# Amalgamation of Diverse Data Types and Sources to Facilitate Data Analytics of Engineering Interpretation and Historical Data*

**Cesar Velasquez[1], Ivan Olea[1], and Russell Roundtree[1]**

Search and Discovery Article #42483 (2019)**
Posted January 6, 2020

*Adapted from oral presentation given at 2019 International Conference and Exhibition, Buenos Aires, Argentina, August 27-30, 2019
 DOI:10.1306/42483Velasquez2019

[1]IHS Markit, London, United Kingdom (cesar.velasquez@ihsmarkit.com)

## Abstract

Field development planning of unconventional assets can benefit from learnings from historical completion practices and their production performance results. For this, it is necessary to have a consistent, integrated data source. The Oil and Gas business has been challenged for decades to have industry standard data schemas adopted across the entire industry when consensus on those data elements lags the need to capture new data types. In the age of "if it can be digitized, it will be digitized", proliferation of data types and data volumes are outpacing our historical ability to respond. As our industry becomes ever more "data driven", the need to efficiently use diverse data types from many sources and make these data available to a host data analytics tools to extract information and knowledge from those data requires a new approach. This work outlines a general procedure to create an integrated data management solution with the aim of implementing a customizable schema and rules-based engine. The challenge is to validate, manage and consolidate the data with different sources and formats into a standardized single trusted version. Once the fundamental Data Engineering task is done a wealth of data analysis techniques were applied to establish the relationship between well completion parameters and long-term production of an unconventional asset. Different databases of an unconventional play in North America were used, facilitating the interconnection between tables due to the standardization in most column names, primary key parameters, etc. However, this process involved a dynamic workflow to load and inspect the quality and representativeness of the data each time new information was loaded, using business rules specifically created for each attribute. Once the standardized product was generated (SQL database), it was connected to the data analytics tool to evaluate the EUR and its relationship with completion parameters. In most Upstream data analytics projects, it is common to spend more time preparing the data than analyzing it. Therefore, the integrated solution recommended here was extremely useful to facilitate the analysis, adding a high level of confidence, consistency, and representativeness on the analysis performed. This approach can be used as a guide to improve the field development plan of a play and can also be extended to host training data for Artificial Intelligence and Machine Learning techniques.

# Amalgamation of Diverse Data Types and Sources to Facilitate Data Analytics of Engineering Interpretation and Historical Data

**Authors:** Cesar Velasquez, Ivan Olea, Russell Roundtree

# Introduction

In the age of “if it can be digitized, it will be digitized”, proliferation of data types and data volumes are outpacing our historical ability to respond. As our industry becomes ever more “data driven”, the need to efficiently use diverse data types from many sources and make these data available to a host data analytics tools to extract information and knowledge from those data requires a new approach.

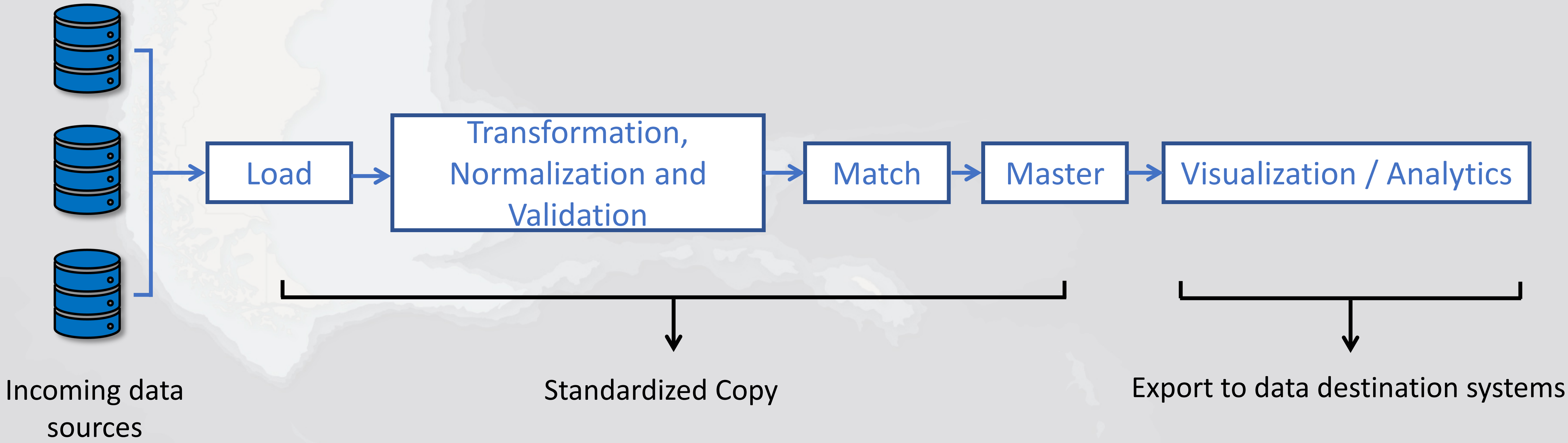
# Outline

- Objective
- Standard Data Management Workflow:
  - Pre-Match Processing.
  - Match Processing.
  - Post-Match Processing.
- Unconventional Reservoir - Visualization / Data Analysis
- Conclusions

# Objective

This work outlines a general procedure to create an integrated data management solution with the aim of implementing a customizable schema and rules-based engine. The challenge is to validate, manage and consolidate the data with different sources and formats into a standardized single trusted version. Once the fundamental Data Engineering task is done a wealth of data analysis techniques were applied to establish the relationship between well completion parameters and long-term production of an unconventional asset.

# Standard Data Management Workflow

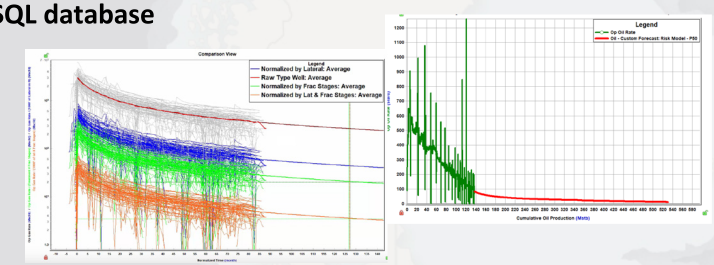


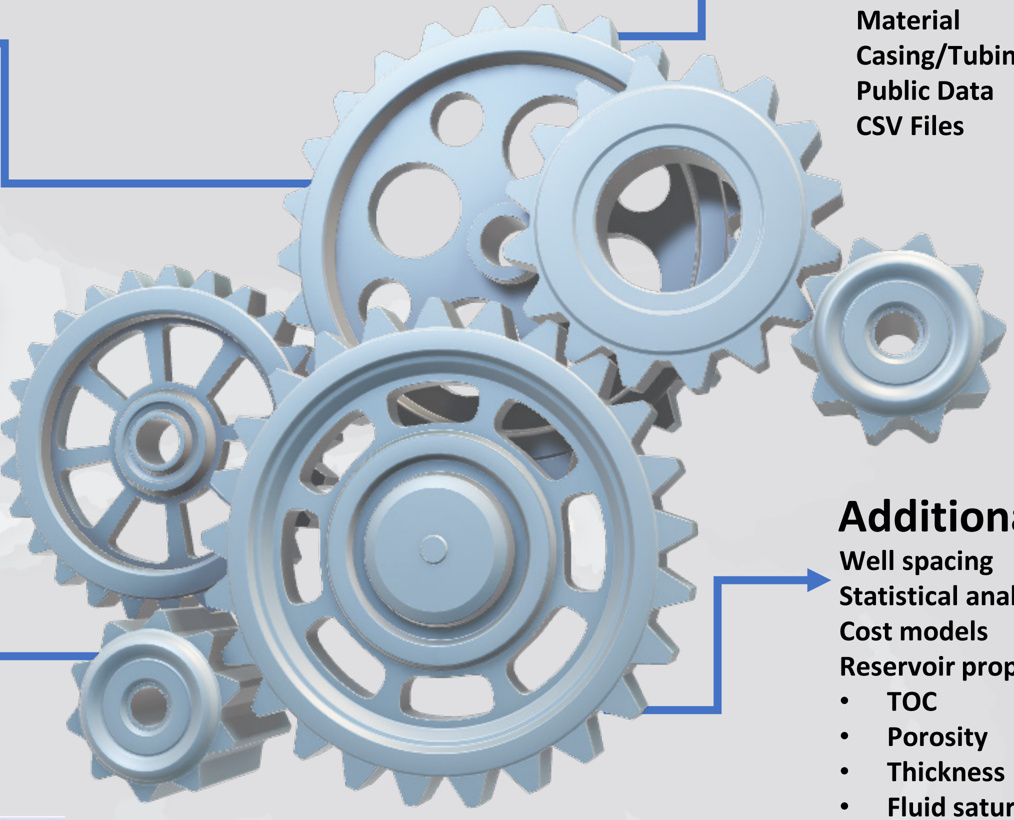
# Sources to be interconnected

IHS Markit / Harmony Enterprise

## Production Database

Production history
Cumulative production (3, 6, 12, 18, 36 months)
DCA - Decline Curve Analysis
EUR – Estimated Ultimate Recovery
SQL database

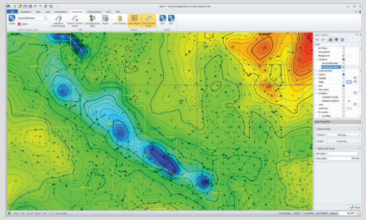


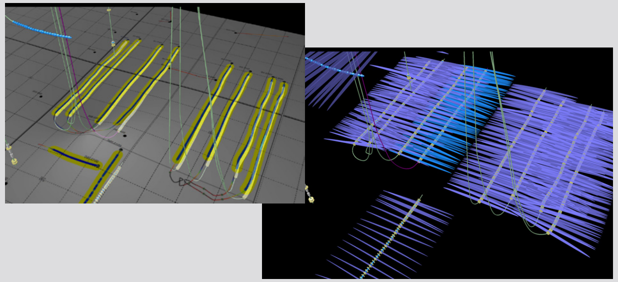


IHS Markit / AccuMap

## Completion Database

Historical well Completions
Perforation
Material
Casing/Tubing
Public Data
CSV Files

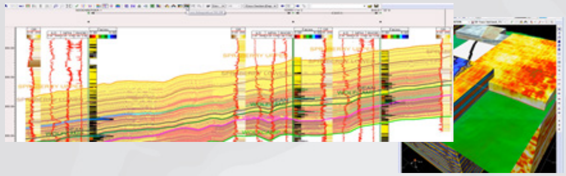




IHS Markit / Kingdom

## Geological Model

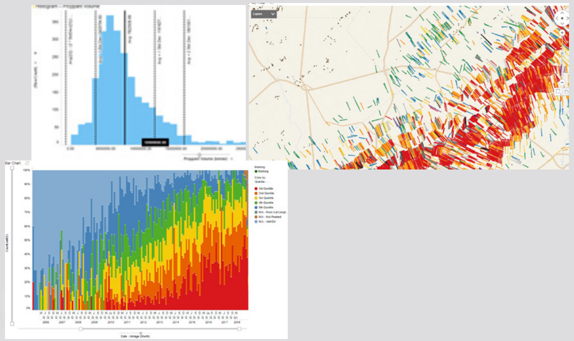
LogData
Surface / Zone selected
Structural model
Petrophysical model
SQL Database



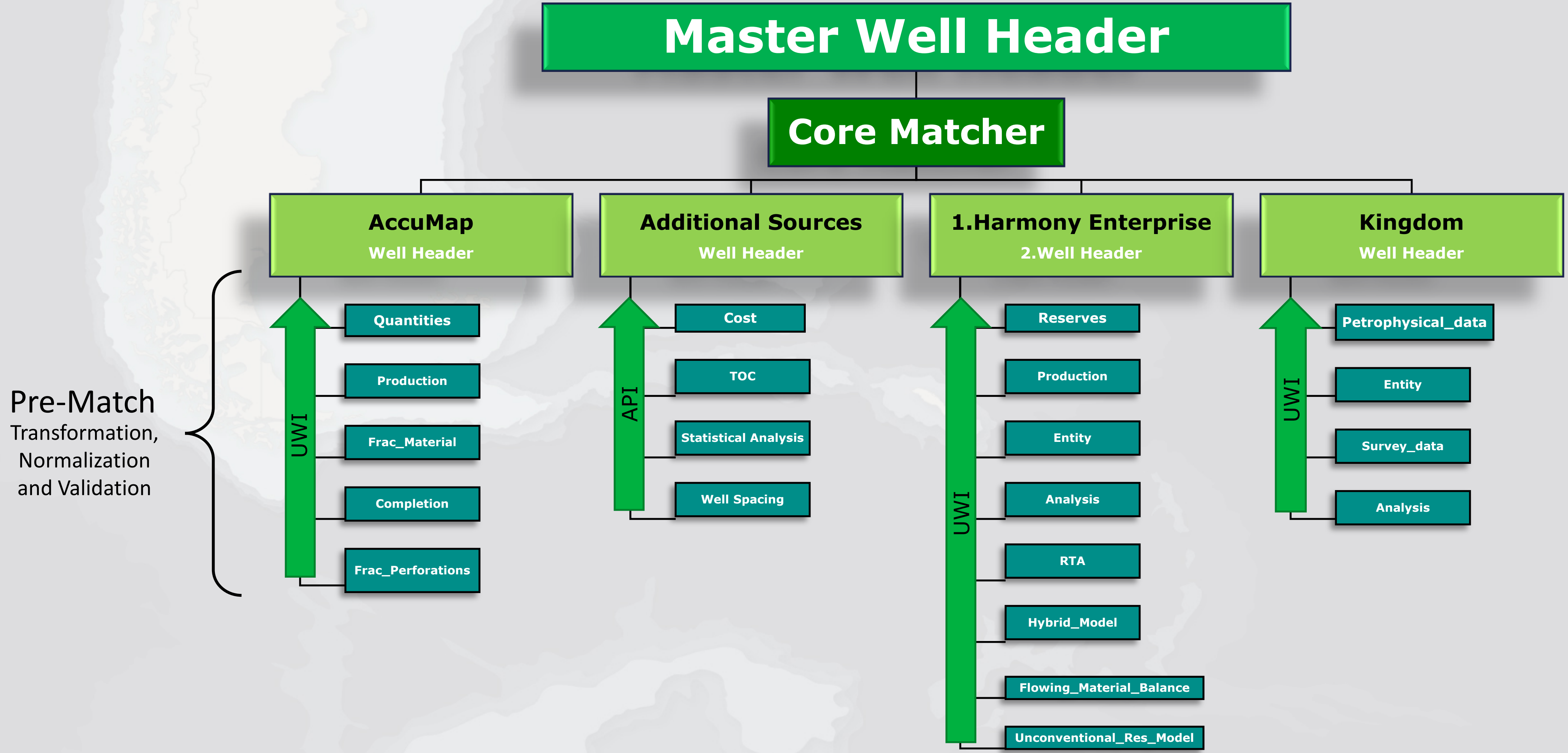
## Additional Sources

Well spacing
Statistical analysis
Cost models
Reservoir properties:
- TOC
- Porosity
- Thickness
- Fluid saturation

