MARKLOGIC WORLD 19

LLNL reins in data to support the stockpile stewardship program

Valerie Noble

Directorate Applications Manager

Lawrence Livermore National Laboratory

LLNL Reins in Data to Support the Stockpile Stewardship Program

Valerie Noble, LLNL, Directorate Applications Manager

May 14, 2019

LLNL-PRES-773226 This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Lawrence Livermore is a Nuclear Weapon Physics and Engineering Laboratory

Livermore, California (45 miles east of San Francisco)

Site 300 Experimental Test and HE Operations
(11 miles^2, 15 miles east of LLNL)

- **Established in 1952**
- **~7,300 employees (ST&E: 47% Ph.D.)**
- **7.1 million SF, 684 facilities**
- **Annual federal budget: ~$1.6B (63% stockpile stewardship)**

LLNL's core mission is nuclear security

Stockpile Stewardship is Cradle to Grave

Stockpile Stewardship is an integrated, multi-disciplinary program to ensure the nuclear deterrent without nuclear testing

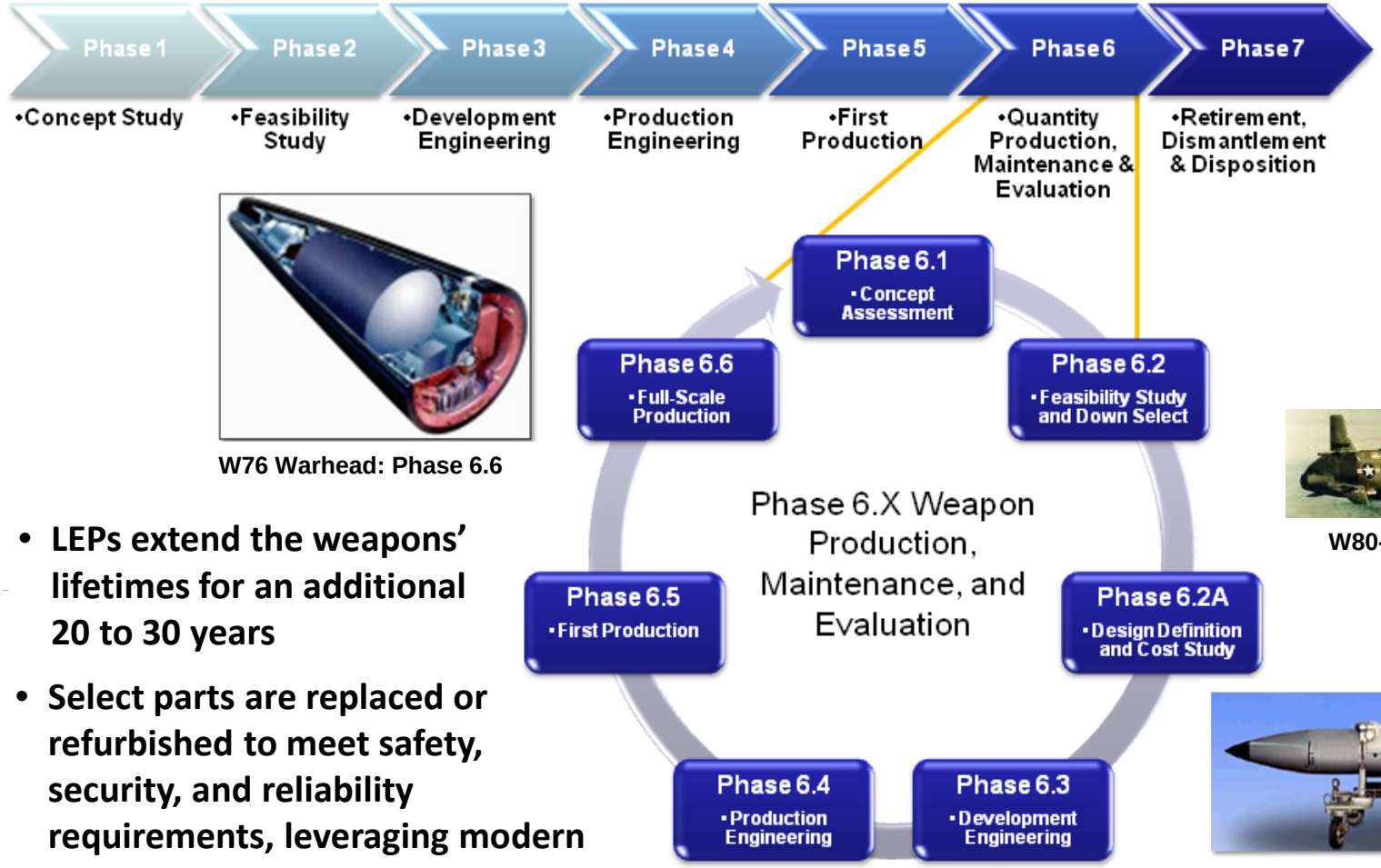
"You can imagine that if you had car [25-30 years] old, you'd probably want to get a new one, but you can't … You've got to sustain that car – by the way, you aren't allowed to start it to see if it works, you're not allowed to test it – but it has to work with 100% reliability when you do – God forbid – ever have to start it."

