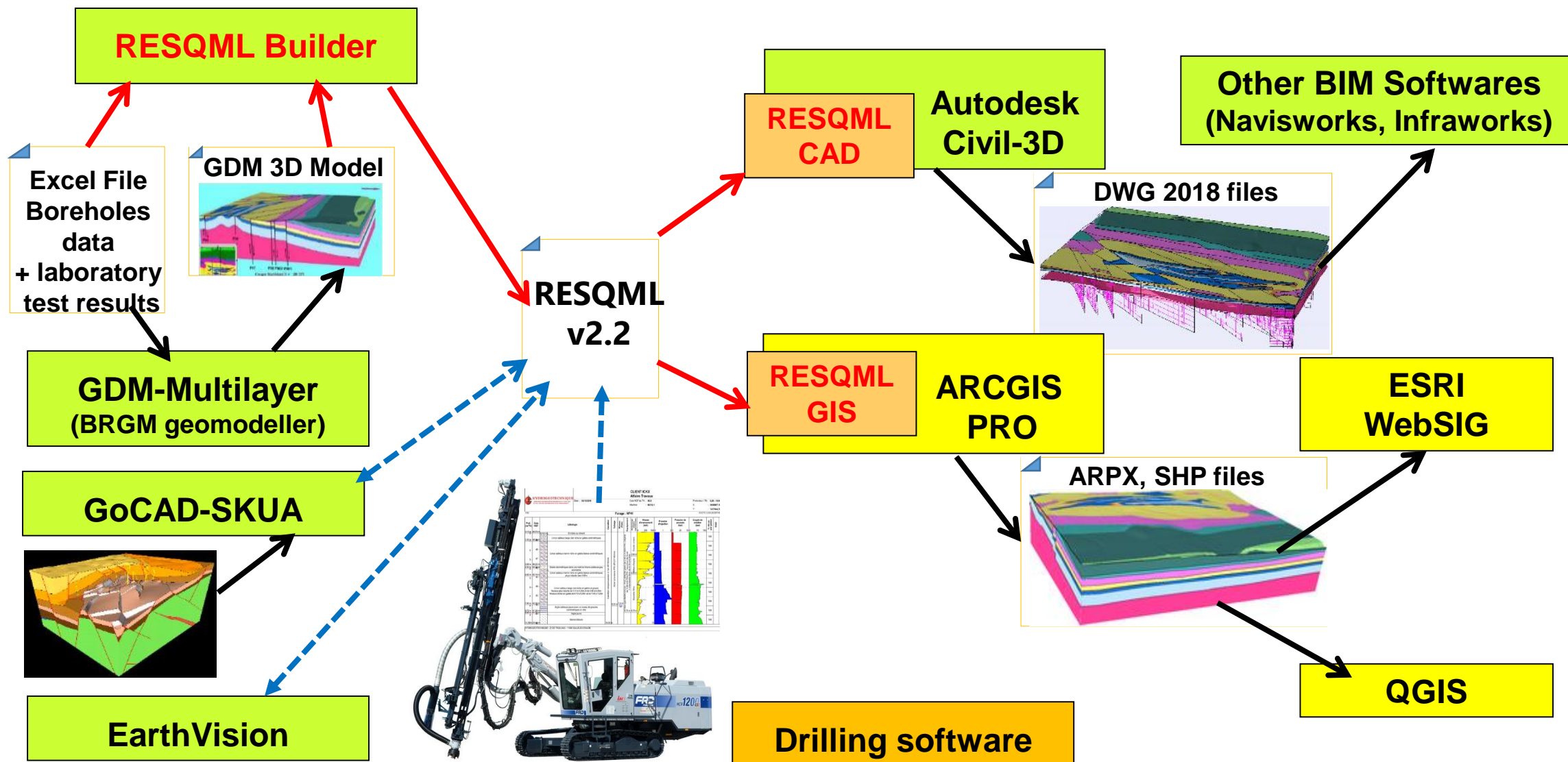
mdValues at /RESQML/c48fdec2-108b-4462-922a-75797fd81051/ ...

| Table | |
|-------|------|
| 0 | 0.0 |
| 1 | 3.0 |
| 2 | 4.0 |
| 3 | 5.5 |
| 4 | 9.0 |
| 5 | 28.0 |

code_litho at /RESQML/c48fdec2-108b-4462-922a-75797fd81051/ ...