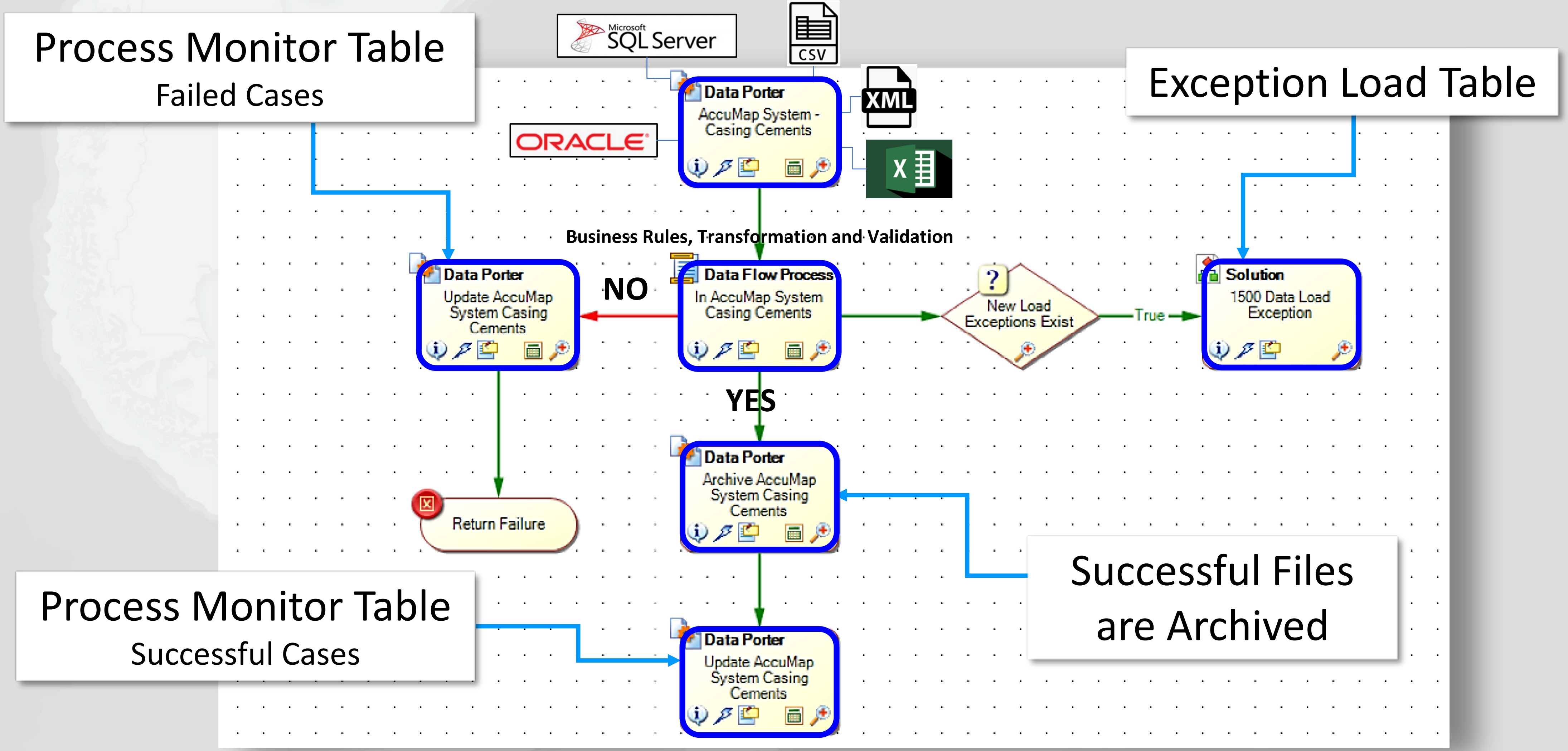
CSV Files



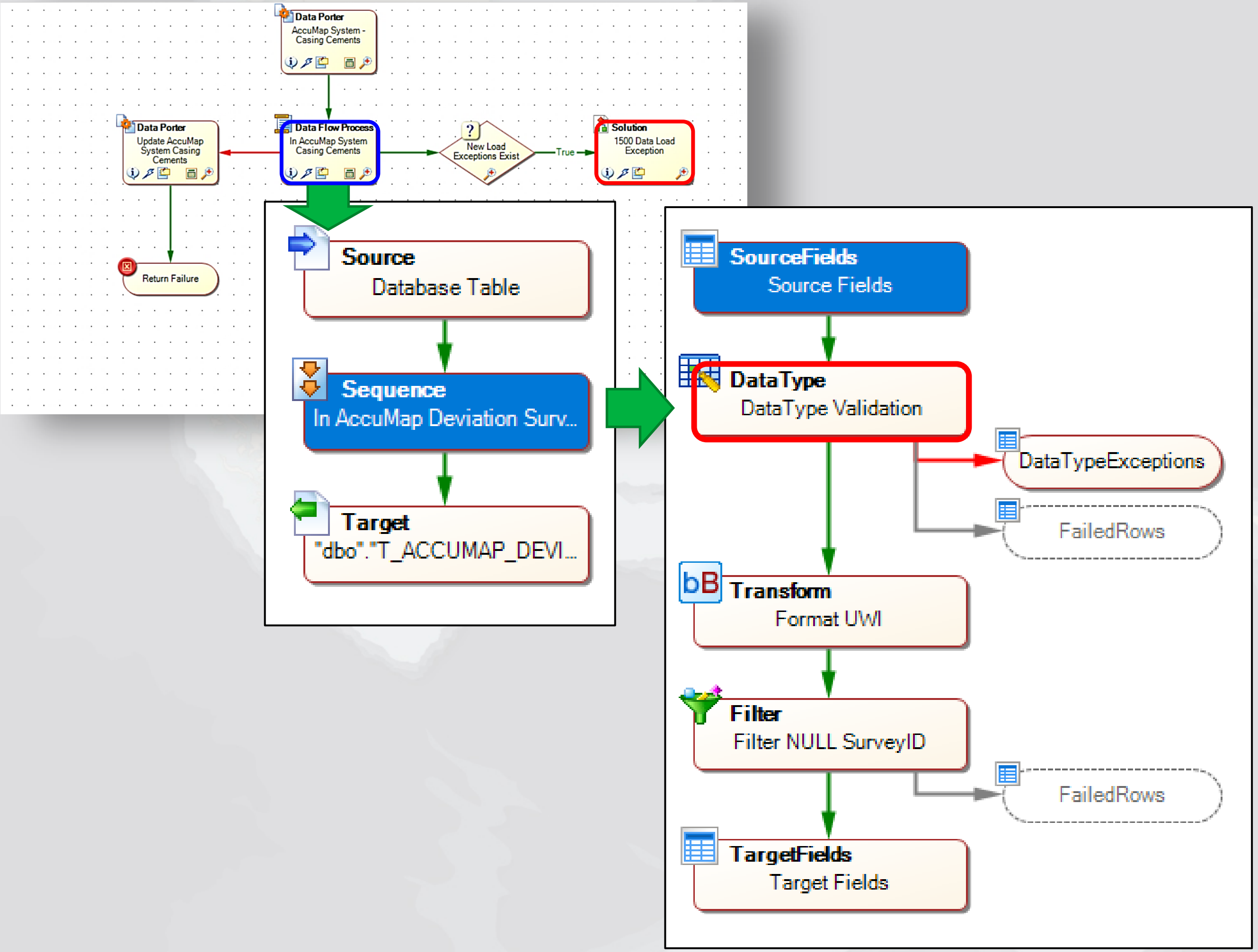
# Technical Workflow / Schematic



# Pre-Match Process

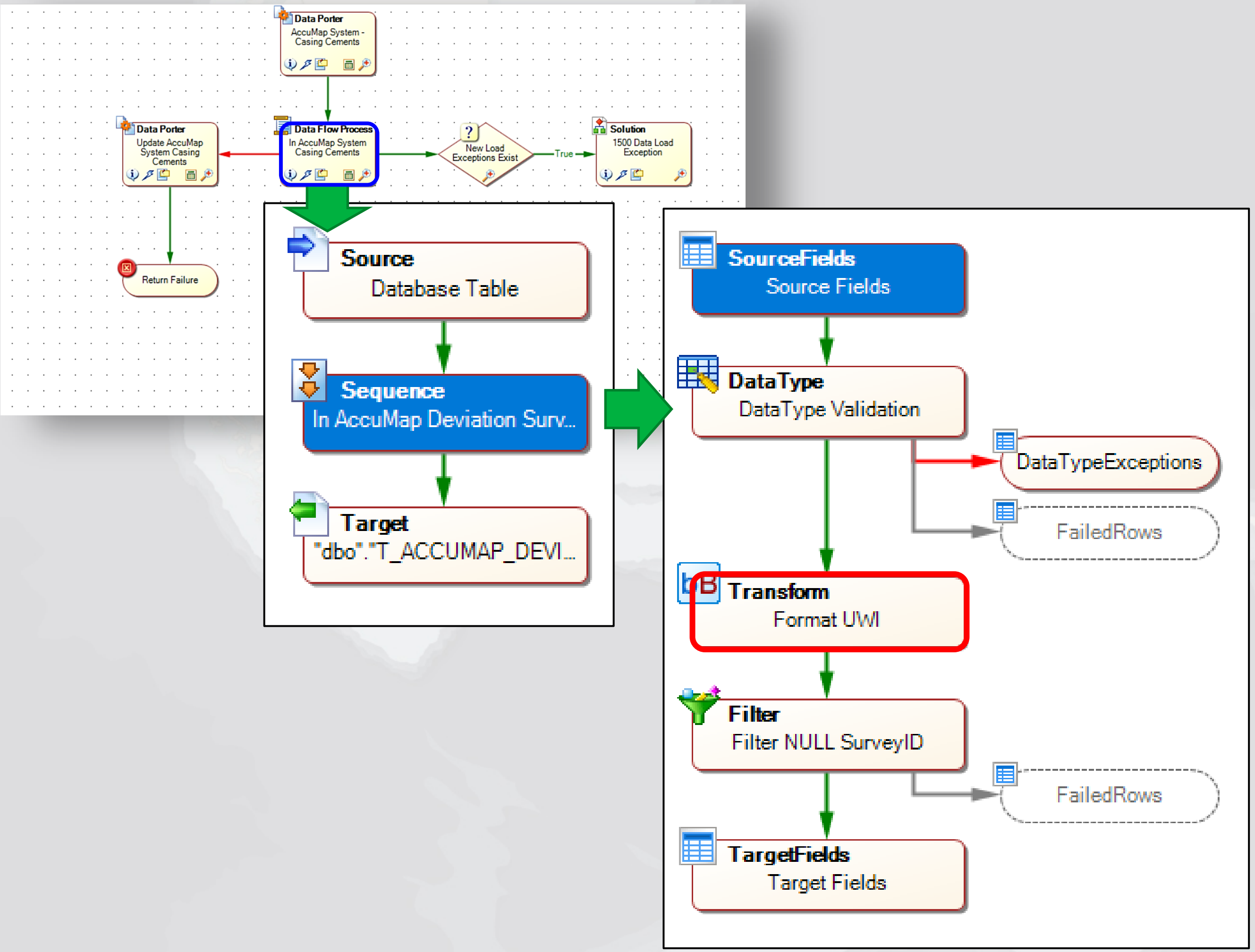


# Pre-Match Process




| Attribute Name | Data Type | Exception Code |
|---|---|---|
| **UWI** | NVARCHAR | Field Length Exceeded ✓ |
| **TUBING_ID** | INT | Invalid Integer ✓ |
| **RIG_RELEASE_DATE** | DATETIME | Invalid Date ✗ |
| **WATER_SATURATION** | NUMBER | Invalid Number ✓ |

# Pre-Match Process



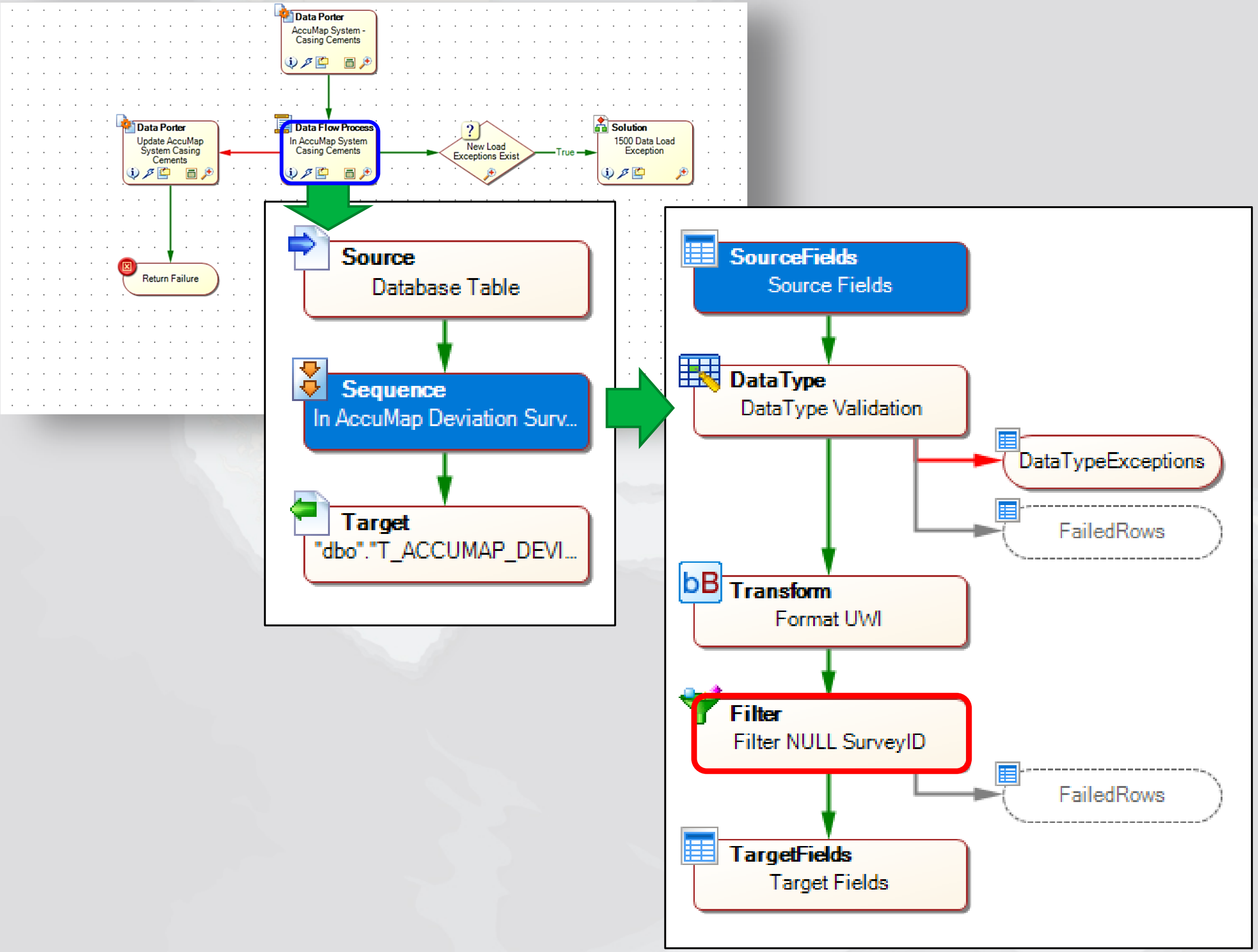

| UNFORMATED UWI |
|---|
| 1W50601801010000 |
| 1W50632001010000 |
| 1W50622404010000 |

**Business Rule** SQL Server

Left({INPUT}.[Input UWI],3 )+'/'+substring({INPUT}.[Input UWI],4,2 )+'-'+substring({INPUT}.[Input UWI],6,2)+'-'+substring({INPUT}.[Input UWI],8,3)+'-'+substring({INPUT}.[Input UWI],11,4)+'/'+substring({INPUT}.[Input UWI],15,3)

| UNFORMATED UWI | FORMATED UWI |
|---|---|
| 1W50601801010000 | **100/01-01-060-18W5/0** |
| 1W50632001010000 | **100/01-01-063-20W5/0** |
| 1W50622404010000 | **100/01-04-062-24W5/0** |

