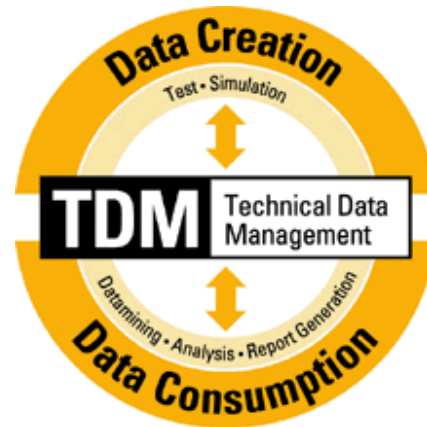


Raw Data to Results: Test Data Management Architectures in the Petabyte Age

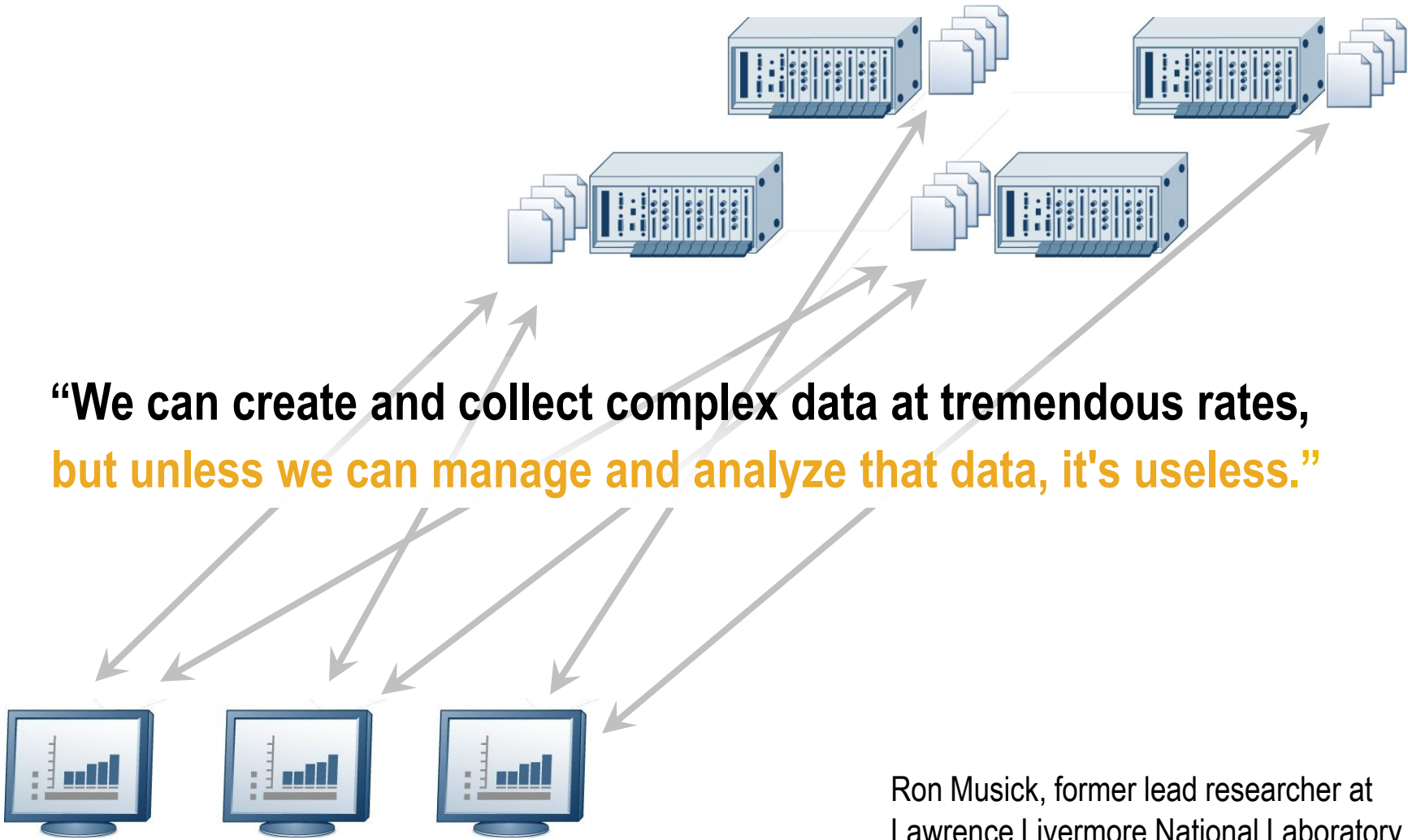


Caroline Bright

*Data Management Product Manager
Automotive Testing Expo 2008*

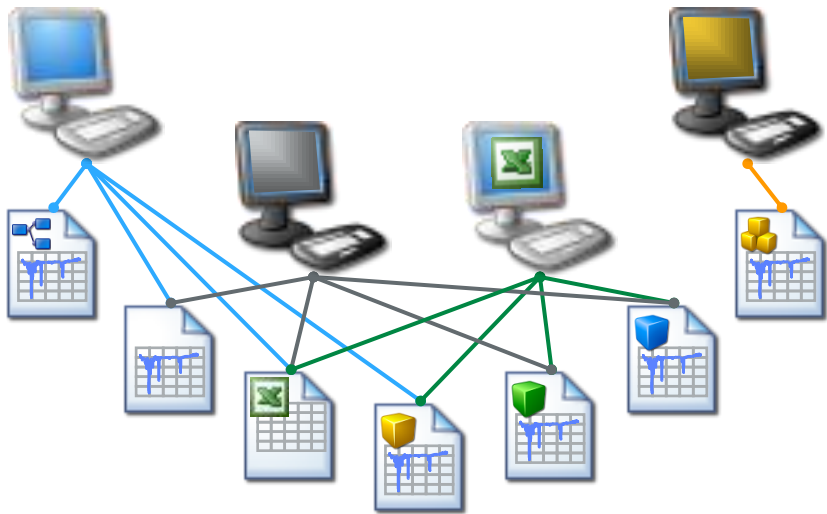
The Data Challenge