…Bruce Goodwin, former Principal Associate Director, LLNL

Design and Manufacturing	***Surveillance***	***Assessment and Certification***
Ability to design and manufacture weapons, dismantle weapons, and store components	Ability to accurately determine the state of health of the stockpile	Ability to quantify with confidence the safety, reliability and effectiveness of the stockpile

The Stockpile Stewardship Program has successfully maintained the nuclear deterrent without nuclear testing since 1992

Current focus is on Life Extension Programs (LEPs)



W76 Warhead: Phase 6.6

W80-4 Warhead: Phase 6.3

- **LEPs extend the weapons' lifetimes for an additional 20 to 30 years**
- **Select parts are replaced or refurbished to meet safety, security, and reliability requirements, leveraging modern technologies**

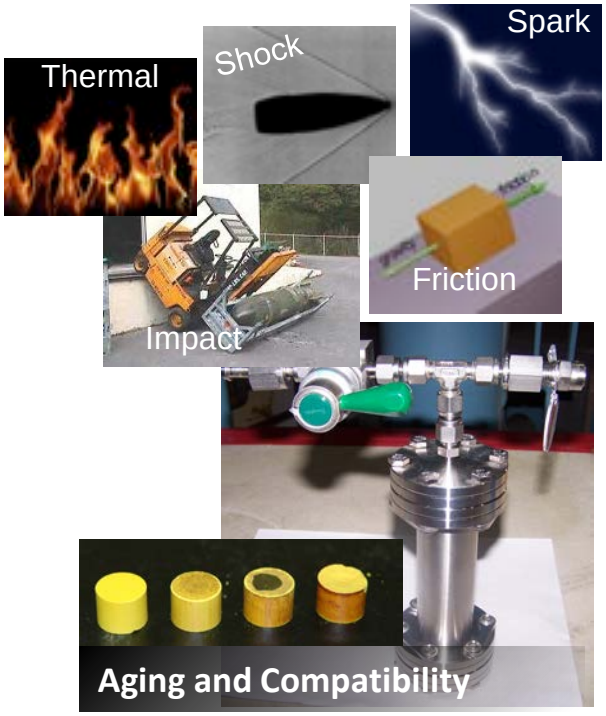
B61 Bomb: Phase 6.3

LEPs generate petabytes of data

- Tests conducted during an LEP may cost millions of dollars to execute, so we should treat the resultant data as a million dollar asset
  - Test data from previous LEPs was stored across many different file systems, each controlled by various groups and individuals with their own data management protocols and priorities
  - Much of the data collected was lost or made useless due to hardware/technology obsolescence, personnel changes, or lack of context

Why is this so hard?

- **An LEP generates large volumes of data from multiple sources and in many formats:**
  - Full system engineering tests
  - Material tests
  - Flight tests
  - Component level tests
- **Datasets are unpredictable, and highly variable**
- **Contextual information about the data is just as important as the data itself**
- **If post-processing is performed, the results must be captured and associated with the raw data**
- **Individual files can be multi-terabyte in size**



Many options were explored – without success

- **Document management system**

  Limitations

  - Data sets are extremely large and can quickly take up terabytes of space
  - Data sets have unique storage needs to help facilitate interpretation and analysis

- **Fileshare**

  Limitations

  - Can not easily associate context or metadata
  - Downstream search and discovery of data is extremely limited
  - Folder structures can be difficult to manage with multiple layers of folders

- **Relational database**

  Limitations

  - Lack of predictable, heterogeneous data sets drives continuous schema changes
  - Requires coupling to a datastore to accommodate large files