# Pre-Match Process



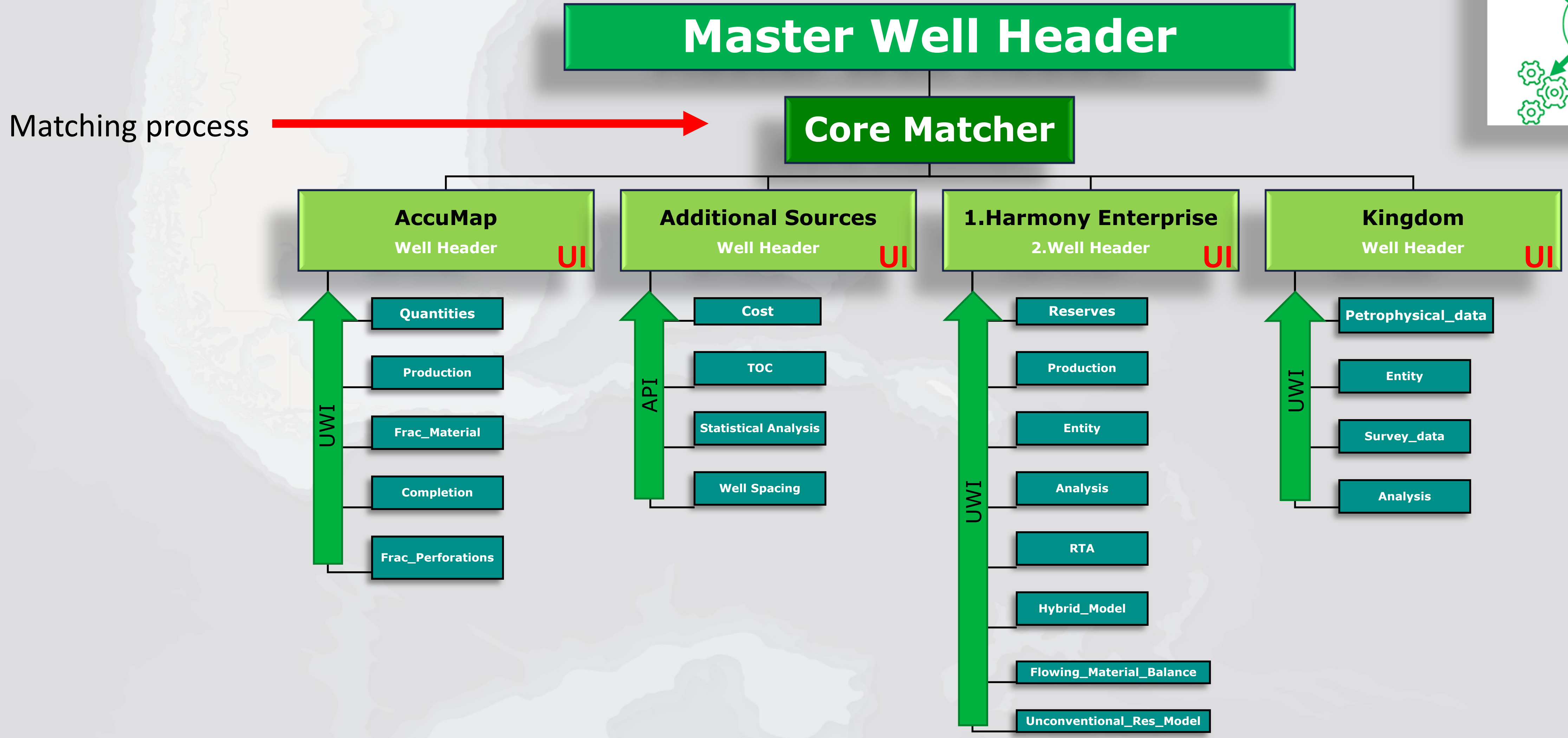

| PRE_FRAC_COST | PRE_FRAC_COST_PER_STAGE | FRAC_COST |
|---|---|---|
| 2176955.36 | 140540.41 | 6941524.36 |
| 1285860 | 45840.63 | 3987525 |
| NULL | NULL | NULL |
| 731660 | 75933.54 | 3438640 |
| NULL | NULL | NULL |
| NULL | NULL | NULL |
| 8306749 | 56856.78 | 2059599 |

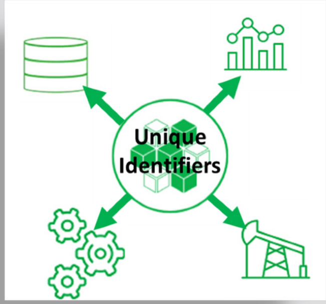
**Business Rule** Microsoft SQL Server

Filter = IF {INPUT}.[INPUT Value] IS NULL THEN 0
Else = 1

| PRE_FRAC_COST | PRE_FRAC_COST_PER_STAGE | FRAC_COST |
|---|---|---|
| 2176955.36 | 140540.41 | 6941524.36 |
| 1285860 | 45840.63 | 3987525 |
| 731660 | 75933.54 | 3438640 |
| 8306749 | 56856.78 | 2059599 |

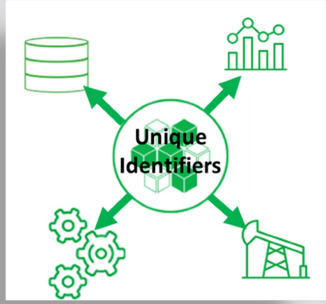
# Technical Workflow / Schematic



AAPG ICE 2019 International Conference & Exhibition
AAPG 27-30 August Buenos Aires, Argentina

# Matching Process




Harmony Well Header

| WELL_ID | PRIMARY_FLUID | GIS_STATUS_ID | DLS | FIELD_NAME |
|---|---|---|---|---|
| 0032484A-B09E-439F-A4CF-04330EAAC9E9 | Gas | 1 | 100/10-23-064-21W5/0 | TONY CREEK NORTH |
| 005EC69D-AAD4-45FF-B434-1E79FBDB1 | | | 3W5/0 | KAYBOB SOUTH |
| 0061B2AD-52DC-4B30-89FA-8CFBAE74 | | | 5W5/0 | TWO CREEK |
| 00703C37-A53D-42E9-A8B6-B63C7BB7F | | | 0W5/0 | KAYBOB |

AccuMap Well Header

| UWI | LAH | FLUID | MODE |
|---|---|---|---|
| 100/01-01-060-18W5/0 | Deeper- | Shale Gas | Flowing |
| 100/01-01-063-20W5/0 | Develo | Shale Gas | Pumping |
| 100/01-04-062-24W5/0 | Develo | Shale Gas | Pumping |
| 100/01-06-044-05W5/2 | Deeper- | Shale Gas | Flowing |

Kingdom Well Header

| ID | NAME | LATITUDE |
|---|---|---|
| 1000 | 100/16-33-062-20W | 54.3978005 |
| 1001 | 102/16-33-062-20W | 54.3977105 |
| 1002 | 100/09-33-062-20W | 54.4068816 |
| 1003 | 100/15-33-062-20W | 54.3976215 |

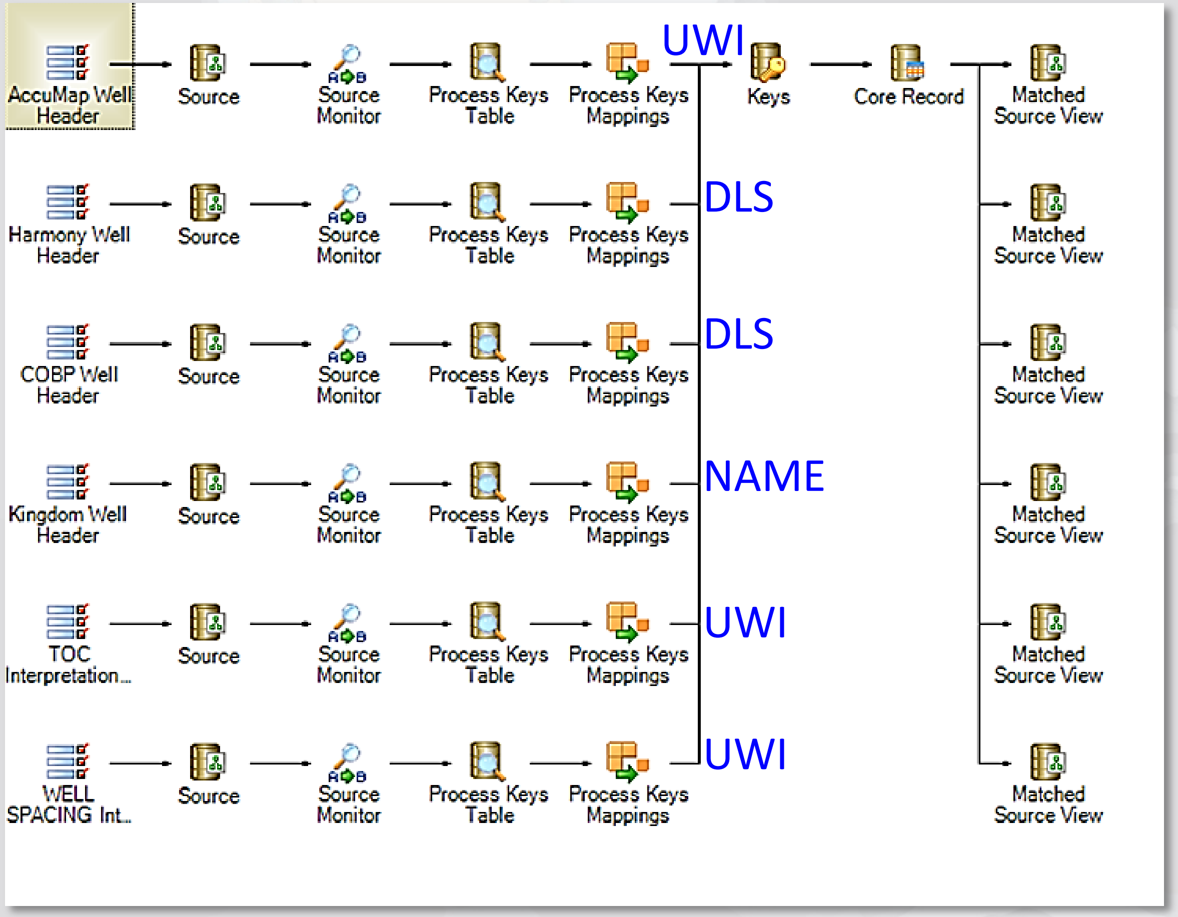
| UWI | | | |
|---|---|---|---|
| 100/01-01-063-20W5/0 | | | |
| 100/01-04-062-24W5/0 | | /0 | 488 |
| 100/01-06-062-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 100/02-06-062-20W5/0 | 402 |
| 100/01-06-062-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 102/04-05-062-20W5/0 | 418 |

| Matching Attributes |
|---|
| Wellbore ID |
| License Number |
| Lahee Name |
| Cost Center |
| Well ID |
| Well Name |

## Data Illustrator

| Field Name | Illustrated Name |
|---|---|
| API | UWI |
| DLS | UWI |
| NAME | UWI |

# Matching Process




Harmony Well Header

| PRIMARY_FLUID | GIS_STATUS_ID | DLS |
|---|---|---|
| Gas | 4 | 100/10-23-064-21W5/0 |
| Oil | 1 | 102/16-33-062-20W5/0 |
| NULL | 0 | 100/02-17-064-15W5/0 |
| Oil | 4 | 100/10-33-063-20W5/0 |

AccuMap Well Header

| UWI | LAHEE | OPERATOR |
|---|---|---|
| 100/10-23-064-21W5/0 | Deeper-pool Test | HUSKY OIL OPERATIONS LIMITED |
| 100/01-01-063-20W5/0 | Development | SHELL CANADA ENERGY |
| 100/01-04-062-24W5/0 | Development | ENCANA CORPORATION |
| 100/01-06-044-05W5/2 | Deeper-pool Test | ENCANA CORPORATION |