**“We can create and collect complex data at tremendous rates,
but unless we can manage and analyze that data, it's useless.”**



Ron Musick, former lead researcher at
Lawrence Livermore National Laboratory

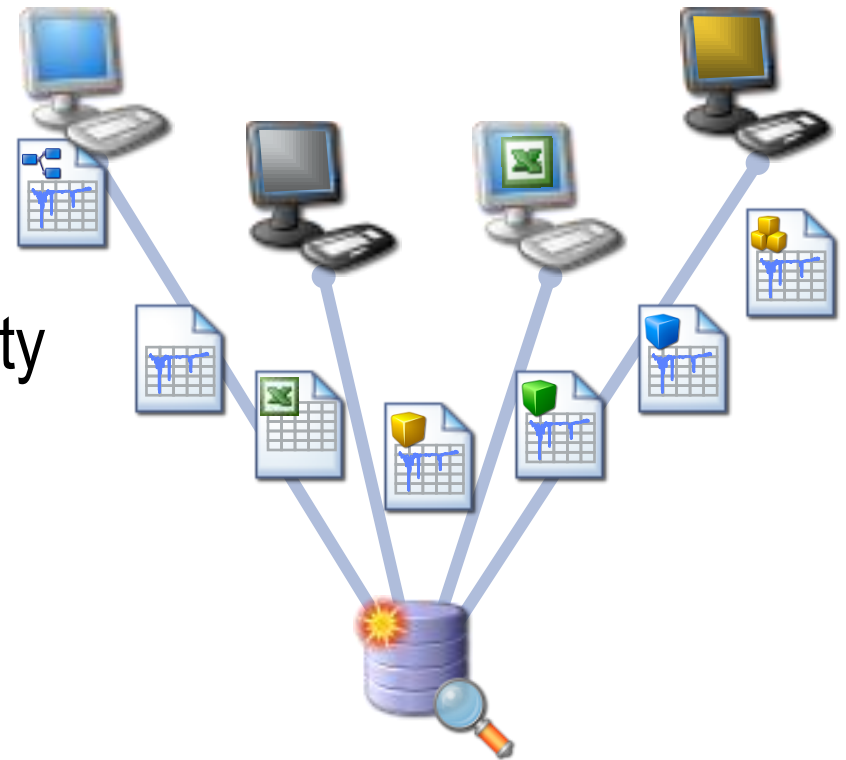
Traditional Reactions to the Data Problem



