How did we get where we are?

What are some of the issues in existing data?
- Obviously bad data (e.g. typos)
- Subtle and more pernicious

Why has it been so difficult to wield potential power of these databases given the advances of modern technical applications?

Technical approach being presented has proven to work (or at least provide a good start) in addressing this problem*

*Does not require new people, more IT people; just some training and organizational commitment
My perspective from a 30 year career in the Petroleum industry

Database issues in ArcGIS – Geoscience specific

My experience as a geoscientist with ArcGIS and databases
• My job was to teach spatial analysis and map based play assessment
• The reality of the job: resolving data issues / creating data sets from graphics and digitized hand drawn illustrations.

Data issue remains one of the big stumbling blocks to ArcGIS usage in geoscience
Perhaps you have heard these comments from a geologist:
• I have found flaws in this database
• I cannot use it with confidence
• I’m starting over from scratch

Reality is:
World is mature with respect to exploration (lots of data out there)
Applications are very capable and not limited in scope or scale

*Why are there so many persistent issues with data quality and data conditioning*
Description of Data Problem

Technical Issues: Data Complexity

Geoscience data has (at least) 5 degrees of freedom

- X Y (spatial coordinates)
- Z (value – e.g. depth of formation)
- Geologic time
- Changing geologic classification schemes (an evolving scientific discipline)

Data sets that are captured reflect the concepts and time schemas that were in place the data was captured

- Litho-stratigraphy
- Stratigraphic “type sections” or locales
- Biostratigraphy data
- Palynology data
- Chrono-stratigraphy from different schema around the world
- Radiometric Dates – (needs reverse translation)

Note:
This discussion is about the consistency of the data such that it can be analyzed as it currently exists. Interpreted tops, unconformity surfaces, etc. still are the primary venue of a geoscientist and a separate discussion
Confounding Issues
(if you don’t know where you are going… all roads lead you there.)

Organizational Issues

Data issues are poorly appreciated at an executive level

Industry has already poured lots of money into the problem

Data problems are often viewed as principally an IT issue, not requiring a cross disciplinary solution

End users (geoscientists) often create their own local data (nut piles) and become married to their own databases and database structures

Sloppy usage by geoscientists and mixing of chrono-stratigraphic and lithologic terms (ongoing issue)

We have met the enemy and he is us!
Why the obsession with geologic time?

- Provides fundamental structure to understand sedimentary stacking patterns and groupings of sediment packages
- Geologic plays and play elements are organized by temporal stratigraphy
- Historic calibration of field sizes are organized by age classification
- These issues apply for both conventional and unconventional exploration

Bottom line
Lack of attention to how data gets classified has profound influence on what constitutes hydrocarbon potential or lack thereof and (by extension) exploration strategy
Desired State of Databases
(easy to describe…. challenging to do)

Uniform and consistent time – stratigraphic formats for all the data
Interpretation consistent with that common temporal framework

How do we get there from here?

First step is to get the existing databases into a consistent modern time framework

- What does a typical “before” data set look like
- How do you deal with databases that have inconsistent attribution
- How can you systematically and thoroughly reclassify your data into a consist framework.
A Consistent Temporal Framework
Lots of “google help” out there albeit dated

Some useful reference to geologic age classifications

But........
This reference is the 2004 standard.

Standards are almost continually being modified
Geologic Stage to MYBP age refinement (1937 – 2004)

Databases frequently lock in the age schemas at the time when they were captured.

In many cases other classifications are also included:
- Litho-stratigraphy
- Type sections or locales
- Local formation names
- Ages having no formal definition (e.g. Middle Cretaceous or Middle Oligocene)

Geologic Ages are a Moving Target

Geologic Time changes through time

International Commission on Stratigraphy

Mesozoic comparison chart (color code according to the Commission de la Carte Geologique de Monde, Paris)
**INTERNATIONAL CHRONOSTRATIGRAPHIC CHART**

[Image of a stratigraphic chart showing the geological time scale from the Archean to the Present, with various stages, series, and epochs, each with their respective numerical ages and GSSP (Global Stratotype Section and Point).]

**Legend**
- **Phanerozoic**: Light blue
- **Neoproterozoic**: Red
- **Mesoproterozoic**: Pink
- **Paleoproterozoic**: Orange
- **Archean**: Green

**Key**
- **Epochs**: Upper, Middle, Lower
- **Series**: Oligocene through Quaternary
- **Stage**: Liassic through Turonian
- **Numerical age (Ma)**: 0.117 to 4800

**Units of all rases are in the process of being defined by Global Boundary Stratotype Section and Points (GSSP) for their lower boundaries, including those of the Archean and Proterozoic, long defined by Global Standard Stratigraphic Ages (GSSA). Charts and detailed information on called GSSPs are available at the website http://www.stratigraphy.org. The URL to this chart is not below.**

**Numerical ages are subject to revision and do not define units in the Phanerozoic and the Ediacaran, only GSSPs do. For boundaries in the Phanerozoic without ratified GSSPs or without constrained numerical ages, an approximate numerical age (-) is provided.**