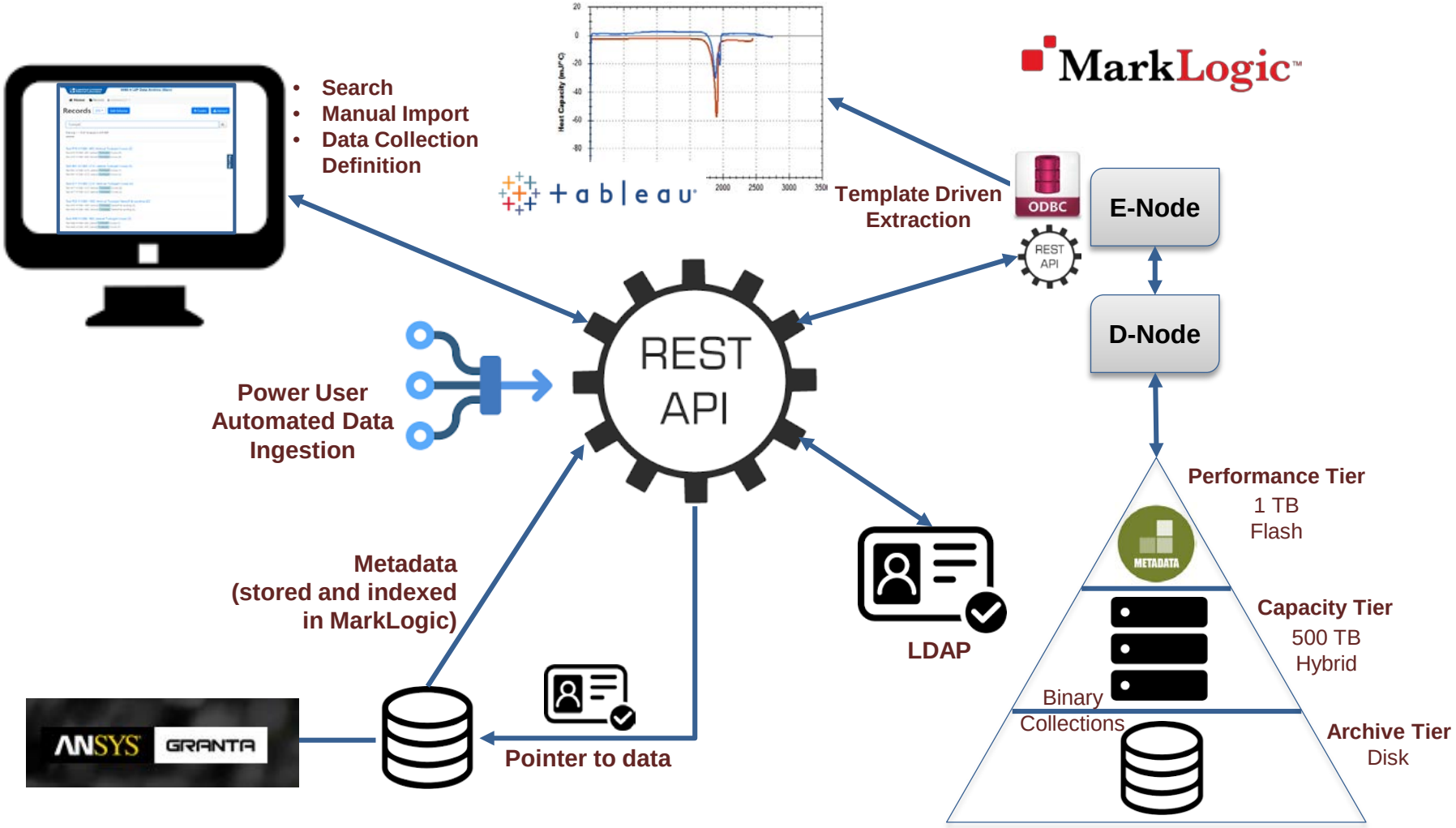
So, why MarkLogic?