| Table | |
|-------|----|
| 0 | 23 |
| 1 | 2 |
| 2 | 10 |
| 3 | 7 |
| 4 | 25 |

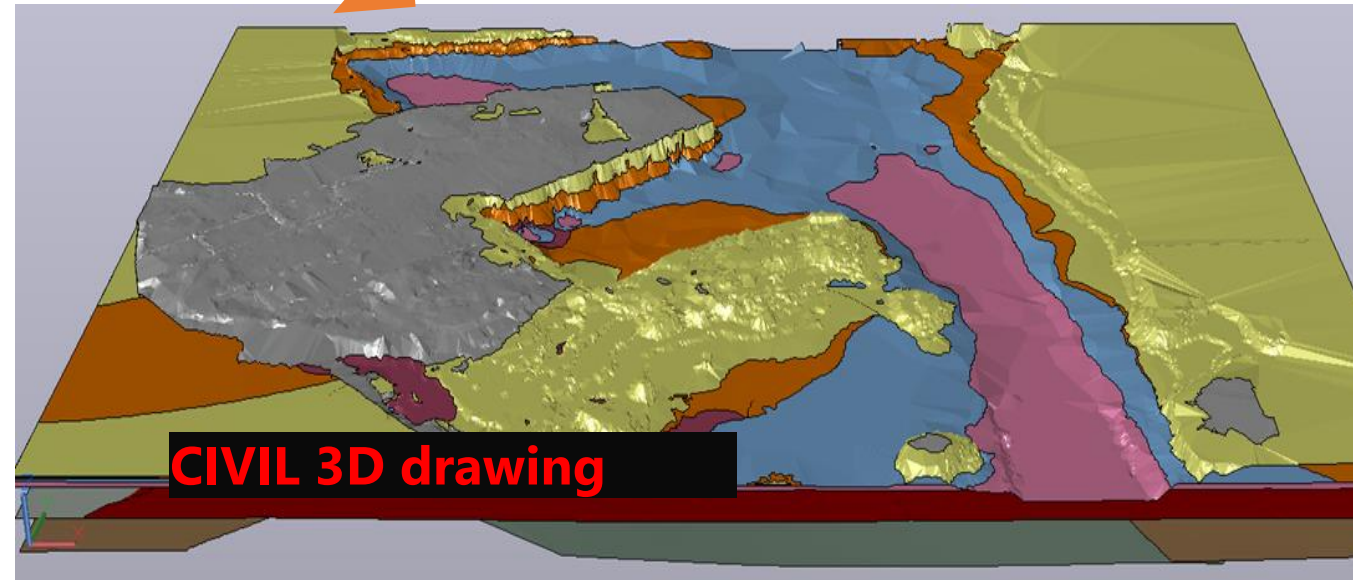
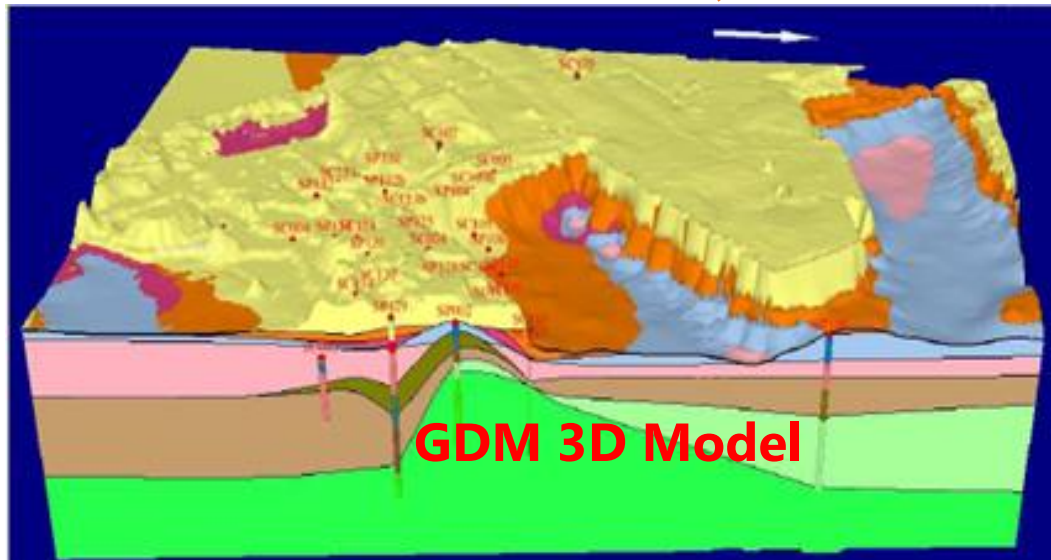
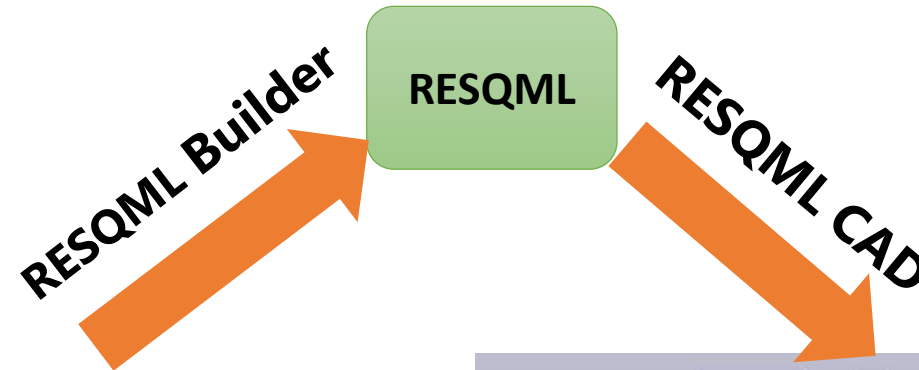
□ Développement d'outils par EGIS



