
Passion for Innovation.
Compassion for Patients.™



New Jersey Alteryx User Group

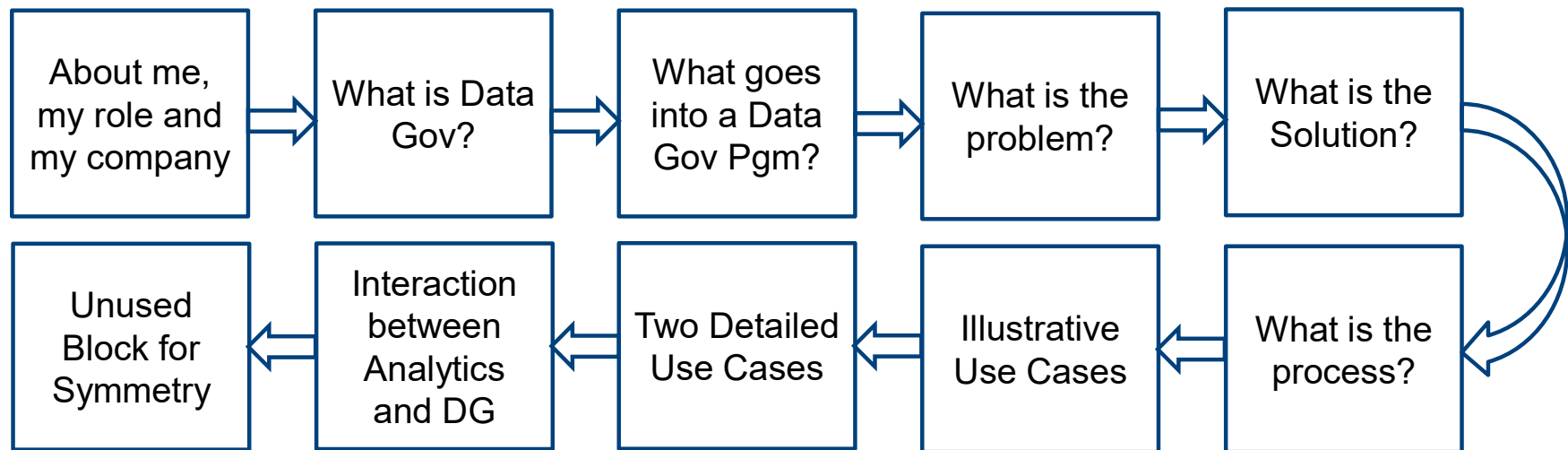
Using Analytics as the Tip of the Spear for Data Governance

December. 16th, 2021

Hayes Williams
R&D Affairs

The views and opinions expressed in this presentation are those of the author and are not intended to reflect the views and/or opinions of Daiichi Sankyo, Inc.

Agenda



Spoiler: Using Alteryx to solve Adhoc data analytics problems for the business generates good will for your analytics program and gives you a direct line into clear Data Governance issues that would have not naturally been recognized or brought forward

Me

Work Background

Consultant long time both big and small firms

Themes: Data and R&D

Celgene – Data Governance for Clin Ops

Daiichi Sankyo – Data Governance for R&D

Alteryx Advanced Certified (expired ☹)

[linkedin.com/in/hayeswilliams/](https://www.linkedin.com/in/hayeswilliams/)





- Daiichi Sankyo is a global pharmaceutical company
- Ranked 24th among top global pharmaceutical companies*
- Worldwide presence
 - Ground presence in 24 countries
 - Manufacturing locations in 6 countries
 - R&D locations in key areas
- In addition to our current portfolio of medicines for cancer and cardiovascular disease, Daiichi Sankyo is primarily focused on developing novel therapies for people with cancer as well as other diseases with high unmet medical needs. With more than 100 years of scientific expertise and a presence in more than 20 countries, Daiichi Sankyo and its 16,000 employees around the world draw upon a rich legacy of innovation to realize our 2030 Vision to become an “Innovative Global Healthcare Company Contributing to the Sustainable Development of Society.”