- Requirements were collected back in 2003 and a prototype web application was created
- With only a one person team and cultural hurdles to overcome, it was never productized – but the concept was well received

2003, Web-Based Prototype

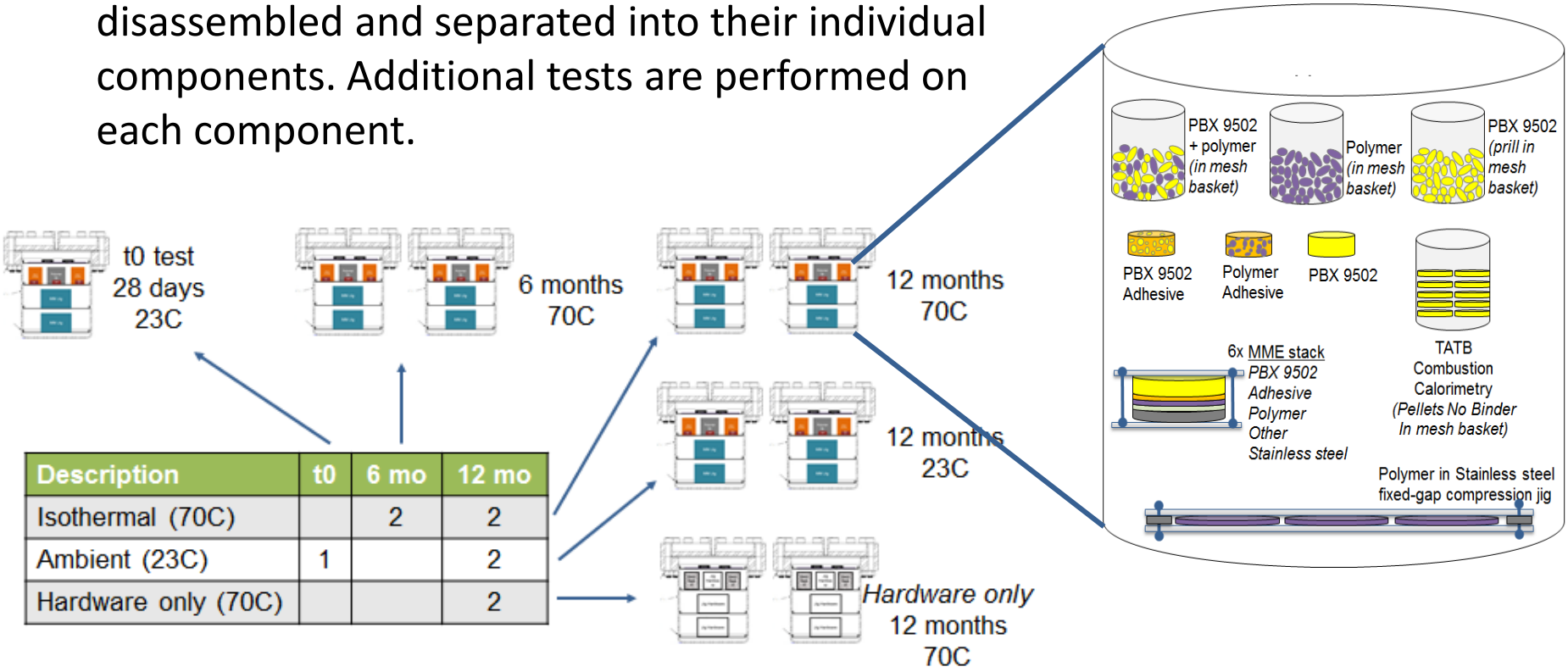
MarkLogic was the first tool that met all of the original 2003 requirements and fully aligned with my vision. It also offers expanded functionality through modern technologies and a flexible framework.