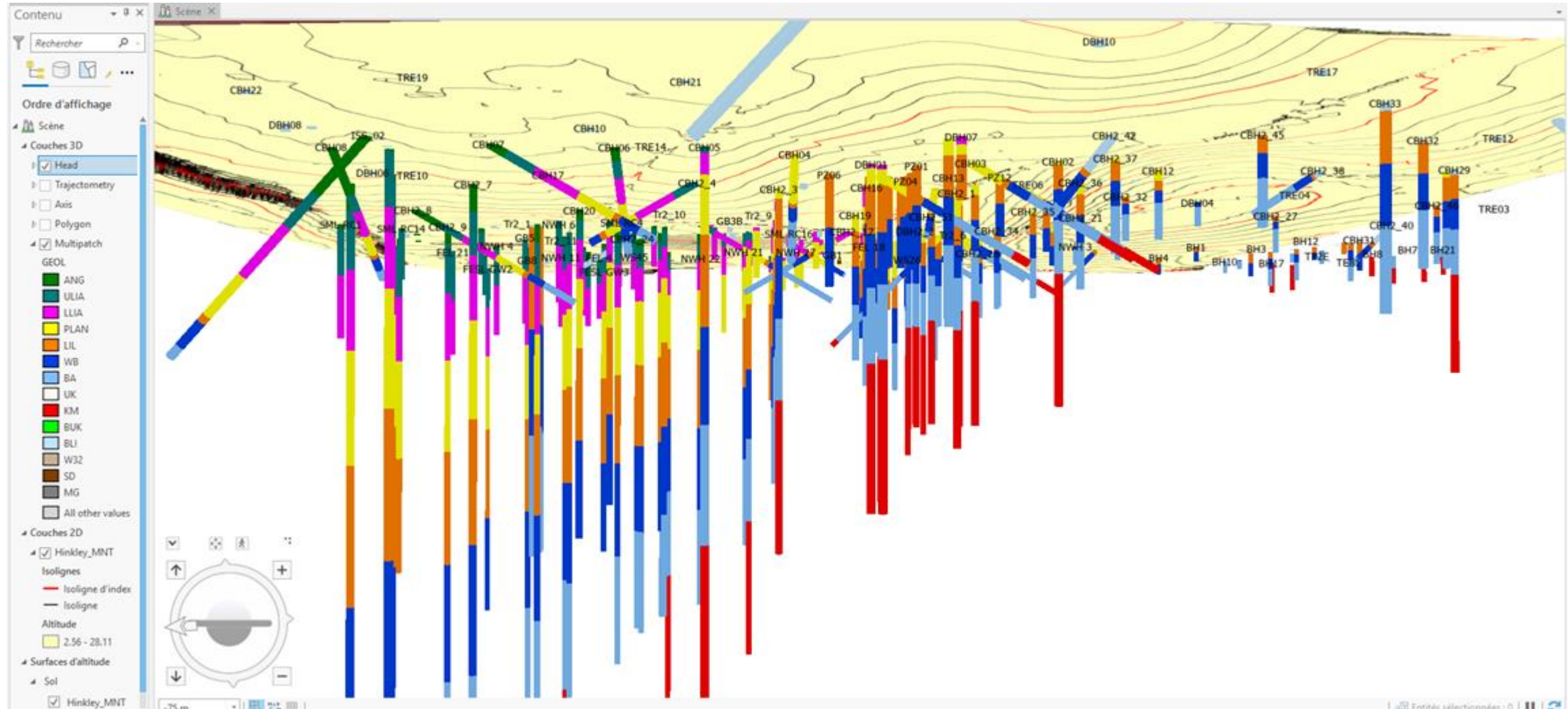
RESQML-CAD : Import sous Civil-3D

| Formation | Type Surf. | Description |
|-----------|------------|------------------------------|
| Rb | EROD | Remblai |
| 3a1 | | Sables coquilliers |
| 3a2 | | Vase, silts, sables argileux |
| 3b | | Tourbe |
| 3c | | Argiles bleues |
| 4a | | Argiles bariolées et tufs |
| 4b1 | | Argiles d'altération |
| 4b2 | | Argiles à blocs |
| 4c | | Calcaires altérés |
| 4d | | Calcaires sains |