Kingdom Well Header

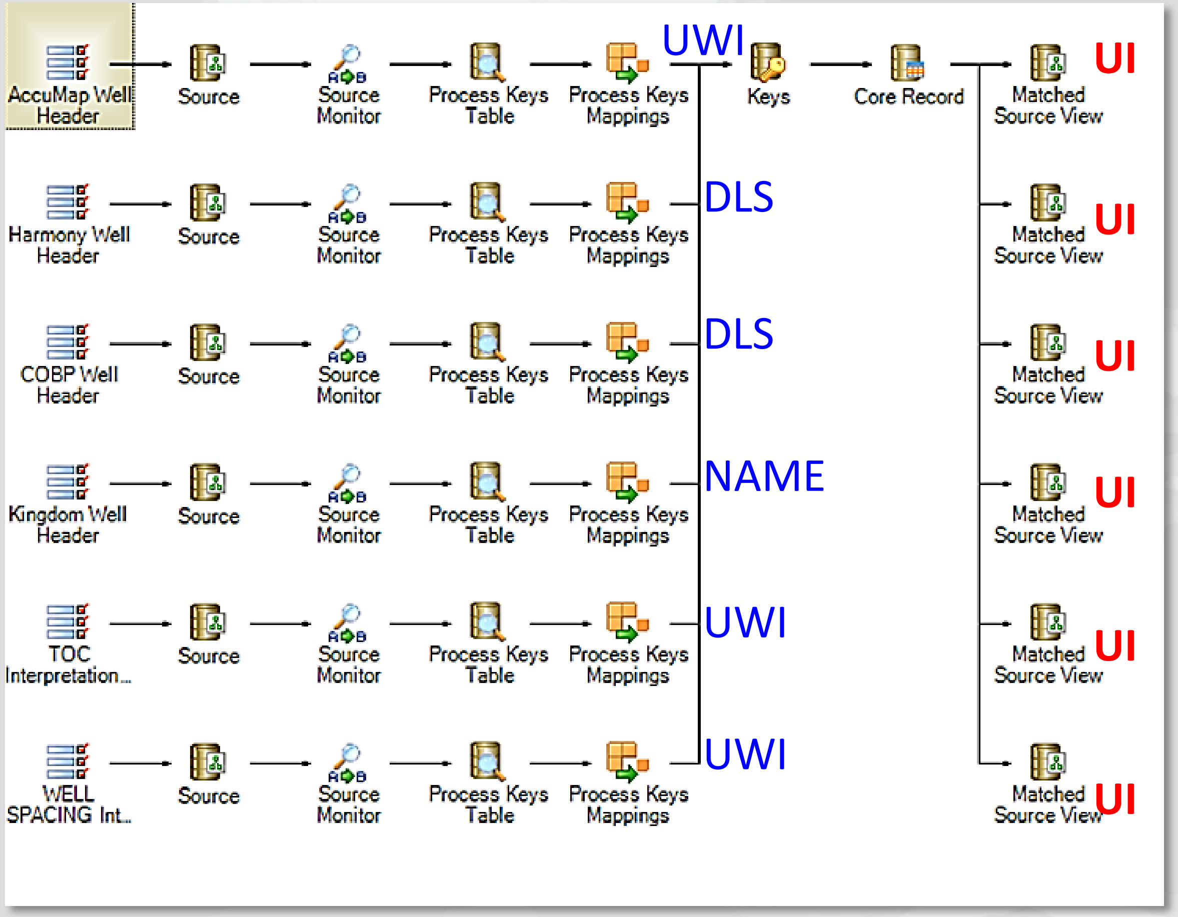
| NAME | WELLNUMBER | SURFACELOCX | SURFACELOCY |
|---|---|---|---|
| 100/16-33-062-20W5/0 | 100/16-33 | 504218.34 | 6027568.54 |
| 102/16-33-062-20W5/0 | 102/16-33 | 504218.74 | 6027558.52 |
| 100/09-33-062-20W5/0 | 100/09-33 | 504450.6 | 6028579.2 |
| 100/15-33-062-20W5/0 | 100/15-33 | 504219.07 | 6027548.62 |

Well Spacing Header

| UWI | WELL_NAME | DISTANCE_M |
|---|---|---|
| 100/10-33-063-20W5/0 | SCL HZ FC23I KAYBOB 1-1-63-20 | 532 |
| 100/01-04-062-24W5/0 | ECA HZ WAHIGAN 1-4-62-24 | 488 |
| 100/01-01-063-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 402 |
| 100/01-06-062-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 418 |

AAPG ICE 2019 AAPG 27-30 August Buenos Aires, Argentina
International Conference & Exhibition

# Matching Process




Harmony Well Header

| UI | PRIMARY_FLUID | GIS_STATUS_ID | DLS |
|---|---|---|---|
| 1000 | Gas | 4 | 100/10-23-064-21W5/0 |
| 1001 | Oil | 1 | 102/16-33-062-20W5/0 |
| 1002 | NULL | 0 | 100/02-17-064-15W5/0 |
| 1003 | Oil | 4 | 100/10-33-063-20W5/0 |

AccuMap Well Header

| UI | UWI | LAHEE | OPERATOR |
|---|---|---|---|
| 1000 | 100/10-23-064-21W5/0 | Deeper-pool Test | HUSKY OIL OPERATIONS LIMITED |
| 1005 | 100/01-01-063-20W5/0 | Development | SHELL CANADA ENERGY |
| 1006 | 100/01-04-062-24W5/0 | Development | ENCANA CORPORATION |
| 1007 | 100/01-06-044-05W5/2 | Deeper-pool Test | ENCANA CORPORATION |

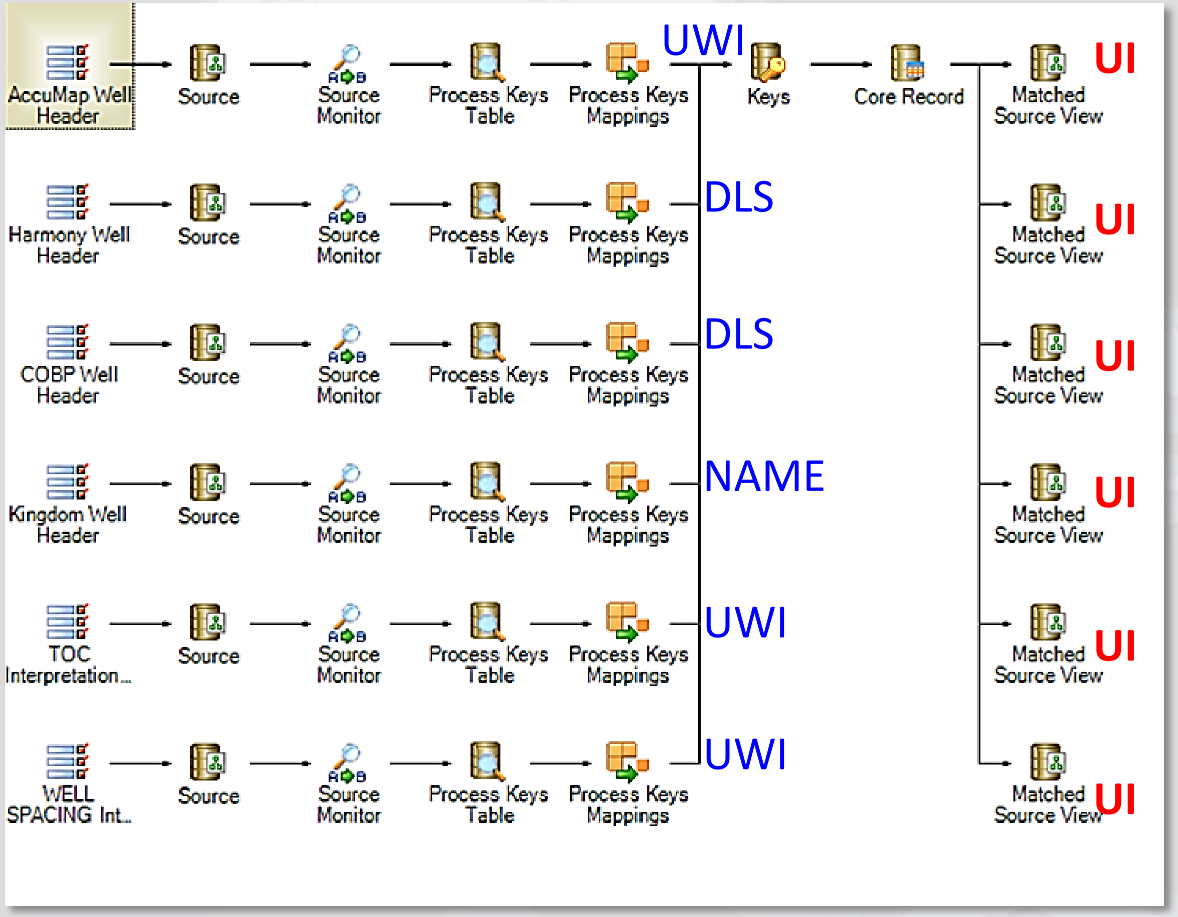
Kingdom Well Header

| UI | NAME | WELLNUMBER | SURFACELOCX | SURFACELOCY |
|---|---|---|---|---|
| 1008 | 100/16-33-062-20W5/0 | 100/16-33 | 504218.34 | 6027568.54 |
| 1001 | 102/16-33-062-20W5/0 | 102/16-33 | 504218.74 | 6027558.52 |
| 1010 | 100/09-33-062-20W5/0 | 100/09-33 | 504450.6 | 6028579.2 |
| 1011 | 100/15-33-062-20W5/0 | 100/15-33 | 504219.07 | 6027548.62 |

Well Spacing Header

| UI | UWI | WELL_NAME | DISTANCE_M |
|---|---|---|---|
| 1003 | 100/10-33-063-20W5/0 | SCL HZ FC23I KAYBOB 1-1-63-20 | 532 |
| 1009 | 100/01-04-062-24W5/0 | ECA HZ WAHIGAN 1-4-62-24 | 488 |
| 1005 | 100/01-01-063-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 402 |
| 1011 | 100/01-06-062-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 418 |

# Matching Process




Harmony Well Header

| UI | PRIMARY_FLUID | GIS_STATUS_ID | DLS |
|---|---|---|---|
| 1000 | Gas | 4 | 100/10-23-064-21W5/0 |
| 1001 | Oil | 1 | 102/16-33-062-20W5/0 |
| 1002 | NULL | 0 | 100/02-17-064-15W5/0 |
| 1003 | Oil | 4 | 100/10-33-063-20W5/0 |

AccuMap Well Header

| UI | UWI | LAHEE | OPERATOR |
|---|---|---|---|
| 1000 | 100/10-23-064-21W5/0 | Deeper-pool Test | HUSKY OIL OPERATIONS LIMITED |
| 1005 | 100/01-01-063-20W5/0 | Development | SHELL CANADA ENERGY |
| 1006 | 100/01-04-062-24W5/0 | Development | ENCANA CORPORATION |
| 1007 | 100/01-06-044-05W5/2 | Deeper-pool Test | ENCANA CORPORATION |

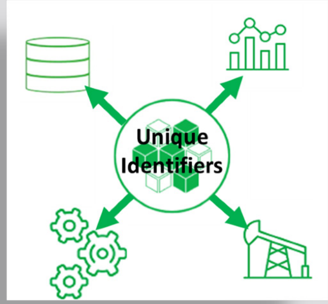
Kingdom Well Header

| UI | NAME | WELLNUMBER | SURFACELOCX | SURFACELOCY |
|---|---|---|---|---|
| 1008 | 100/16-33-062-20W5/0 | 100/16-33 | 504218.34 | 6027568.54 |
| 1001 | 102/16-33-062-20W5/0 | 102/16-33 | 504218.74 | 6027558.52 |
| 1010 | 100/09-33-062-20W5/0 | 100/09-33 | 504450.6 | 6028579.2 |
| 1011 | 100/15-33-062-20W5/0 | 100/15-33 | 504219.07 | 6027548.62 |

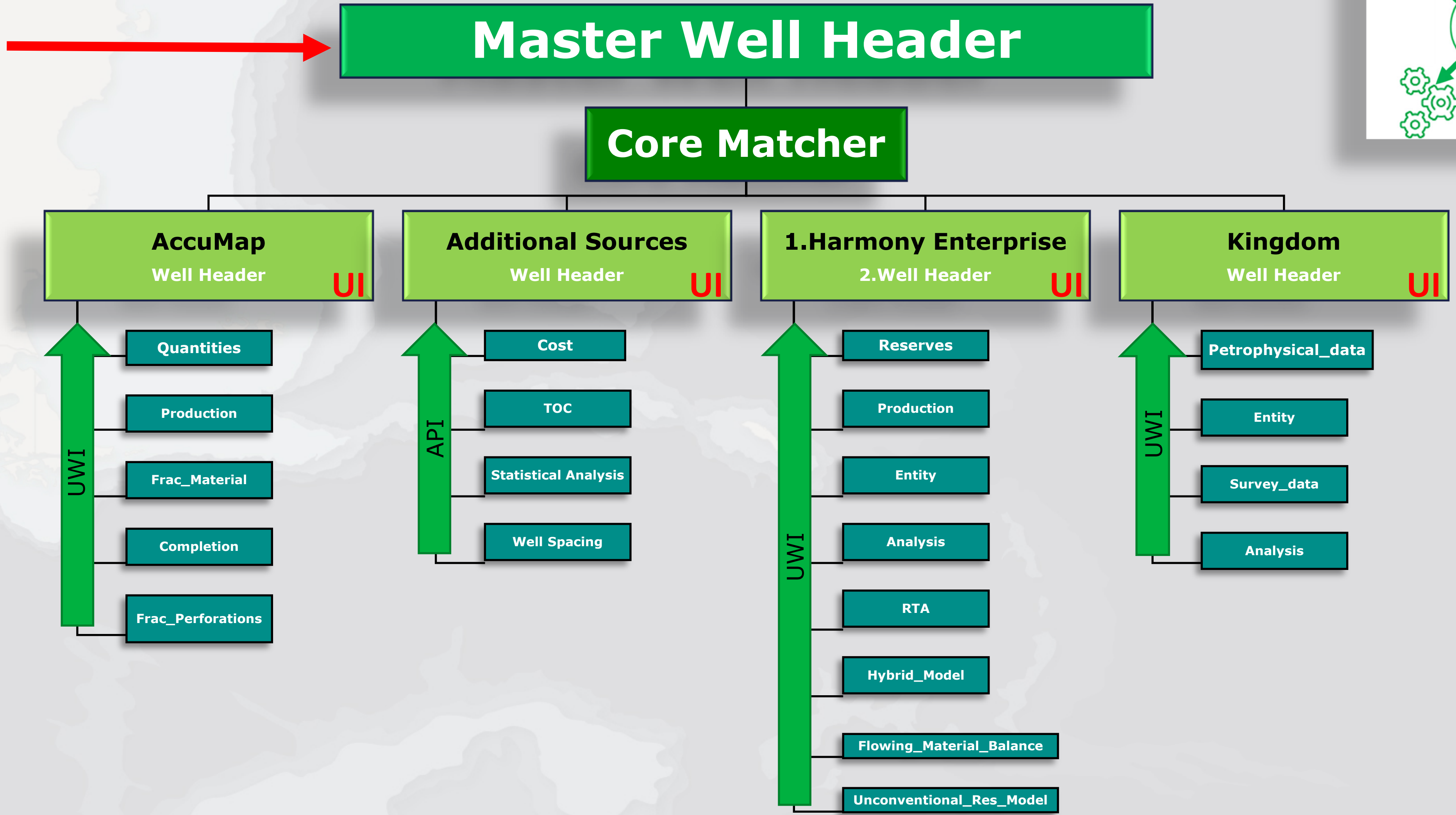
Well Spacing Header

| UI | UWI | WELL_NAME | DISTANCE_M |
|---|---|---|---|
| 1003 | 100/10-33-063-20W5/0 | SCL HZ FC23I KAYBOB 1-1-63-20 | 532 |
| 1009 | 100/01-04-062-24W5/0 | ECA HZ WAHIGAN 1-4-62-24 | 488 |
| 1005 | 100/01-01-063-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 402 |
| 1011 | 100/01-06-062-20W5/0 | CHEVRON HZ KAYBOBS 1-6-62-20 | 418 |