**Numerical ages for all systems except Lower Paleozoic, Permian, Triassic, Cretaceous, and Cenozoic are taken from "A Geologic Time Scale 2012" by Gradstein et al. (2012), those for the Lower Paleozoic, Permian, Triassic and Cretaceous were provided by the relevant ICS subcommissions.**

**Coloring follows the Commission for the Geological Map of the World (http://www.cgmw.org)**

**Chart drafted by K.M. Cohen, S.C. Finney, P.L. Gibbard (International Commission on Stratigraphy, February 2014).**

## Translation Table of International Standard Geologic Ages

Table shown below is based on GTS 2014 Classification Scheme

Coarser time intervals provided for flexibility

Can be joined to your data (once conditioned) to standardize it to the International Commission on Stratigraphy GTS2014*

### Period | Epoch | Stage | Young Age MYBP | Early Age MYBP | Interval_length_MY
---|---|---|---|---|---
Neogene | Miocene | Aquitanian | 20.43 | 23.03 | 2.6
Paleogene | Oligocene | Chattian | 23.03 | 28.4 | 5.37
Paleogene | Oligocene | Rupelian | 26.4 | 33.9 | 5.5
Paleogene | Eocene | Priabonian | 33.9 | 37.2 | 3.3
Paleogene | Eocene | Bartonian | 37.2 | 40.4 | 3.2
Paleogene | Eocene | Lutetian | 40.4 | 48.6 | 8.2
Paleogene | Eocene | Ypresian | 48.6 | 55.3 | 7.2
Paleogene | Paleocene | Thanetian | 55.8 | 56.7 | 2.9
Paleogene | Paleocene | Selandian | 56.7 | 61.7 | 5
Paleogene | Paleocene | Danian | 61.7 | 65.5 | 3.8
Cretaceous | Late Cretaceous | Maastrichtian | 85.5 | 70.8 | 14.7
Cretaceous | Late Cretaceous | Campanian | 70.8 | 33.5 | 12.3
Cretaceous | Late Cretaceous | Santonian | 83.5 | 35.5 | 2.5
Cretaceous | Late Cretaceous | Coniacian | 85.8 | 39.3 | 3.5
Cretaceous | Late Cretaceous | Turonian | 80.3 | 93.5 | 4.2
Cretaceous | Late Cretaceous | Cenomanian | 93.5 | 98.6 | 5.1
Cretaceous | Early Cretaceous | Albian | 98.6 | 112 | 12.4
Cretaceous | Early Cretaceous | Aptian | 112 | 125 | 13
Cretaceous | Early Cretaceous | Barremian | 125 | 130 | 5
Cretaceous | Early Cretaceous | Hauterivian | 130 | 136.4 | 6.4
Cretaceous | Early Cretaceous | Valanginian | 136.4 | 140.2 | 3.8
Cretaceous | Early Cretaceous | Herrinian | 140.2 | 145.5 | 5.3
How do we get these databases under control and deal with all the legacy data chaos?
Extract a unique age list from your Data.

Process your data (Origin Data) using the frequency tool to produce a unique list of age names in a results table more commonly called a Frequency Table.

Build this using the highest resolution age intervals that are in the data table.
Examples of issues in Database Tables

- **Multiple spellings / Litho-stratigraphy**
- **Litho-stratigraphy**
- **Older or alternate age classification**
- **Inconsistent spelling / multiple entries for same interval**
- **Litho-stratigraphy / type section**
- **Biostratigraphy**
- **Litho-stratigraphy**
- **Litho-stratigraphy**
- **Inconsistent spelling / multiple entries for same interval**
- **Multiple spellings / Litho-stratigraphy**
- **Inconsistent spelling / multiple entries for same interval**
- **Trailing space**
Add a column in your Frequency Table to contain correct ages that are standardized: The resultant table is called a Translation Table

**Note:**
This remapping step is the most important step to get right and will require geoscience experience

Add a column in your Origin Table to receive the remapped attributes
Join the **Translation Table** to the **Origin Table**

To make permanent export the feature data set with the join in place
Interactive graphs based on statistics from reclassified age attributes

Age intervals summarized in:

- Total Organic Carbon
- Hydrogen Index
- Vitrinite Reflectance
Recommendations

• Use the frequency tool in ArcGIS to thoroughly look through your existing data
• Commit to a chronostratigraphic standard (and age resolution) everyone wants to work in and carefully reconcile your data to it
• Build and maintain the translation tables
  ➢ For your local data
  ➢ For vendor data
  ➢ For the age model standard

Initial processing and standardization will take a little addition time but there are big benefits:

1. Future raw data sets will have a process in place to deal with them more efficiently
2. All your age specific analysis will become much less daunting to do:
   o Hydrocarbon discovered volumes by age
   o Geologic parameter by age
   o Play assessments
3. Data cleanup product is data that is vetted and able to be loaded into more formal databases (e.g. Oracle)
The requisite sunset picture denoting 

THE END
Resources

**Current International Commission on Stratigraphy: Time Stratigraphic Chart**

http://www.chronos.org