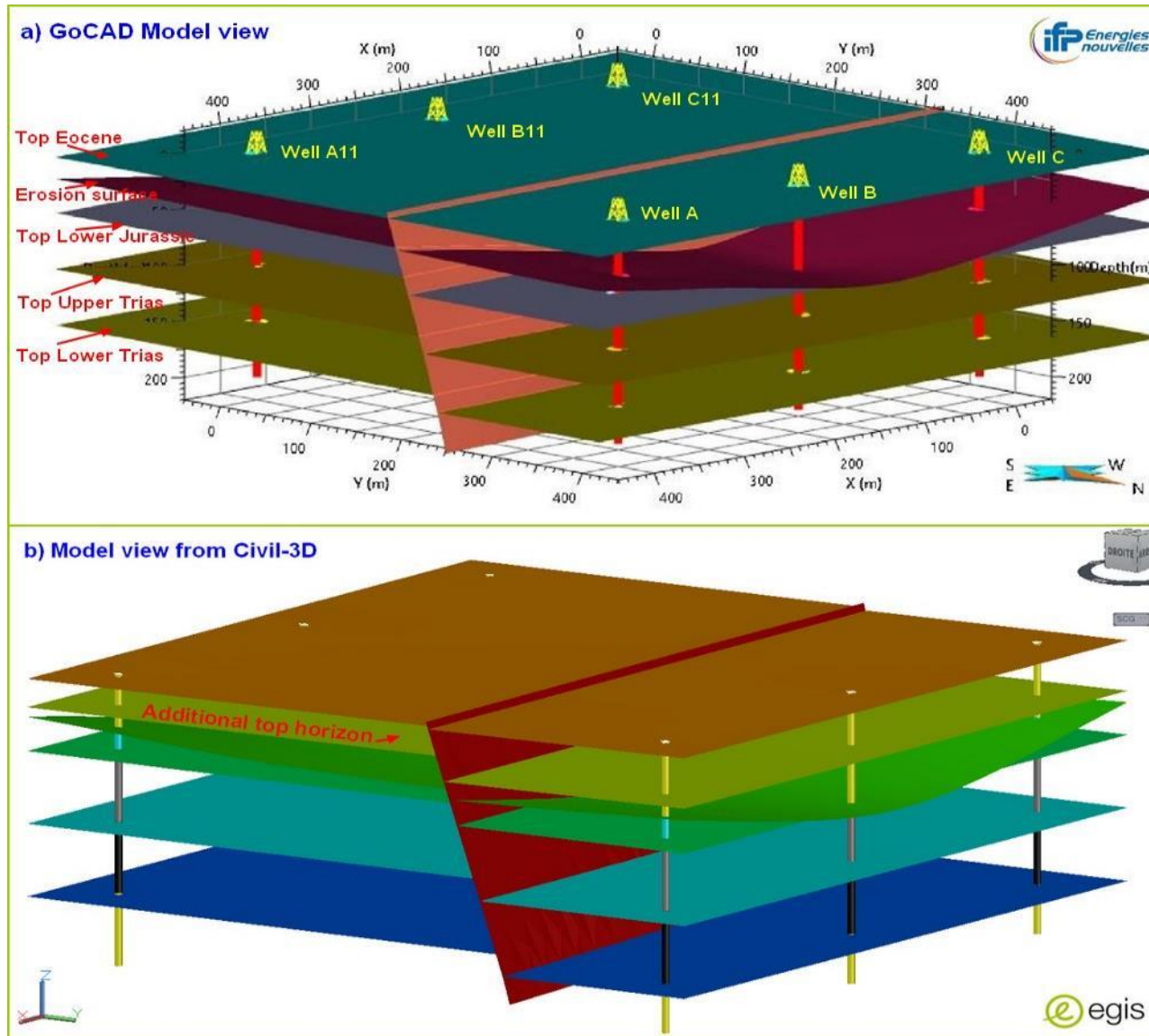
Guadeloupe Harbour expansion 3D Geotechnical Model



RESQML-GIS POUR ARCGIS PRO

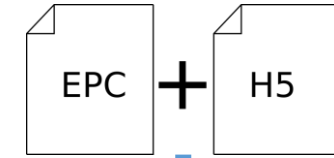


Vertical and inclined boreholes of the Hinkley Point EPR project (UK). The DTM comes from a LandXML import



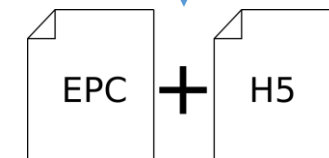
GoCAD Model delivered by IFP Energies nouvelles

GoCAD



RESQML Converter
Migration 2.0.1 to 2.2

RESQML WebStudio
Manually add attributes
+ color info and validation



Civil-3D +
RESQML-CAD