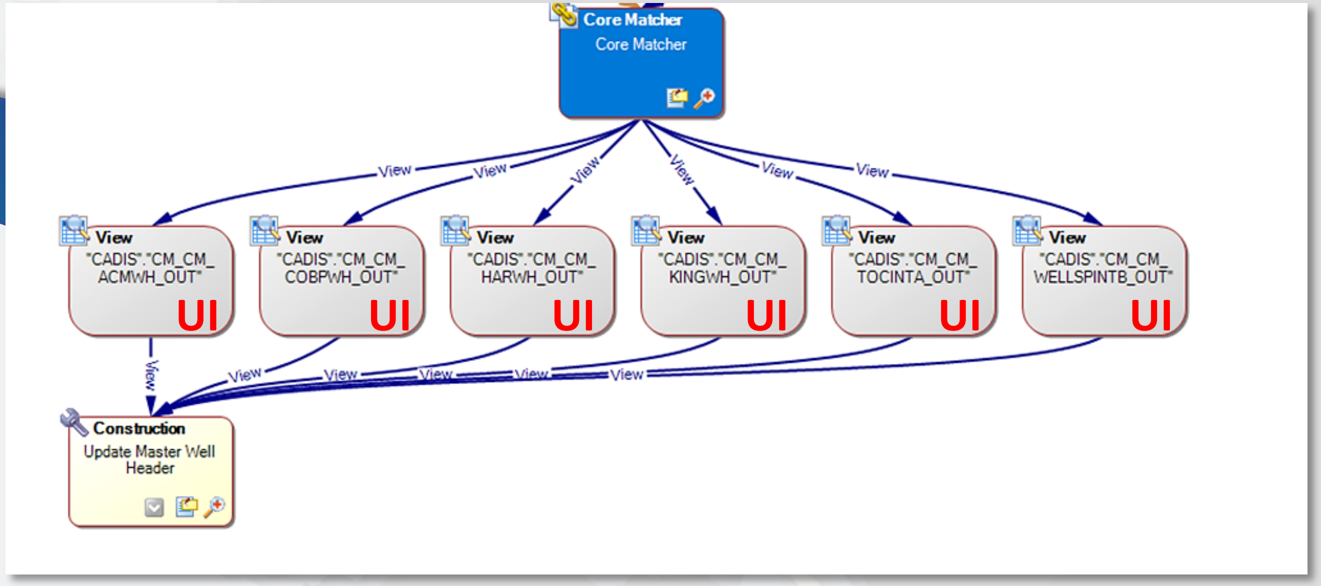
# Technical Workflow / Schematic

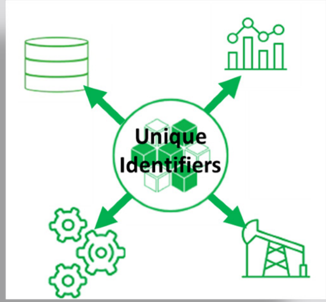



Mastering process



# Mastering Process

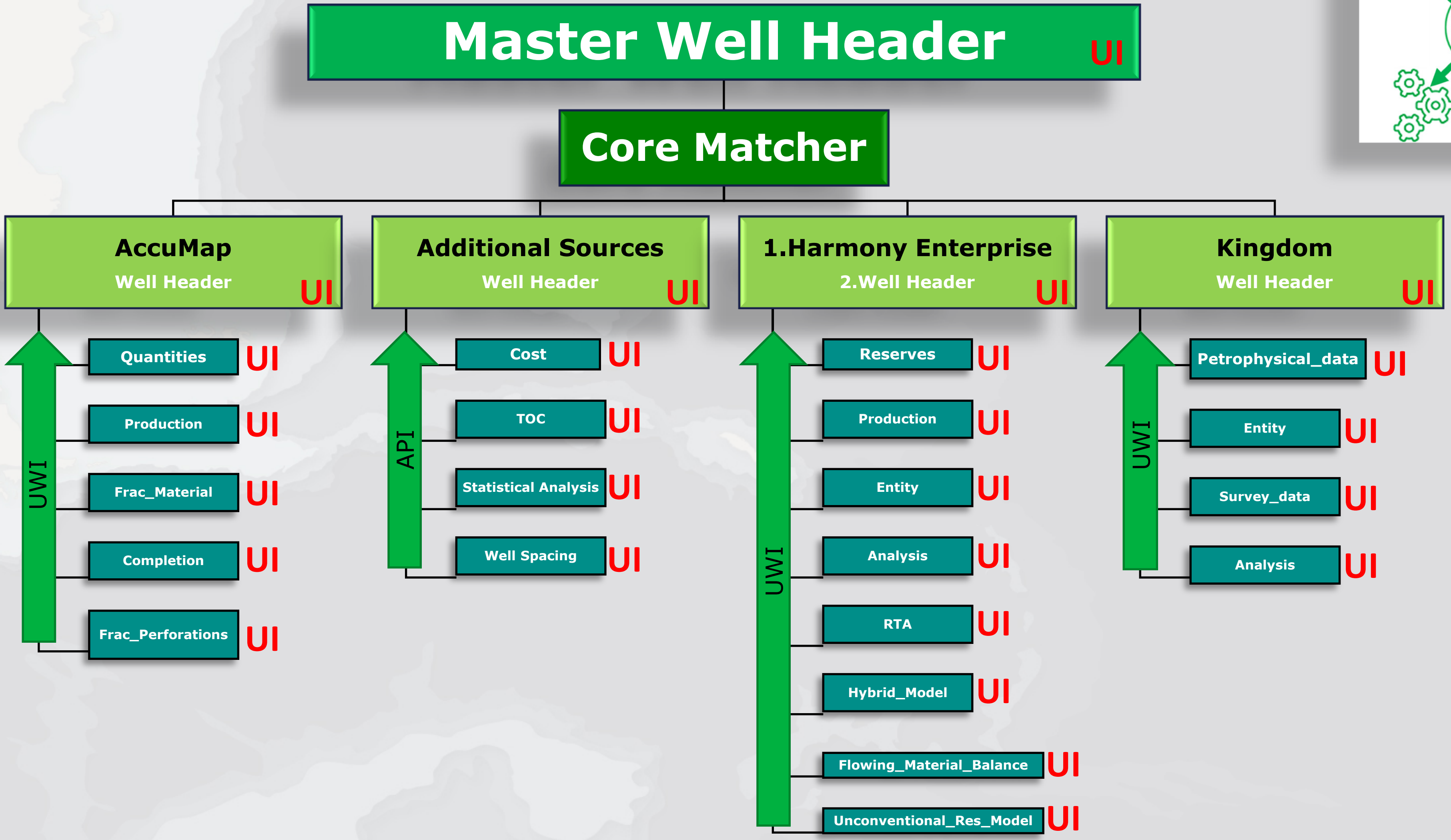


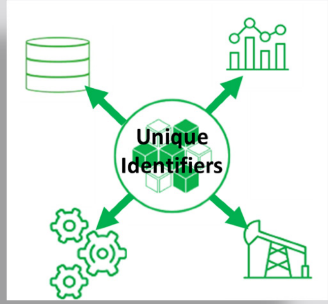




## WELL HEADER - MASTER TABLE

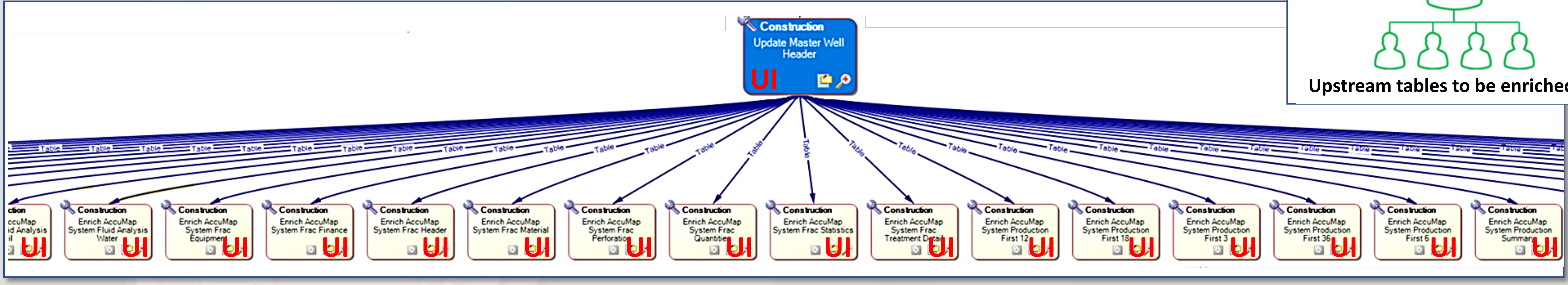
| Harmony Well Header | | Kingdom Well Header | | AccuMap Well Header | | Well Spacing Header | |
|---|---|---|---|---|---|---|---|
| UI | DLS | UI | NAME | UI | UWI | UI | UWI |
| 1000 | 100/10-23-064-21W5/0 | | | 1000 | 100/10-23-064-21W5/0 | | |
| 1001 | 102/16-33-062-20W5/0 | 1001 | 102/16-33-062-20W5/0 | | | | |
| 1002 | 100/02-17-064-15W5/0 | | | | | | |
| 1003 | 100/10-33-063-20W5/0 | 1003 | 100/10-33-063-20W5/0 | | | 1003 | 100/10-33-063-20W5/0 |
| 1004 | 100/01-01-060-18W5/00 | 1004 | 100/01-01-060-18W5/00 | 1004 | 100/01-01-060-18W5/00 | 1004 | 100/01-01-060-18W5/00 |
| 1005 | 100/01-01-063-20W5/0 | | | 1005 | 100/01-01-063-20W5/0 | 1005 | 100/01-01-063-20W5/0 |
| 1006 | 100/01-06-044-05W5/2 | 1006 | 100/01-06-044-05W5/2 | 1006 | 100/01-06-044-05W5/2 | 1006 | 100/01-06-044-05W5/2 |

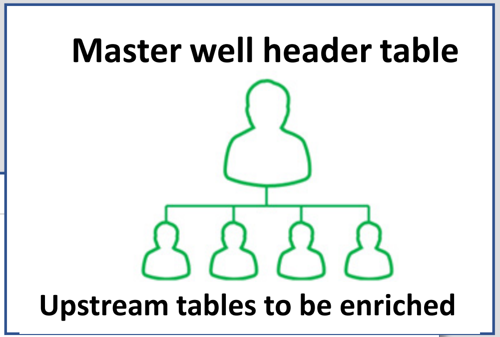
# Technical Workflow / Schematic



# Mastering Process



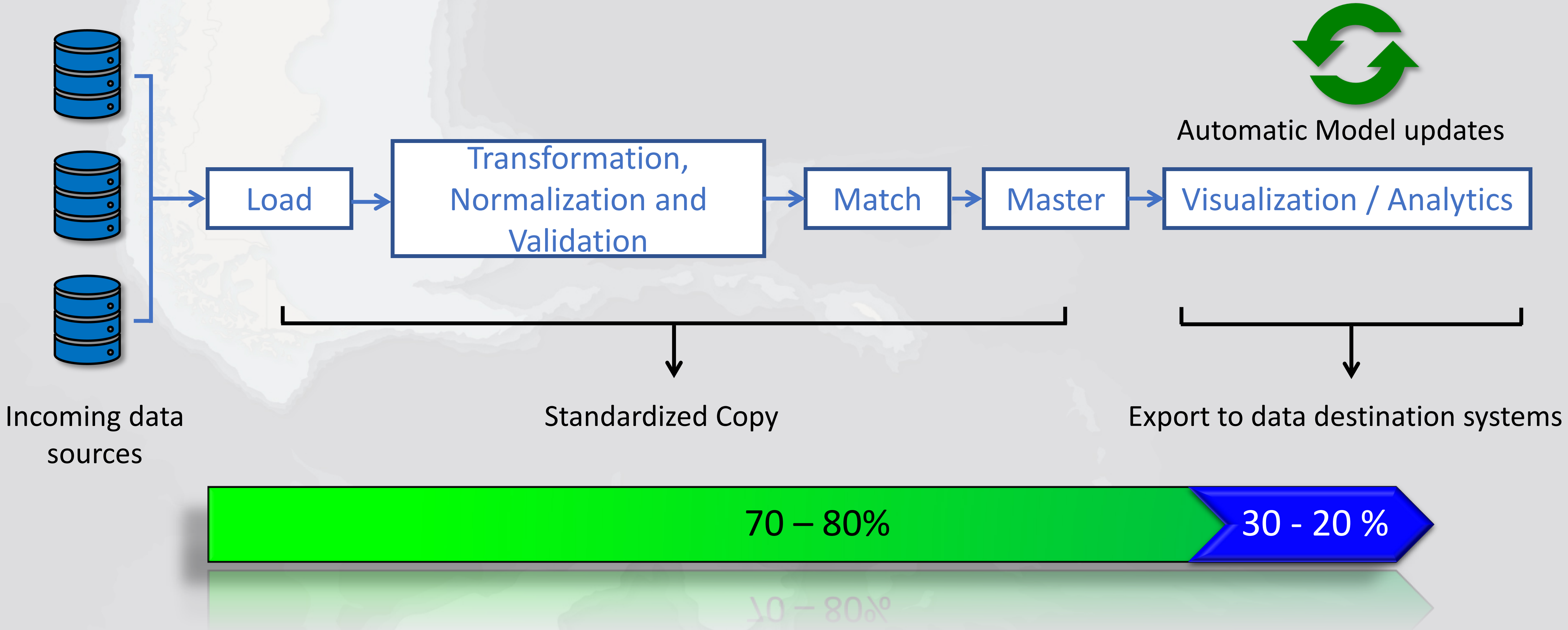

Harmony UNCONVENTIONAL_RES_MODEL

| UI | LINEAR_FLOW_PARAMETER | FRACTURE_SPACING | WIDTH_SIMULATED_RES_VOL | OGIP | RECOVERY_FACTOR | TIME_FOR_END_OF_LINEAR_FLOW |
|---|---|---|---|---|---|---|
| 5061 | 20016.8279 | NULL | NULL | 92613.5817 | 0.87553121 | 27.8464048 |
| 5334 | 31324.3223 | 177.880109 | 92.6923505 | 105801.222 | 0.78015649 | 16.8176132 |
| 5586 | 7700.14759 | 214.840332 | 55.0544154 | 490.233486 | 1.50994431 | 78.7509476 |
| 5967 | 29071.0677 | 130.030621 | 78.8230624 | 187629.457 | 0.80038728 | 45.6221895 |