Deploying a robust and flexible architecture



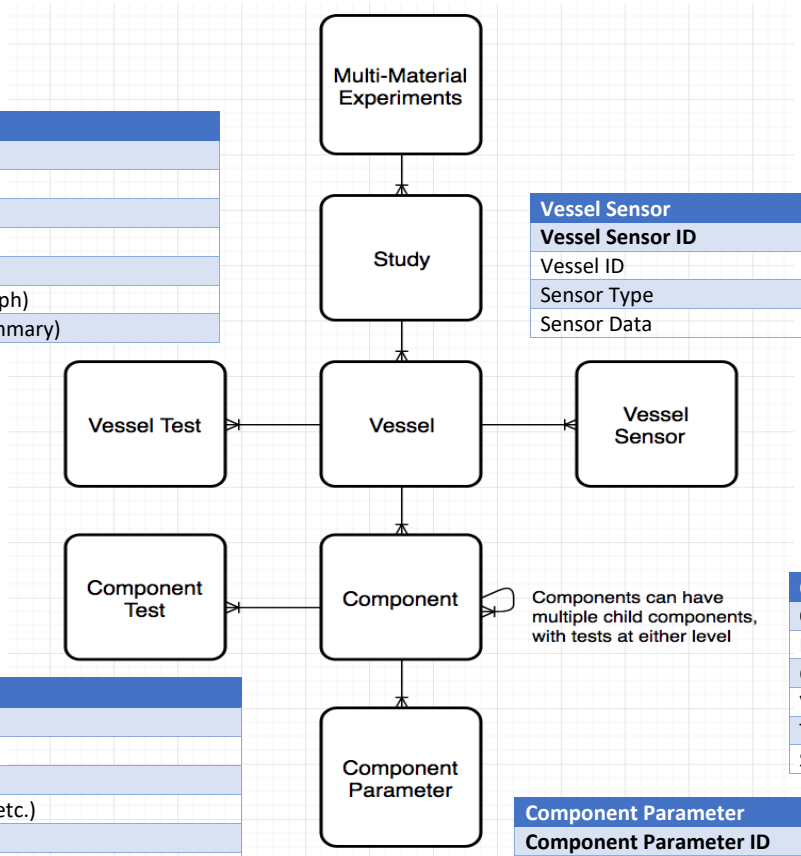
Case study: Multi-material experiments

- **Summary:** Various explosives, polymers, adhesives, etc. are placed in a vessel and then heated over a period of time to simulate aging. At different time intervals, gas is sampled from all of the vessels, and a subset of the vessels are disassembled and separated into their individual components. Additional tests are performed on each component.



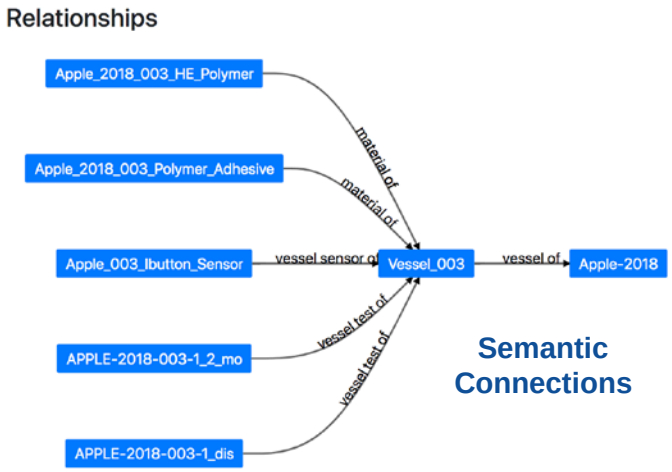
Description	t0	6 mo	12 mo
Isothermal (70C)		2	2
Ambient (23C)	1		2
Hardware only (70C)			2

Relational view implemented using MarkLogic's multi-model and semantic triples



Vessel Test
Vessel Test ID
Vessel ID
Test Date
Time Aged
Test Description
Test Data (Chromatograph)
Test Data (Chemical Summary)

Vessel Sensor
Vessel Sensor ID
Vessel ID
Sensor Type
Sensor Data

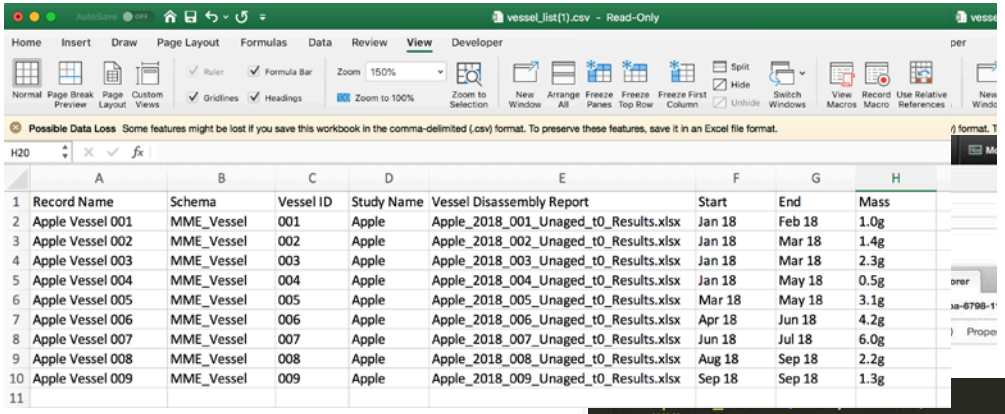


Component
Component ID
Name
Component Parent ID
Vessel ID
Type (pellet, jig, wafer, puck, etc.)
Serial Number

Component Test
Component Test ID
Component ID
Test Date
Test Type (DSC, TGA, etc.)
Test Description
Test Data (varies based on test type)

Component Parameter
Component Parameter ID
Component ID
Name
Type (width, depth, mass, etc.)
Units
Value