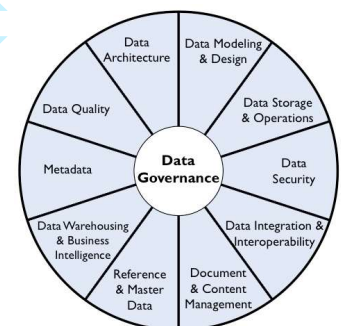
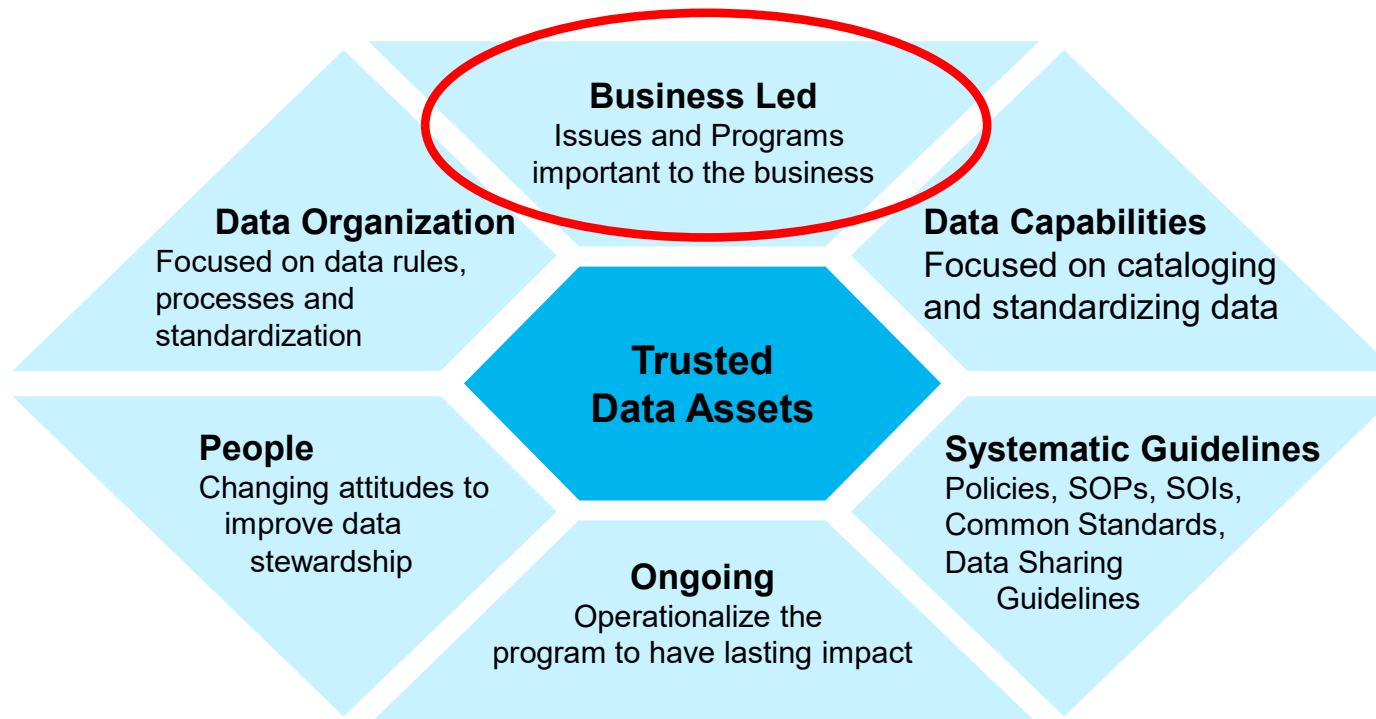
Good Data Governance leads to:

- which leads to*
better business outcomes



Data Governance program deliverables

The R&D Data Governance program delivers trust in the data to improve business decision making, analytics, usage and by extension business results through the following key areas:



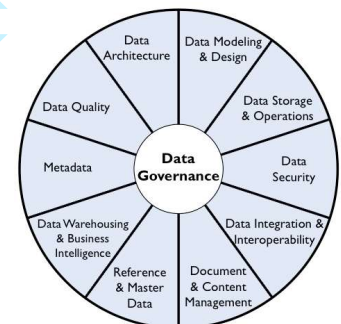
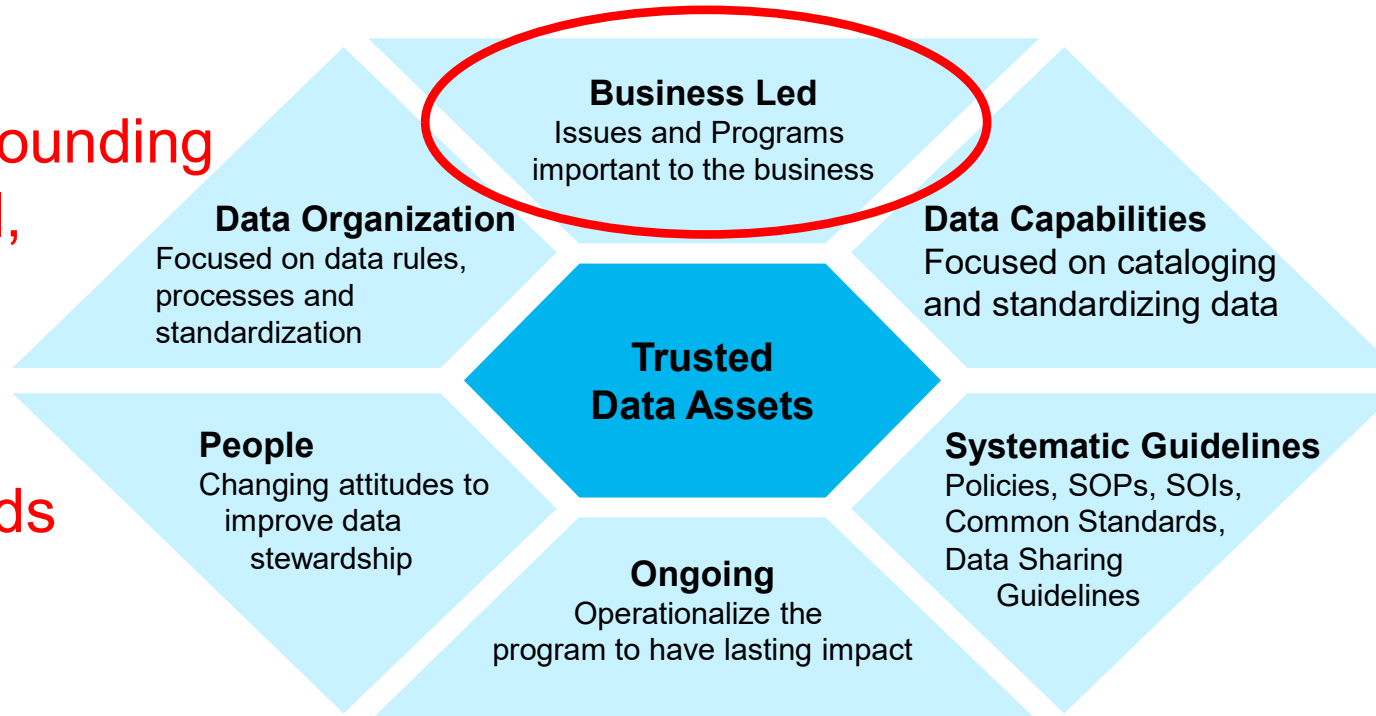
Copyright © 2017 DAMA International

