

# ~~Archiving~~ Strategies for Synchrotron Data

Presented by Kevin D. Jones, Ph.D.

April 15, 2019

NASPI - San Diego, CA

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

# OUTLINE

# Language Matters

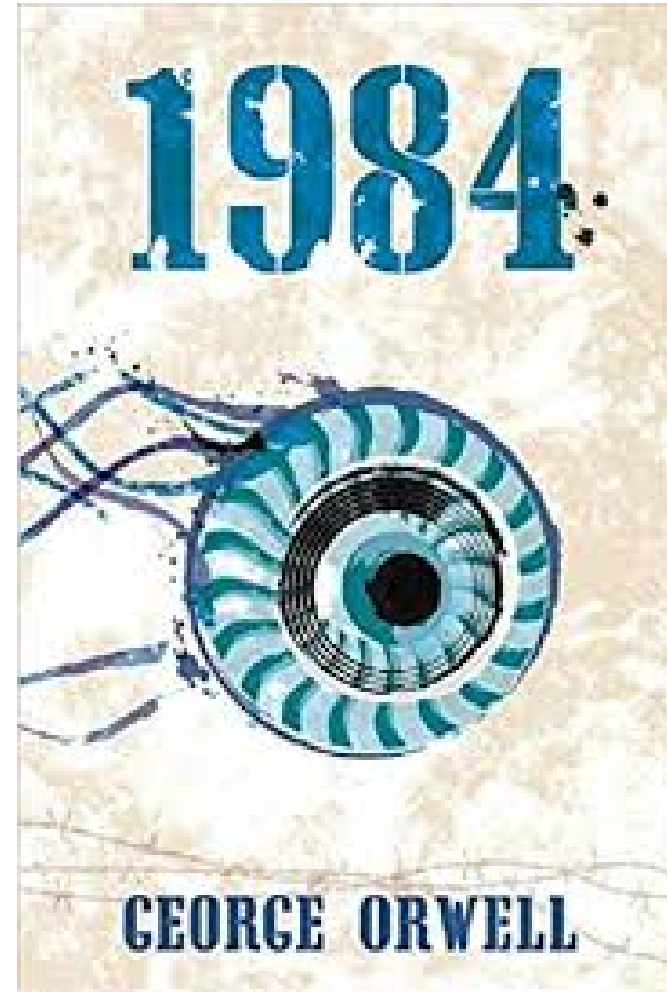
1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching



# Language Matters

“Archive”, “Historian” implies...

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching



# Language Matters

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

“Storage” - OK

“Database” – OK

How can we expose our underlying requirements through our language?

Data-as-a-Service

Data-as-a-Utility

# Keep is the Answer!

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

*Inclusion:*

Measure Everything; Keep Everything

*Retention:*

Keep Everything, Forever

*Compression:*

**NO LOSSY COMPRESSION**

# No Value Prejudgments!

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

Who are you to decide  
whether data will be valuable  
in the future or not?

We are only stewards  
of the data.

# No Value Prejudgments!

## *Events vs Ambient Data*

Define “Event” – Not easy, huh?

Events are not data, they are metadata - a pass-by-reference link to the data source.

“Ambient” data is still information rich

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching



# Systems and Architecture

*...or, finding things when we need them*

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching



# Systems and Architecture

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

We're often so concerned with just *getting the data in* that we don't consider the future uses requiring *processing the data* or *getting the data out*.

This myopic assumption leads to poor decisions about tools, systems, and architectures.

# 1. Word Jumble You Must Look at Your Data

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

## *Data Quality Matters*

It's wasteful to store garbage data but more importantly...



You only get one opportunity to capture *that* moment at *that* location in *this* universe.

1. Word Jumble

2. To Keep or  
Not to Keep

3. The Anti-Kon-  
Mari Method

4. The Hall Closet

5. Data Binge  
Watching

**QUESTIONS?**

[kevin.d.jones@dominionenergy.com](mailto:kevin.d.jones@dominionenergy.com)