Metadata and data relationships can be uploaded in batch

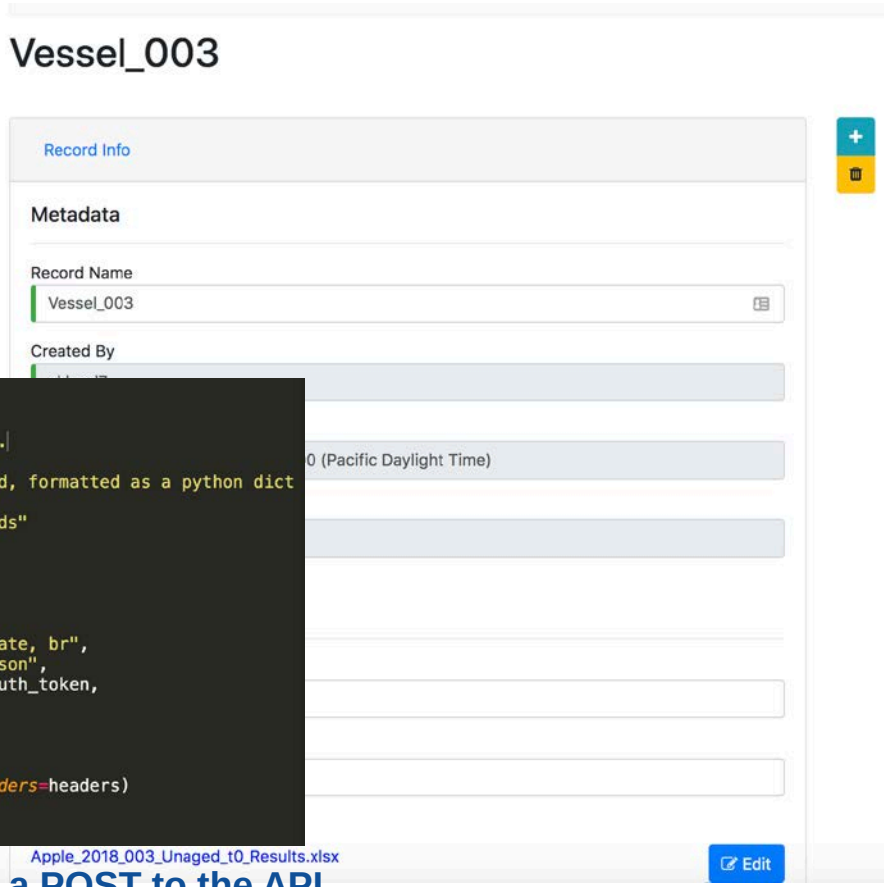


	A	B	C	D	E	F	G	H
1	Record Name	Schema	Vessel ID	Study Name	Vessel Disassembly Report	Start	End	Mass
2	Apple Vessel 001	MME_Vessel	001	Apple	Apple_2018_001_Unaged_t0_Results.xlsx	Jan 18	Feb 18	1.0g
3	Apple Vessel 002	MME_Vessel	002	Apple	Apple_2018_002_Unaged_t0_Results.xlsx	Jan 18	Mar 18	1.4g
4	Apple Vessel 003	MME_Vessel	003	Apple	Apple_2018_003_Unaged_t0_Results.xlsx	Jan 18	Mar 18	2.3g
5	Apple Vessel 004	MME_Vessel	004	Apple	Apple_2018_004_Unaged_t0_Results.xlsx	Jan 18	May 18	0.5g
6	Apple Vessel 005	MME_Vessel	005	Apple	Apple_2018_005_Unaged_t0_Results.xlsx	Mar 18	May 18	3.1g
7	Apple Vessel 006	MME_Vessel	006	Apple	Apple_2018_006_Unaged_t0_Results.xlsx	Apr 18	Jun 18	4.2g
8	Apple Vessel 007	MME_Vessel	007	Apple	Apple_2018_007_Unaged_t0_Results.xlsx	Jun 18	Jul 18	6.0g
9	Apple Vessel 008	MME_Vessel	008	Apple	Apple_2018_008_Unaged_t0_Results.xlsx	Aug 18	Sep 18	2.2g
10	Apple Vessel 009	MME_Vessel	009	Apple	Apple_2018_009_Unaged_t0_Results.xlsx	Sep 18	Sep 18	1.3g
11								