AccuMap PPRODUCTION_FIRST 12_MONTHS

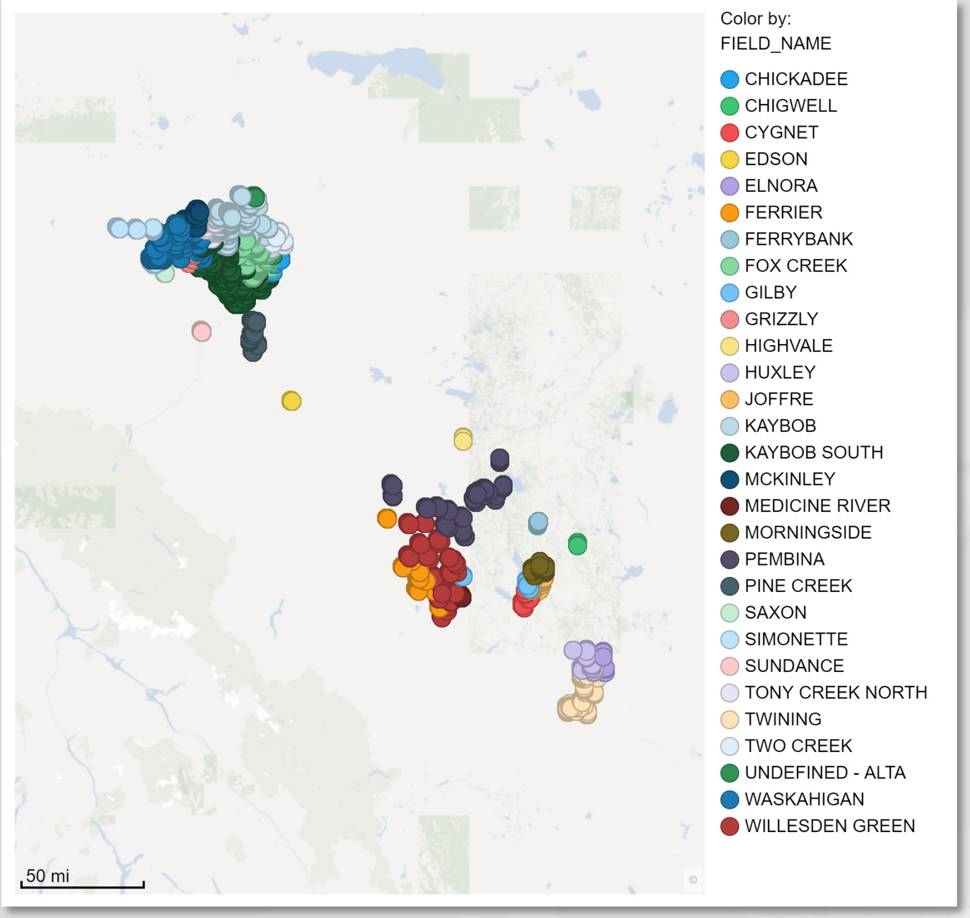
| UI | CUM_PRD_GAS_FIRST12_PROD_E12M12 | AVG_DLY_WATER_FIRST12_PROD_M12D | CAL_DLY_WATER_FIRST12_PROD_M12D | CUM_PRD_WATER_FIRST12_PROD_M12 | AVG_DLY_COND_FIRST12_PROD_M12D |
|---|---|---|---|---|---|
| 5000 | 1230.4 | 64.4 | 53.14 | 3241.4 | 6.64 |
| 5001 | 16184.4 | 28.85 | 25.52 | 8523.8 | 0.42 |
| 5002 | 44901.9 | 0 | NULL | 0 | 0 |
| 5003 | 7506.2 | 0.19 | 1.69 | 50.6 | 22.99 |

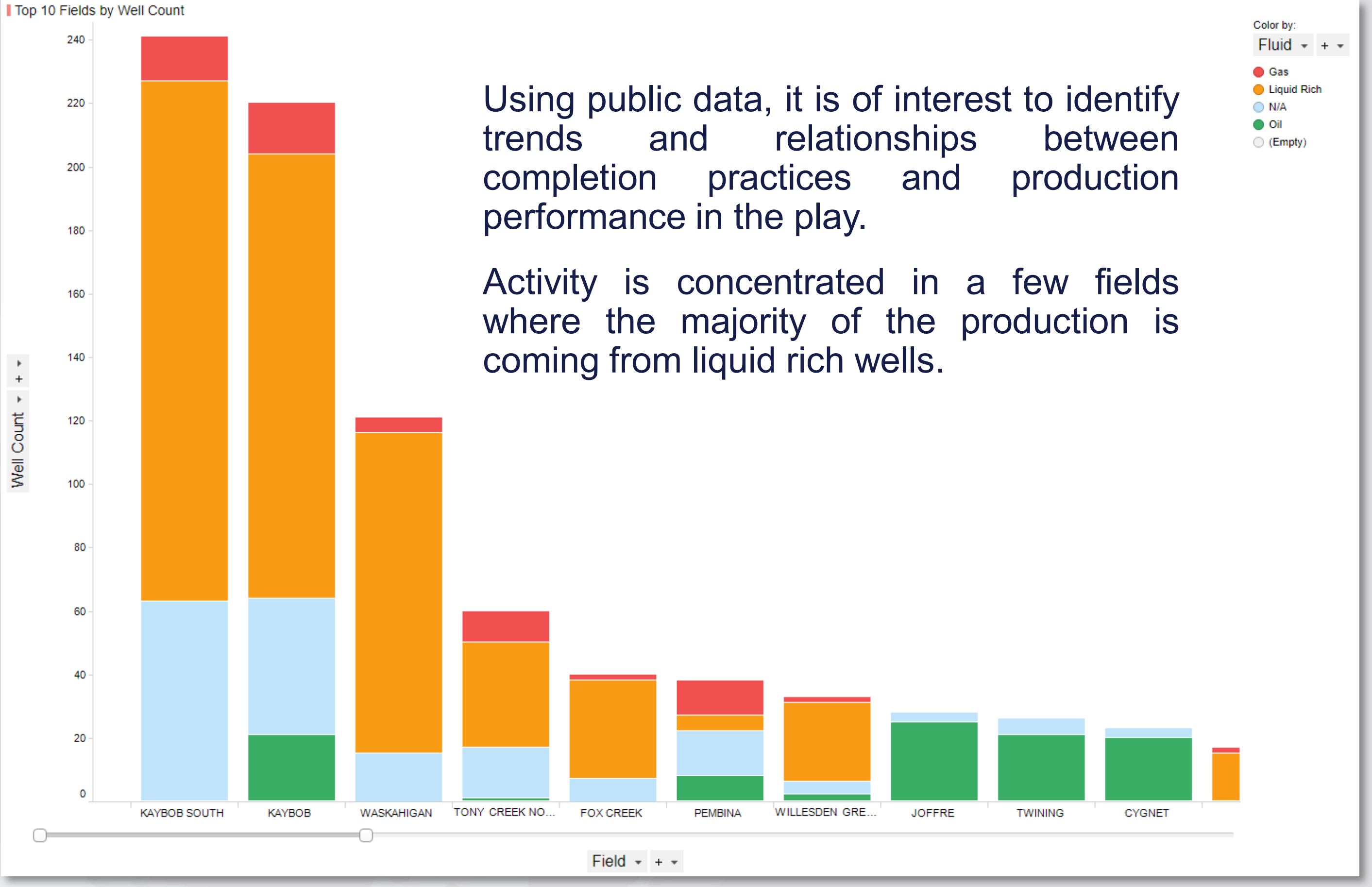
# Standard Data Management Workflow



# VISUALIZATION – DATA ANALYSIS

# Unconventional Reservoir - Visualization / Data Analysis



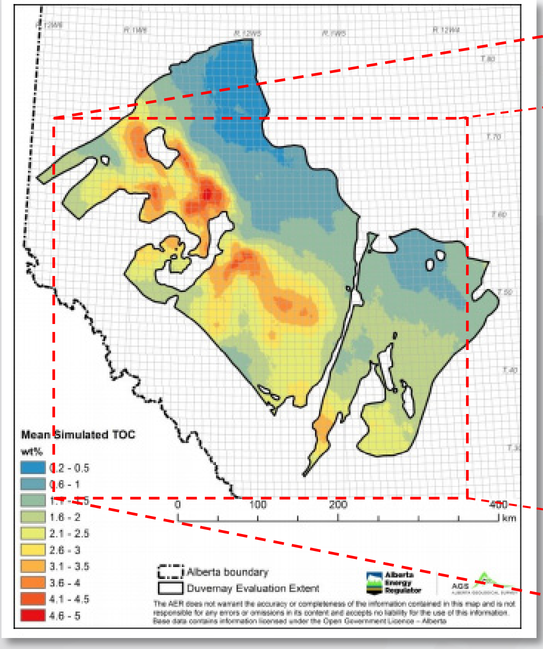

Using public data, it is of interest to identify trends and relationships between completion practices and production performance in the play.

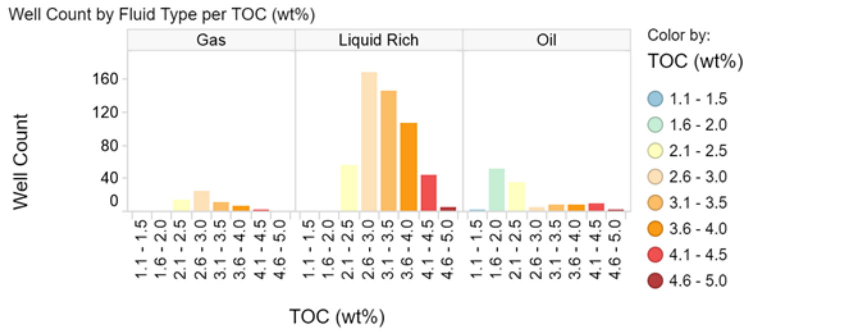
Activity is concentrated in a few fields where the majority of the production is coming from liquid rich wells.

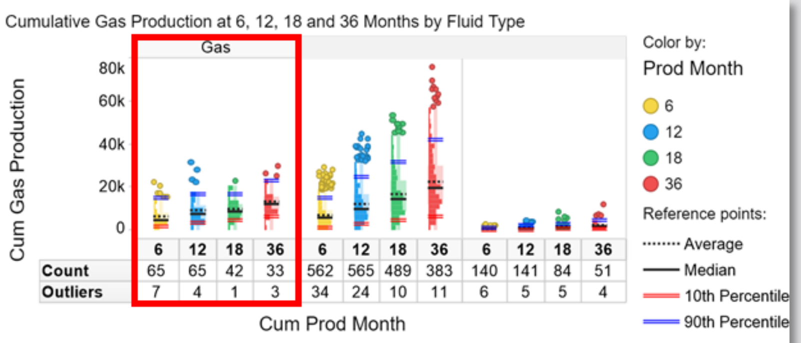
# Unconventional Reservoir - Visualization / Data Analysis

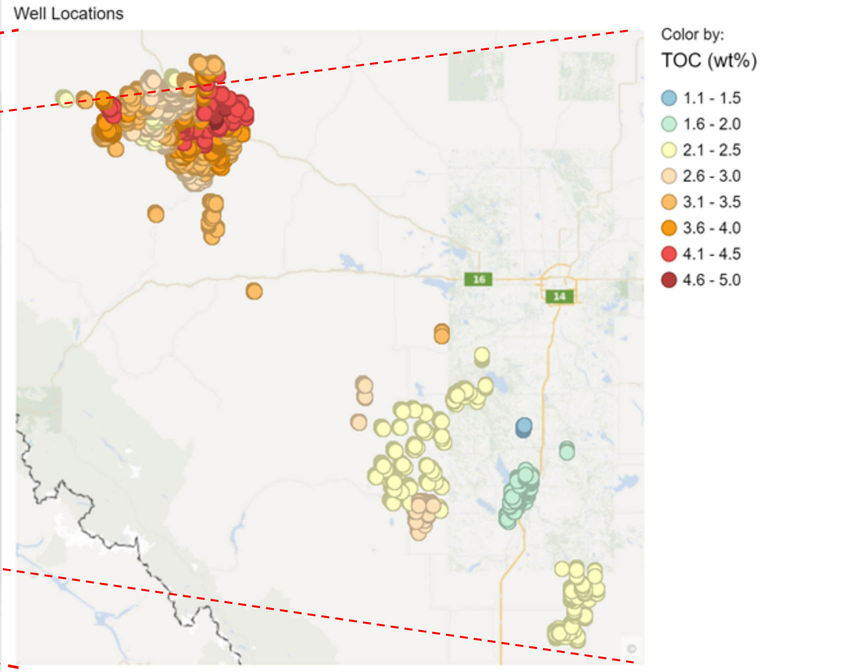
Public production data and hydraulic fracturing completion information from more than 800+ horizontal multi-fractured shale wells was collected.

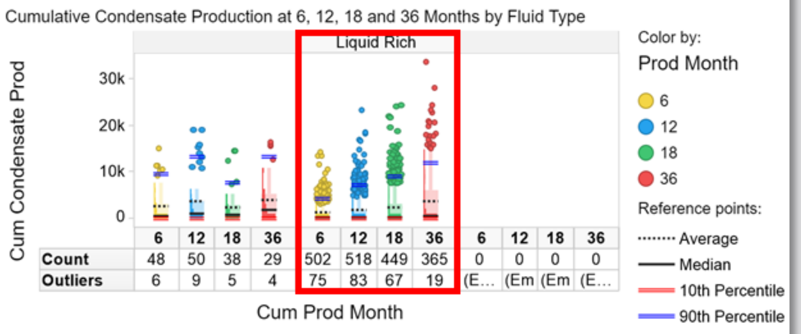
A range of TOC values was assigned per well following published maps by the Provincial Energy Regulator