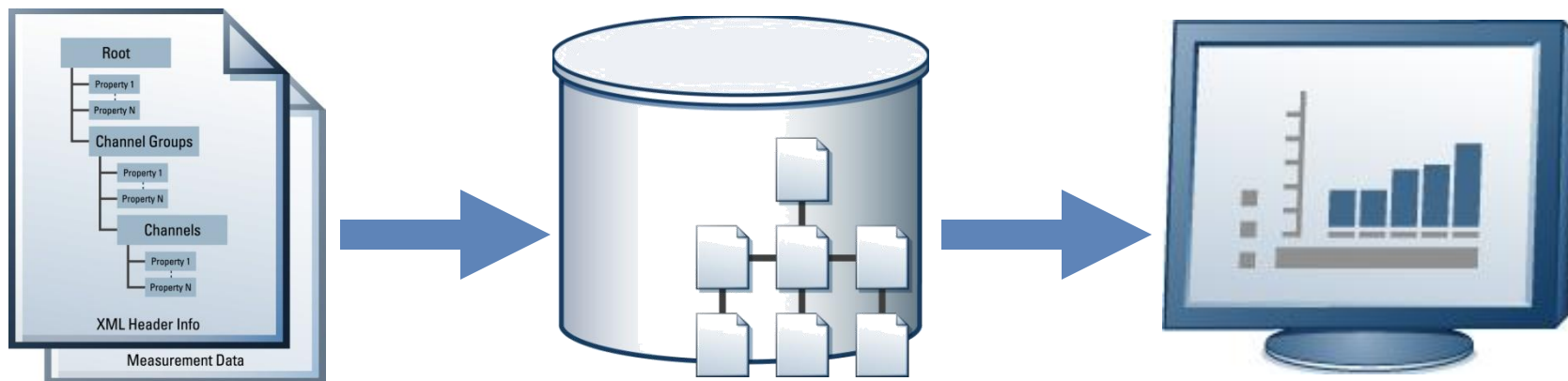
“Every man for them self”
Low initial cost but low productivity

Traditional Reactions to the Data Problem

Top down approach
High cost and debatable productivity



Three Step Architecture for Data Management



TDM Data Model

Indexing

**Search and Post
Processing Interface**

Determining Your Storage Format

When determining the appropriate storage format for your data, consider

1. What will you do with your data once you acquire it?
2. Will you write and read data with the same application?
3. How much data will you acquire?
4. At what rate will you acquire data?
5. Will you need to exchange data with another program?
6. Will you need to search your data files?