To ba
records
values
if

Cre

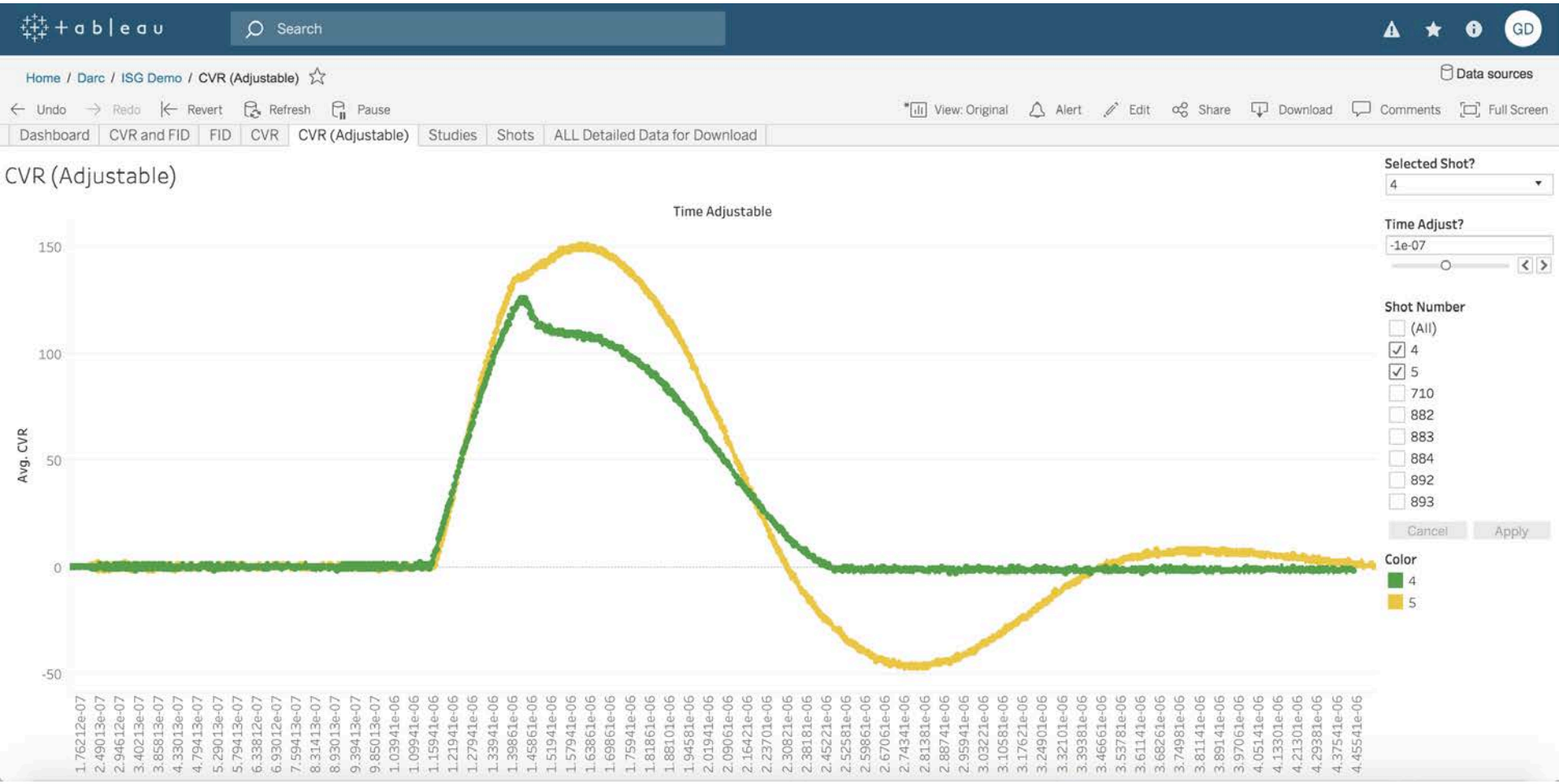
```
"""
Upload a record using our Node API.

:param record: the record to upload, formatted as a python dict
"""
url = "http://localhost:5   00/records"

auth_token = self.authent  cate()
payload = json.dumps(record)
headers = {
    'Accept': "application/json",
    'Accept-Encoding': "gzip, deflate, br",
    'Content-Type': "application/json",
    'Authorization': "Bearer " + auth_token,
    'cache-control': "no-cache",
}

response = requests.request(
    "POST", url, data=payload, headers=headers)

print(response.text)
```