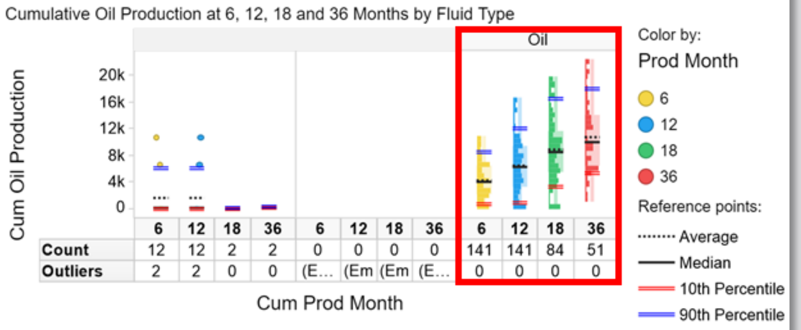


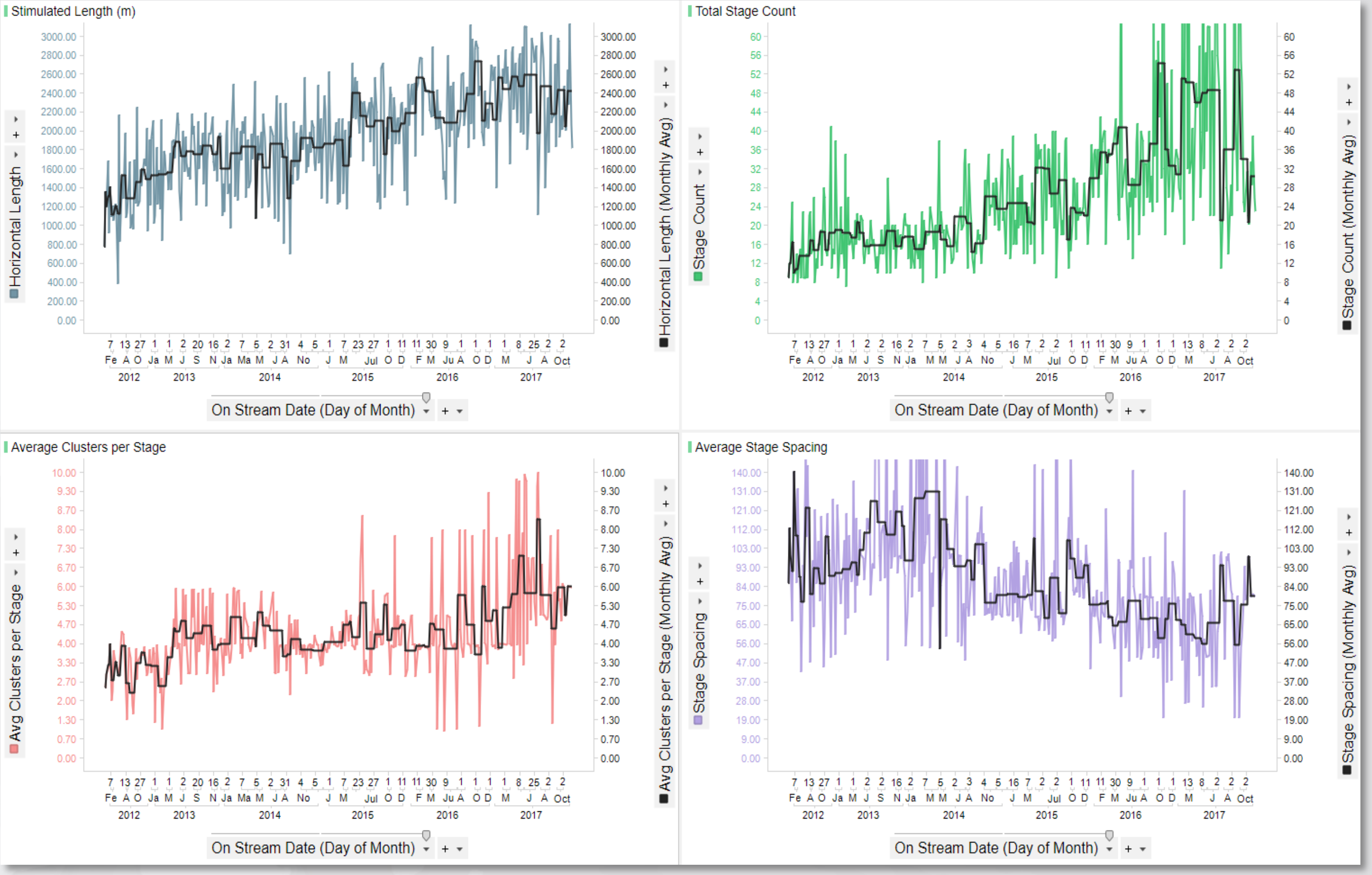




Source: Lyster, S., Corlett, H.J. and Berhane, H. (2017): Hydrocarbon resource potential of the Duvernay Formation in Alberta – update; Alberta Energy Regulator, AER/AGS Open File Report 2017-02, 44 p.

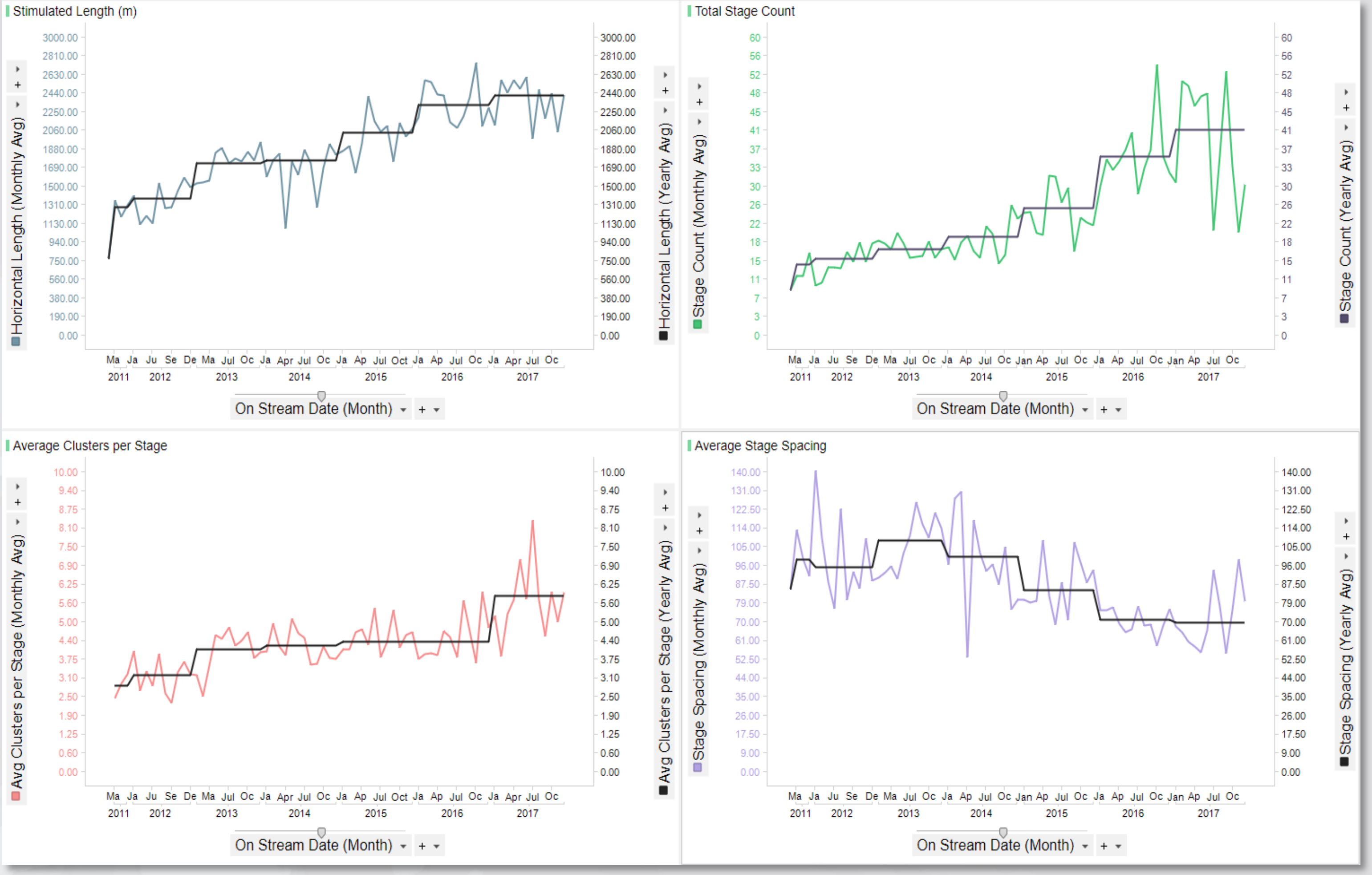
# Unconventional Reservoir - Visualization / Data Analysis

Completion practices have evolved over time by increasing the stimulated horizontal length, stage count and clusters per stage which has led to a tighter stage spacing along the horizontal section.



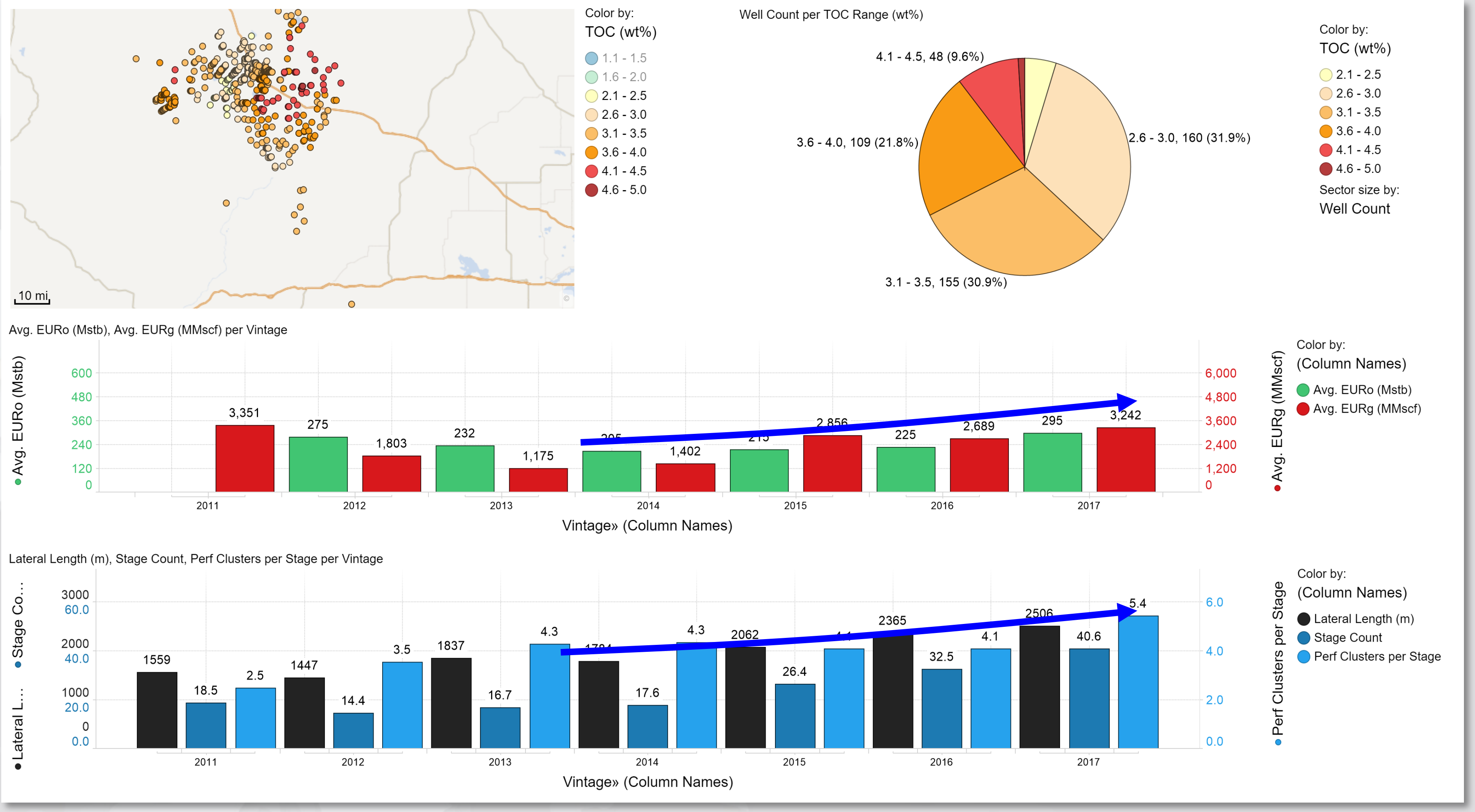
# Unconventional Reservoir - Visualization / Data Analysis

By changing the level of aggregation, it is possible to reduce the noisy profiles and have a clearer picture of the trend changes in completion practices over time.



# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

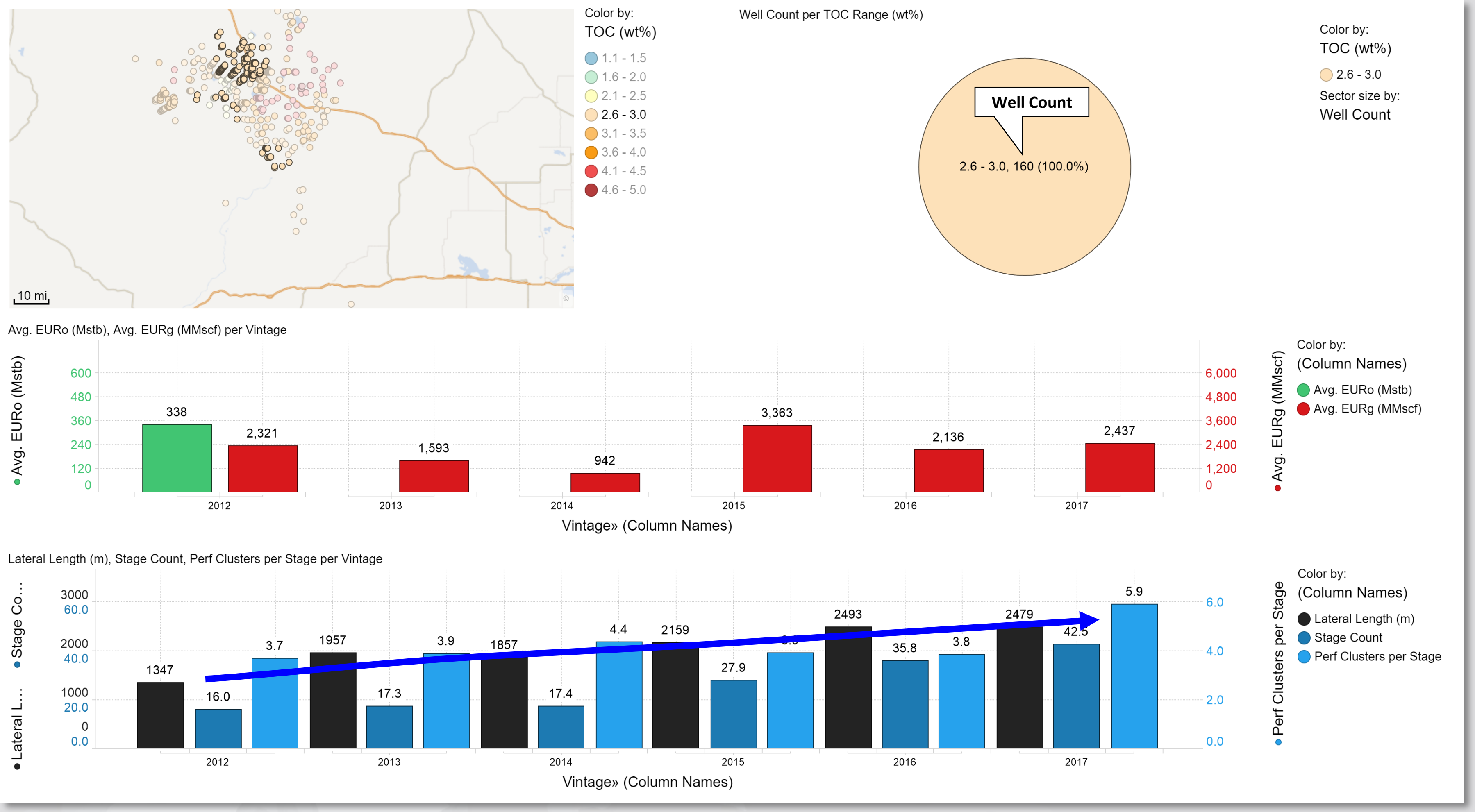


# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

The TOC range selected:

2.6 and 3 (wt%)