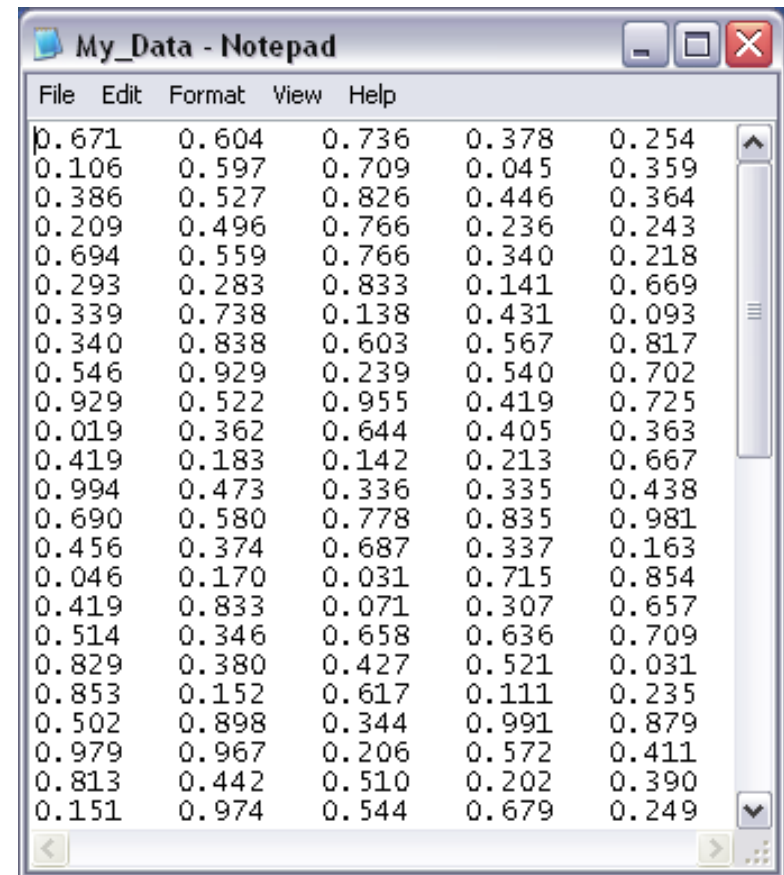
ASCII Files

Characteristics

- Human-readable
- Easily portable to other applications such as Microsoft Excel

Considerations

- Significantly larger disk footprint
- Slow read and write



The screenshot shows a Notepad window with the title 'My_Data - Notepad'. The window contains a table of numerical data with 5 columns and 20 rows. The data is as follows:

0.671	0.604	0.736	0.378	0.254
0.106	0.597	0.709	0.045	0.359
0.386	0.527	0.826	0.446	0.364
0.209	0.496	0.766	0.236	0.243
0.694	0.559	0.766	0.340	0.218
0.293	0.283	0.833	0.141	0.669
0.339	0.738	0.138	0.431	0.093
0.340	0.838	0.603	0.567	0.817
0.546	0.929	0.239	0.540	0.702
0.929	0.522	0.955	0.419	0.725
0.019	0.362	0.644	0.405	0.363
0.419	0.183	0.142	0.213	0.667
0.994	0.473	0.336	0.335	0.438
0.690	0.580	0.778	0.835	0.981
0.456	0.374	0.687	0.337	0.163
0.046	0.170	0.031	0.715	0.854
0.419	0.833	0.071	0.307	0.657
0.514	0.346	0.658	0.636	0.709
0.829	0.380	0.427	0.521	0.031
0.853	0.152	0.617	0.111	0.235
0.502	0.898	0.344	0.991	0.879
0.979	0.967	0.206	0.572	0.411
0.813	0.442	0.510	0.202	0.390
0.151	0.974	0.544	0.679	0.249

XML Files

Characteristics

- Stores complex data structures
- Shows display in a Web browser or in a text editor

Considerations

- Even larger disk footprint
- Front-end schema design
- Does not stream

```
<?xml version="1.0" standalone="yes" ?>
- <LVData>
  <Version>7.0</Version>
  - <DBL>
    <Name>Max</Name>
    <Val>0.81299</Val>
  </DBL>
  - <DBL>
    <Name>Min</Name>
    <Val>-0.86182</Val>
  </DBL>
  - <DBL>
    <Name>Mean</Name>
    <Val>0.10645</Val>
  </DBL>
</LVData>
```