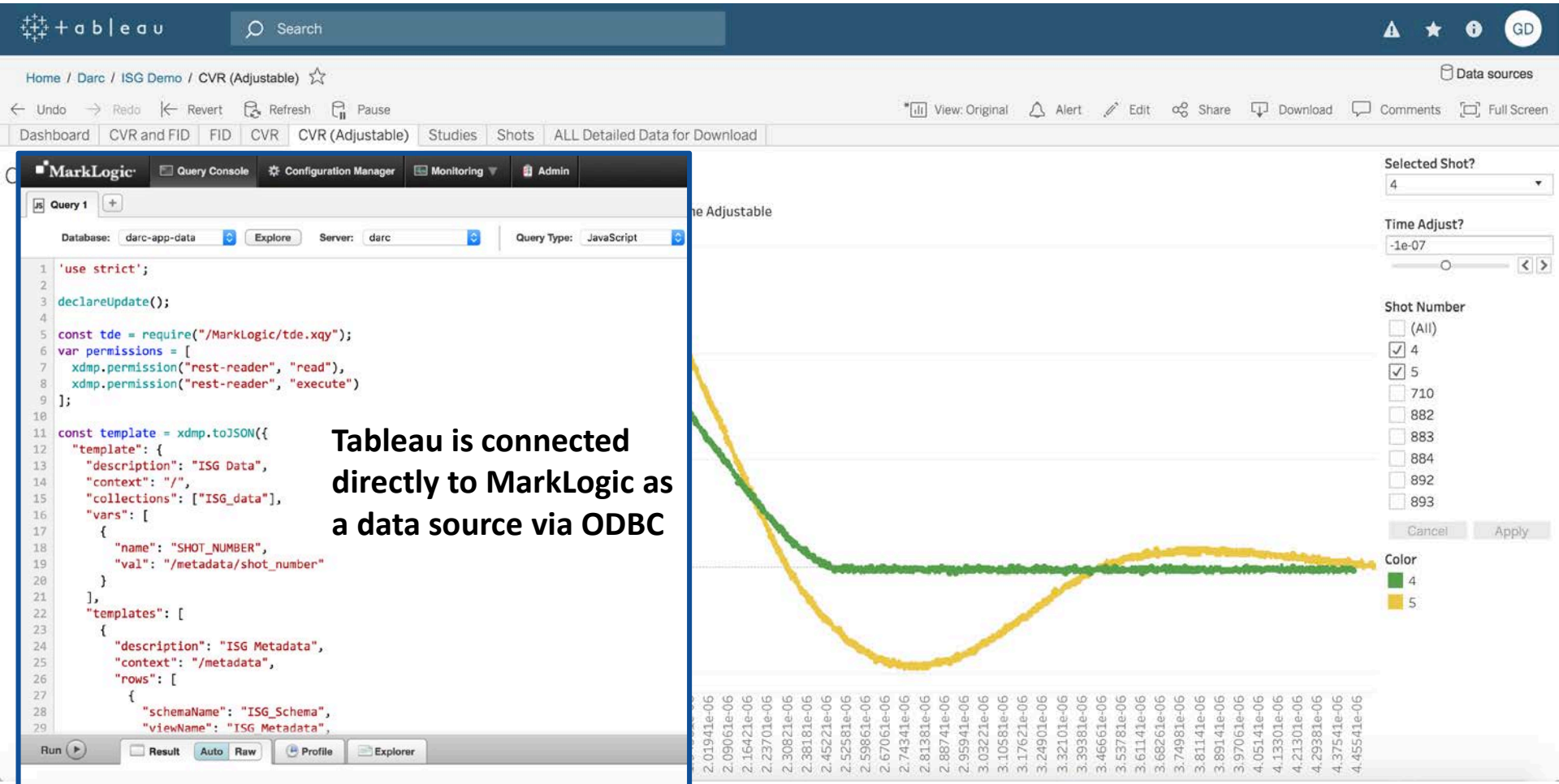
API-upload-code executes a POST to the API.
Standard python code is used to turn the csv docs into JSON for MarkLogic.

The records can be searched for and viewed through the front end UI (with link to related Excel report)

Connecting Tableau to MarkLogic provides an instant UI with faceting and filters



MarkLogic's Query Console enables Template Driven Extraction



In Summary

- Although we are still maturing some of these concepts, we are extremely excited about the potential
- We will continue to expand the framework using the MarkLogic Data Hub to create a single portal into our LEP data

Anxious to come back next year to share our progress!!