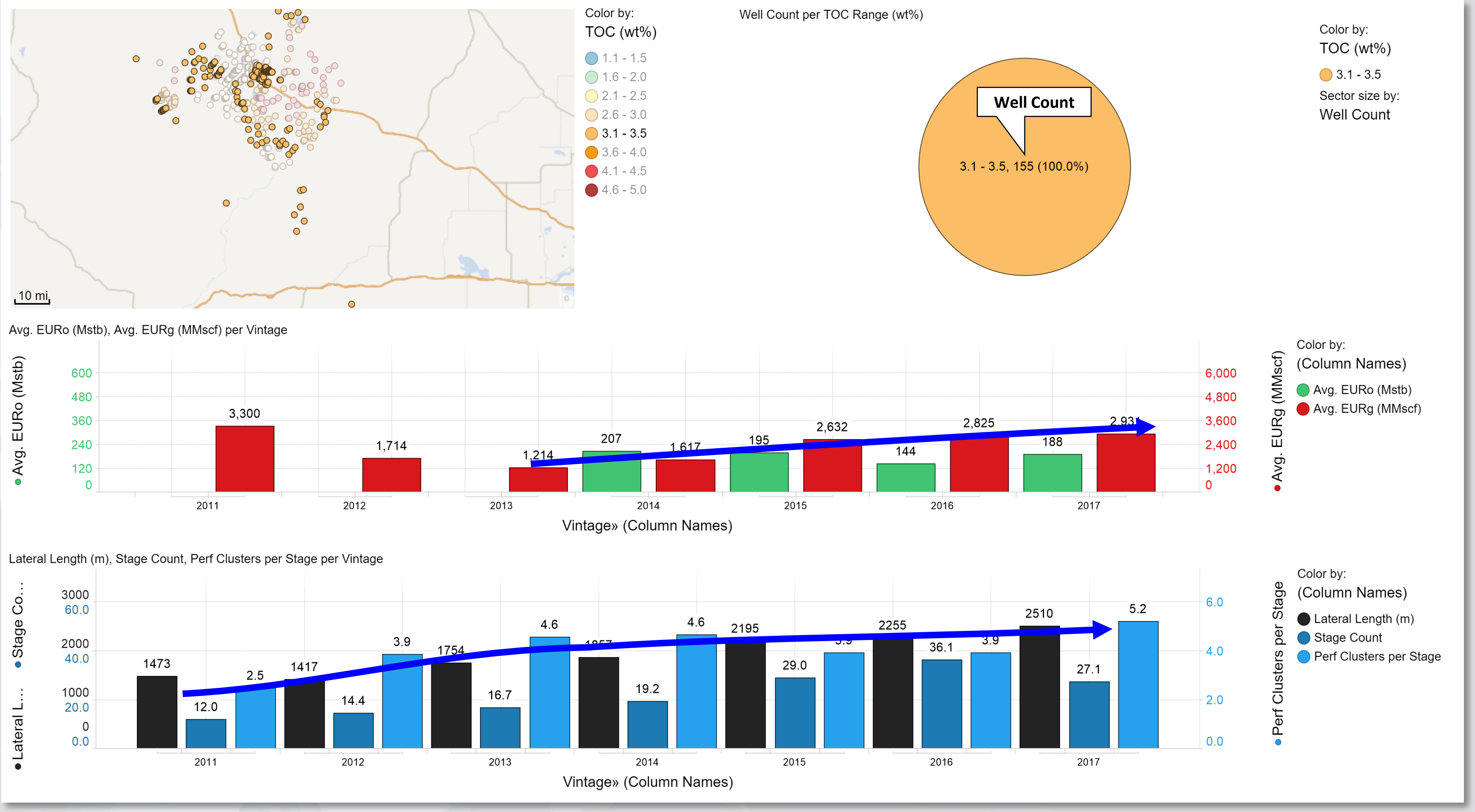


# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

The TOC range selected:

3.1 and 3.5 (wt%)

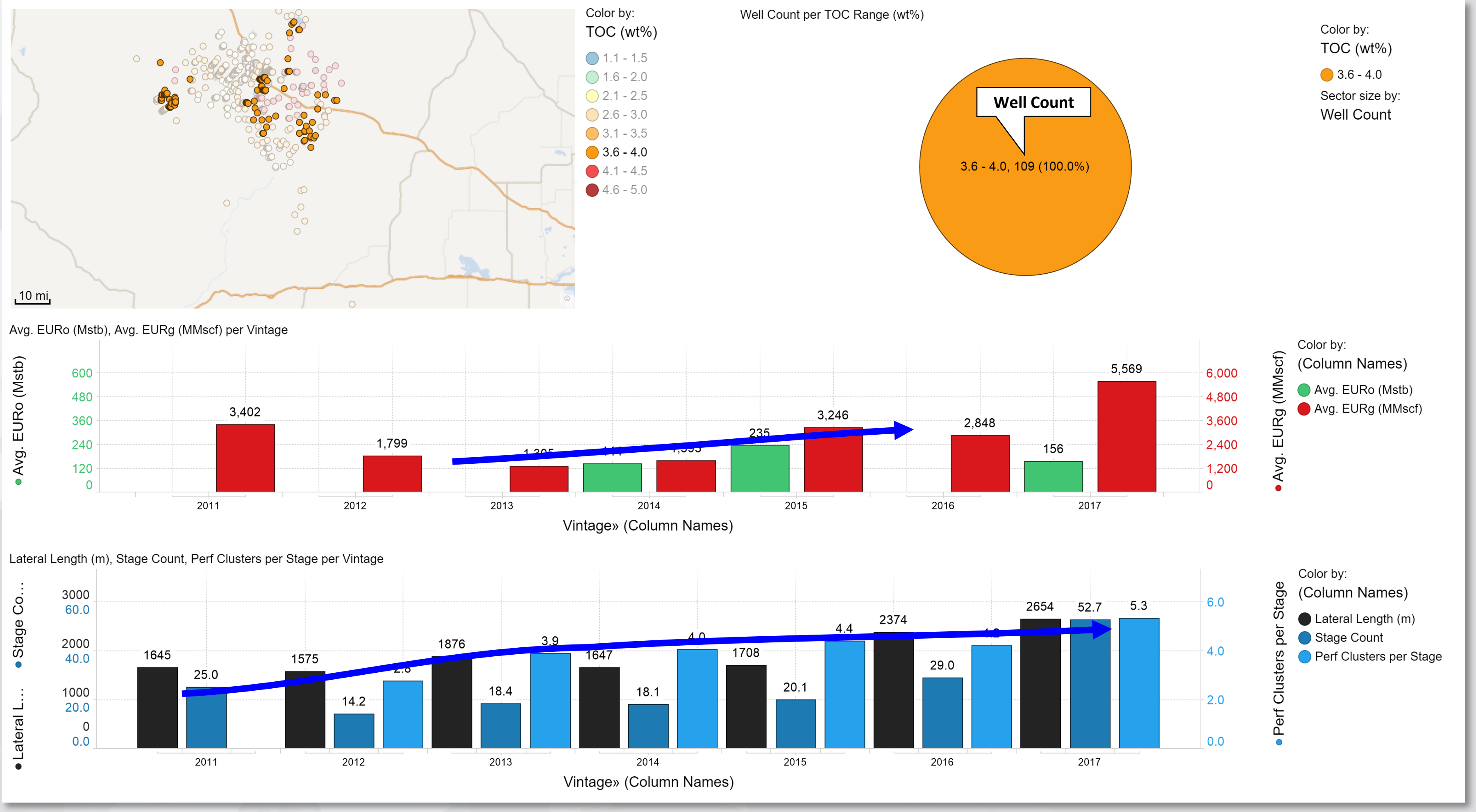


# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

The TOC range selected:

3.6 and 4.0 (wt%)

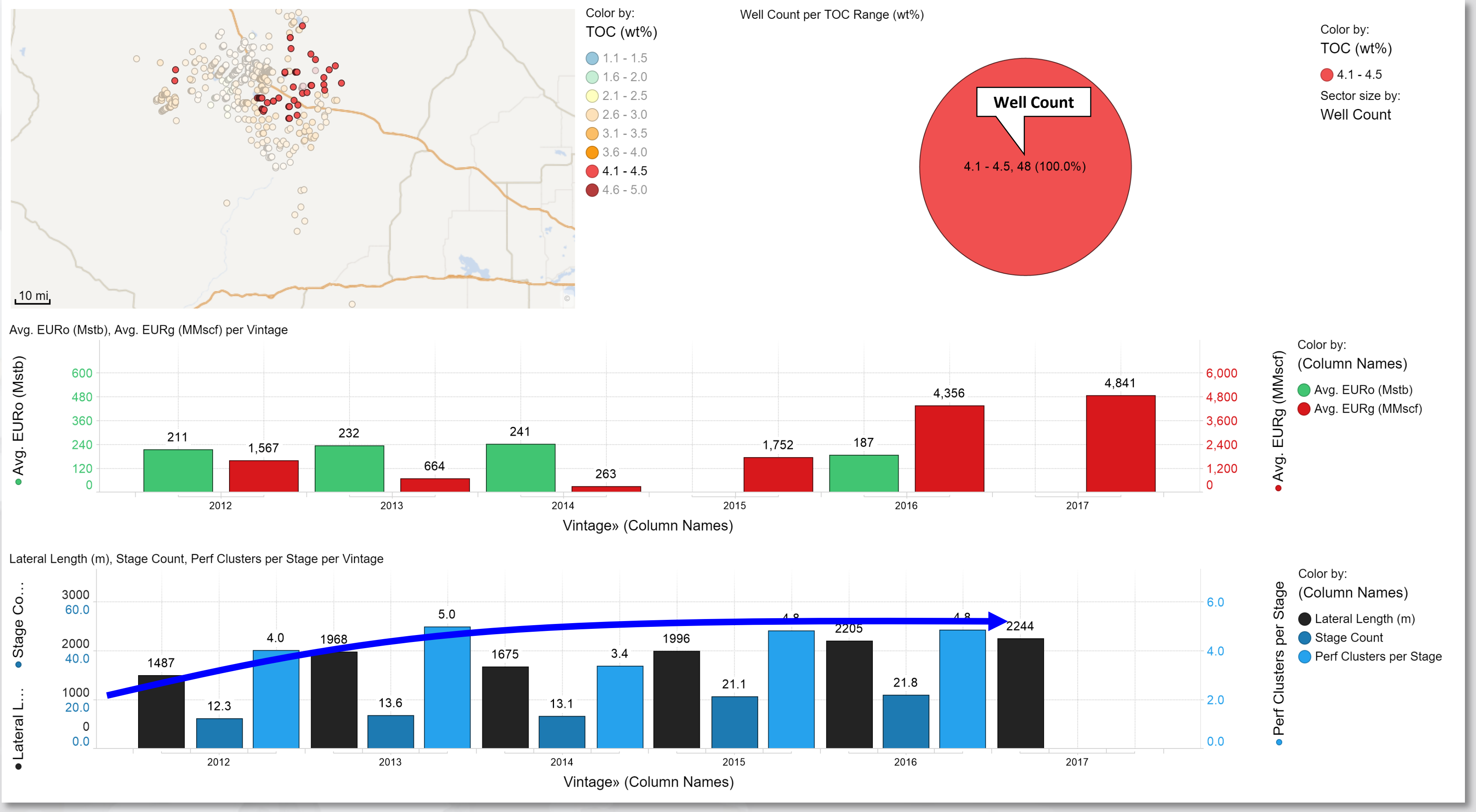


# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

The TOC range selected:

4.1 and 4.5 (wt%)

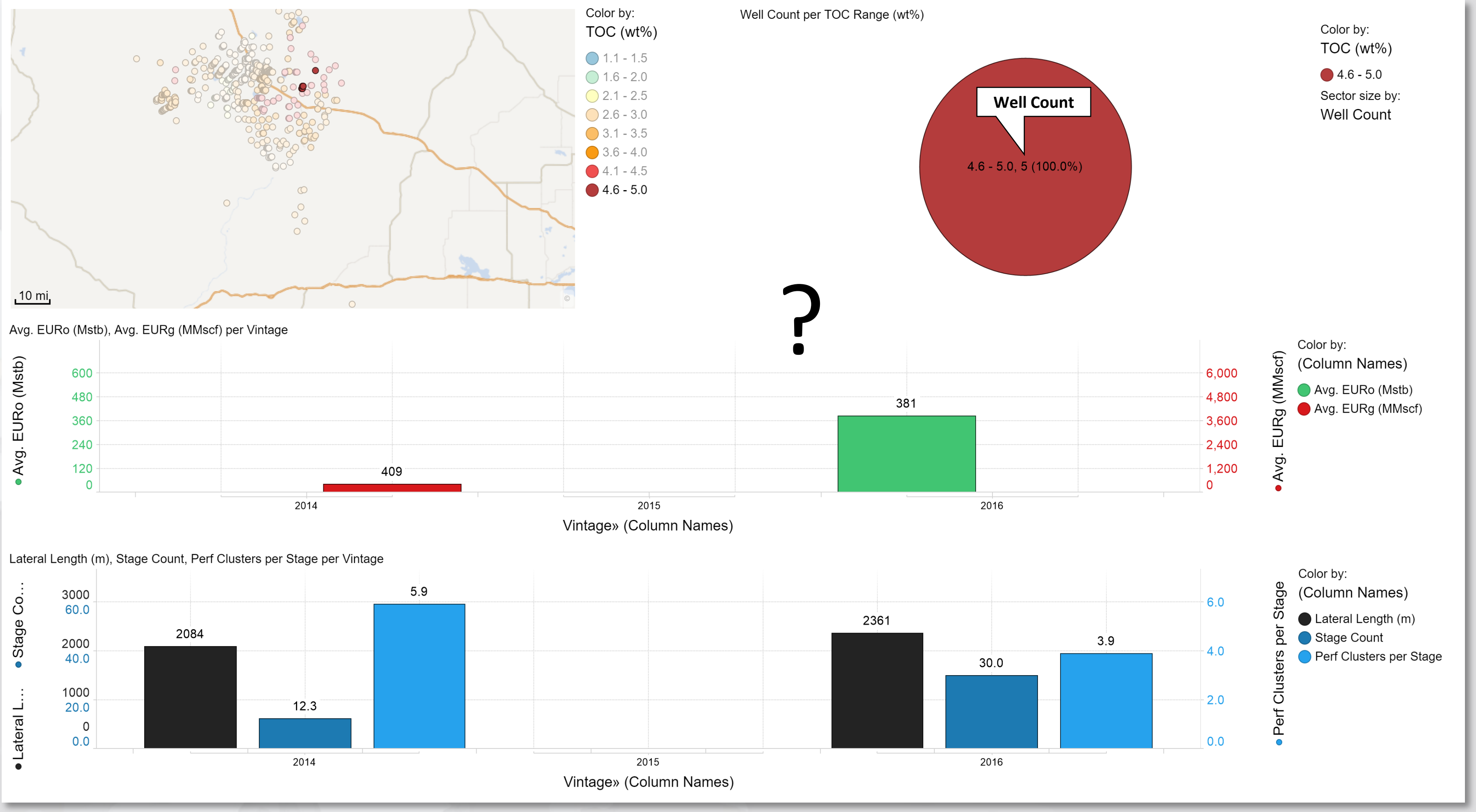


# Unconventional Reservoir - Visualization / Data Analysis

TOC distribution per well and its association with EUR's trends and the stimulation technology used per year. (Kaybob field).

The TOC range selected:

4.6 and 5.0 (wt%)




**This type of analysis should be considered as the preliminary step to a much detailed modeling using AI/ML techniques.**

# Conclusions

- The data management process commonly represents up to 80% of the execution of data analytics / AI / ML projects.
- The value and representativeness of the interpretation is based on the quality of the data.
- Completion practices have changed over time, and the EUR trend follows these changes (increasing horizontal length, number of stages and/or clusters per stage).
- This type of work can also be followed as a first step towards ML algorithms (e.g. training and validation data sets) and applications.

# Thank You