Databases

Characteristics

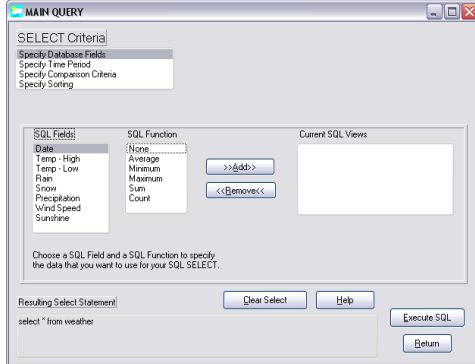
- Store data centrally
- Organize and query test results with Access or MySQL

Considerations

- Programming can be more time intensive
- Require maintenance
- Potentially high cost
- IT

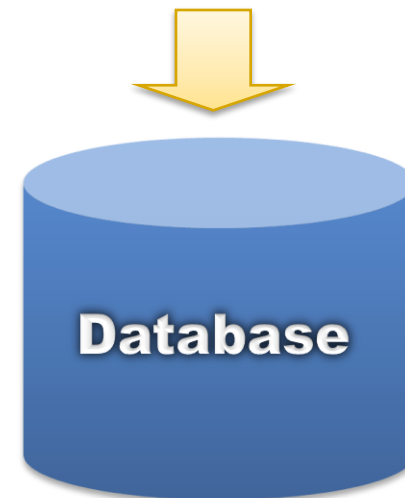
```
/* Connect to data source (in this case dBase files) */
hdbc = DBConnect ("DSN=CVI SQL 2.0 Samples");
if (hdbc <= 0) {ShowError(); {ShowError(); goto Error;}}

/* begin map for constructed SQL statement */
hmap = DBBeginMap (hdbc);
if (hmap <= 0) {ShowError(); goto Error;}
```

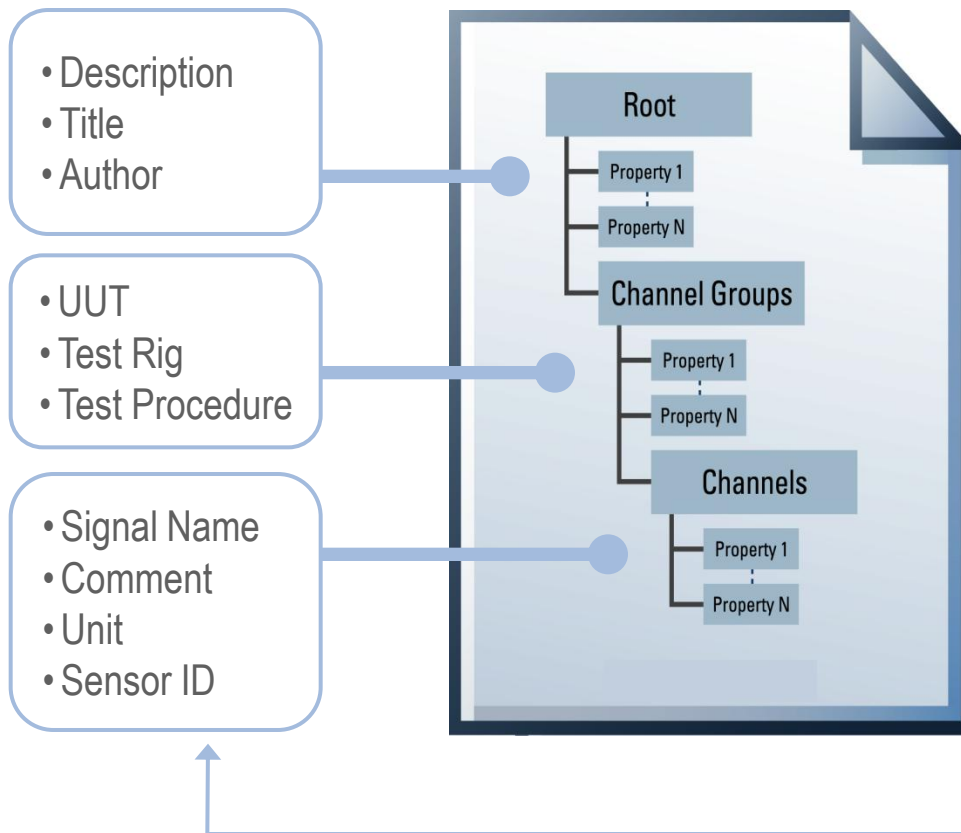


```
ected and the variables where column */
sp, "UUT_NUM", 11, uutNum, &uutStat, "");
>Error(); goto Error;}}
Error(); goto Error;}}
map, "MEAS1", &meas1, &meas1Stat);
>Error(); goto Error;}}
map, "MEAS2", &meas2, &meas2Stat);
>Error(); goto Error;}}

estres, (construct a SQL Select
ent, bind the selected columns to
ables.)
ESTRES");
to Error;}}
```



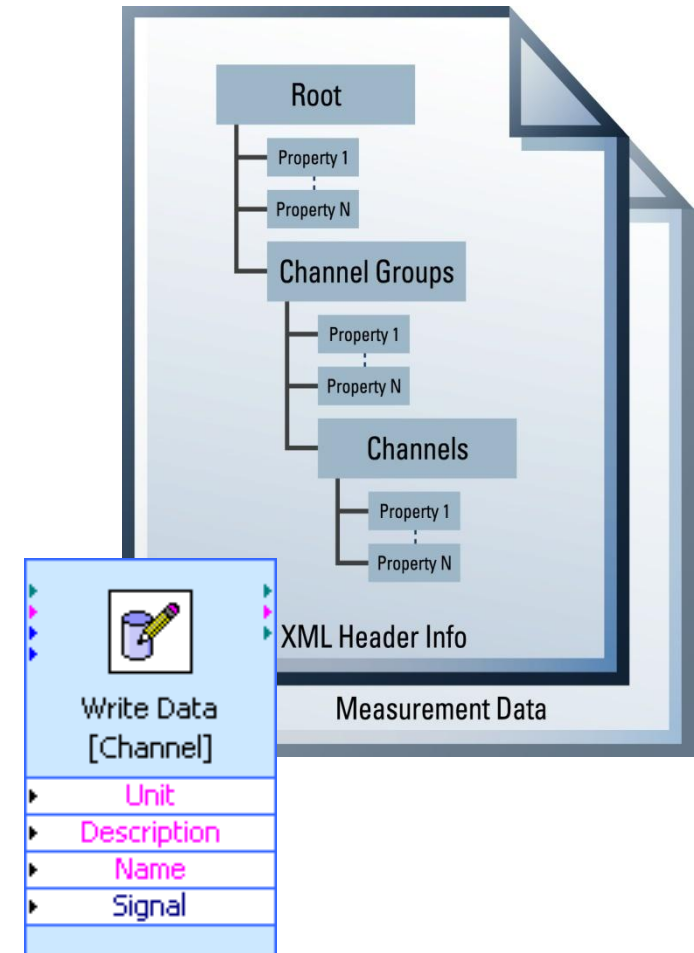
The TDM Data Model



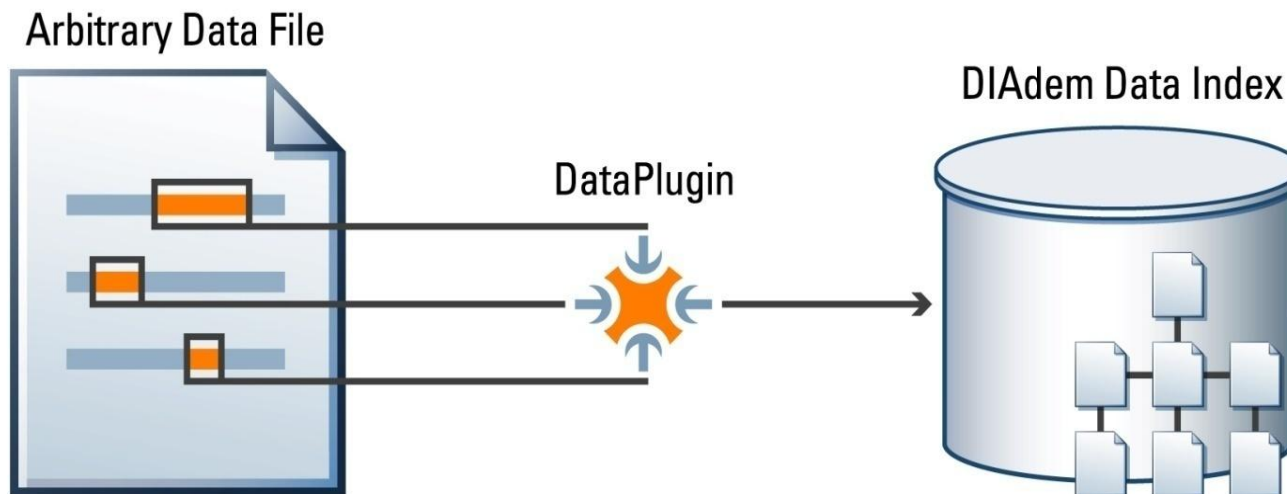
- Simple and flexible
- 3 levels of hierarchy
 - root (e.g. file)
 - groups (e.g. test steps)
 - channels (e.g. signals)
- User defined properties on each hierarchy level

TDM-File format: Optimized for data storage and search

- **Simple:** Easiest approach for storing measurement data in NI software
- **Flexible:** Add custom attributes to every file, channel group and channel
- **Fast:** TDMS, the TDM file for streaming applications
- **Open:**
 - Microsoft Excel and Open Office Add-In
 - Public documentation
 - C DLL for creating files in 3rd party applications



Indexing Technology with NI DataFinder

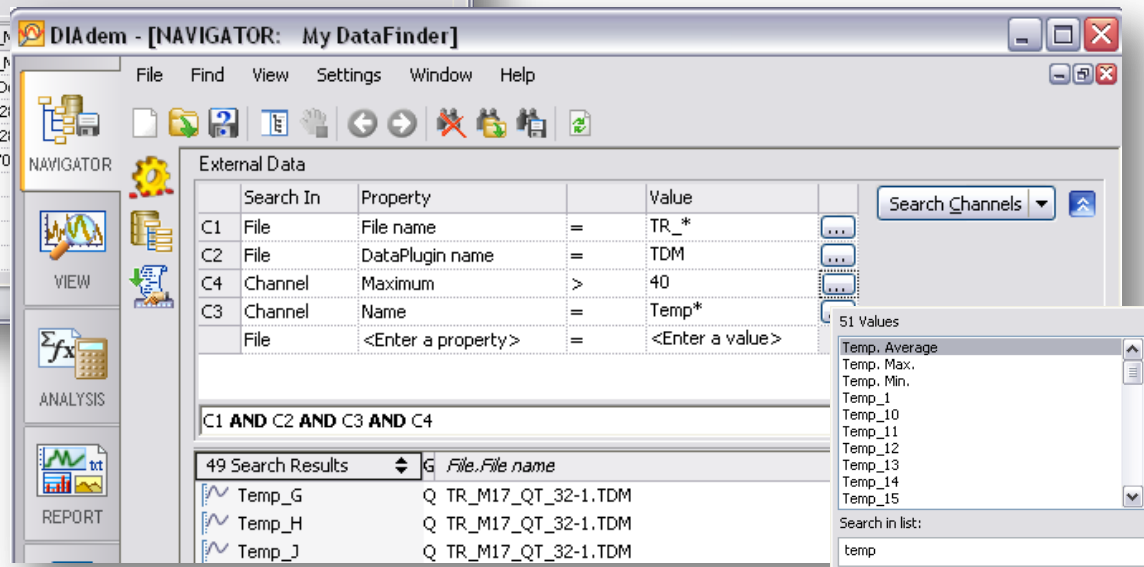
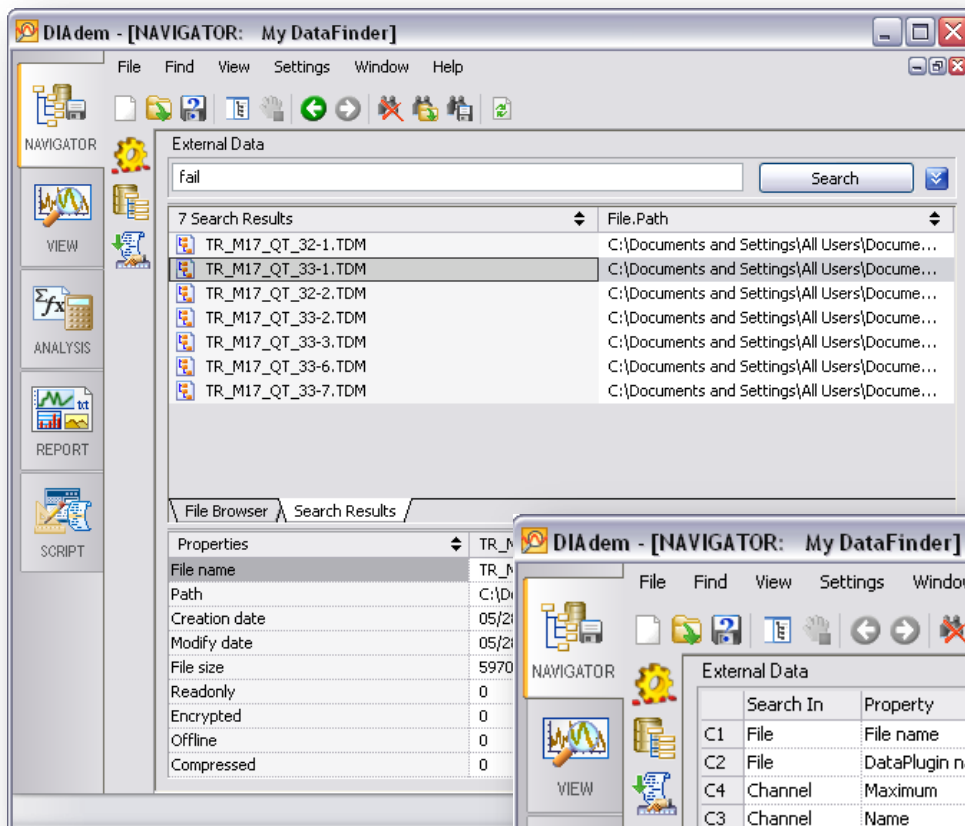


- Stores descriptive information from data files
- It works with **any** data file you have a DataPlugin for *
- Builds and scales automatically
- Requires no IT support to install, configure or maintain

*visit www.ni.com/dataplugins

Data Management and Mining

- Organize and search data using keyword search or an interactive parametric search
- Load data from both files and databases
- Work with up to 2 billion values per channel



Evaluating Climate Data: The DIAdem DataFinder Moves You from Separate Solutions to an Overall Evaluation System

Application: Collecting, analyzing and reporting climate-related meteorological and hydrological data in the Bavarian Forest for preventative environmental protection

Challenge: Designing a data management system to handle large amounts of data, from multiple test stations, written in different formats

Technology: NI DataFinder, DIAdem, TDM file format, NI DataPlugins



“...DIAdem and the DataFinder have achieved an integrated data organization solution that guarantees a high degree of data security and continuous monitoring of the measurement stations.”

Additional Resources

- ni.com/diadem
- ni.com/tdm
- ni.com/datafinder
- ni.com/dataplugins