Data Governance program deliverables

The R&D Data Governance program delivers trust in the data to improve business decision making, analytics, usage and by extension business results through the following key areas:

Ends up sounding theoretical, right?

Which leads to this....



Copyright © 2017 DAMA International

The Problem



Me: What are your business issues which a Data Governance Organization can help solve?

Business: Sure, I will tell you.
What are Data Governance Issues?
(aka what Business Problems are appropriate?)

The Solution



Offer to solve their “spreadsheet” or “data-from-two-systems” adhoc analytical problems for them.

➔ The Data Governance issues show themselves as a bonus.

Simple process: Go to a friendly executive and offer to help with problems that sound like this:

- ❑ Data in disparate sources (multiple spreadsheets, sources) and needs to be combined and cleansed to be suitable for insight
- ❑ Need is adhoc, temporarily repeating (e.g., 6 months) or not general (specific study or function) or has unknown or rapidly changing requirements
- ❑ Resource effort will be notable and error prone with traditional Excel manipulations

Conditions which make the issue unsuitable for an IT implementation so IT may have already said “no”

This description, they understand.

Use Cases

Illustrative list of Use Cases when I tried this:

#	Status	Description	Function
1	Complete	Aggregation of study level data from Excel spreadsheets	Program Management
2	Complete	Aggregation of operational study data from multiple spreadsheets to calculate Cycle Times	Trial Feasibility
3	Complete	Site Initiation overlap and potential indication level conflicts from multiple site sources and indication info	Trial Feasibility
4	Complete	Review categorizations of Protocol Deviations for consistency and Insight	Clinical Process Mgmt
5	In Progress	Review ICFs for completeness, type and version	BioSample Operations
6	In Progress	CRO Site Contracts Investigation	Operations Excellence
7	In Progress	Conversion activities for CTMS Systems Integration	Clinical Operations

Analytics Example 3

Site Initiation: What is the site overlap between our Partner and DS?

Site Lists as Input

DSI (CTMS Extract)

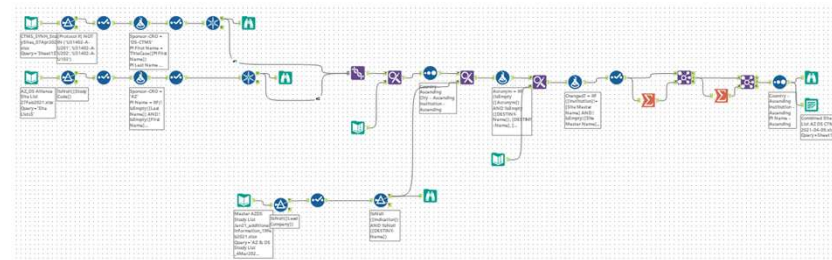
Combined Site List AZ DS CTMS 2021-04-09

Partner

AZ_DS Alliance Site List 27Feb2021

Master AZDS Study List Jan21_additional information_19Feb2021

Standardize and Prep Data



- ☐ Standardize Column names
- ☐ Standardize Country lists
- ☐ Enrich with Study Descriptive
 - Study names which look like IDs should add Indication "Lung"
- ☐ Incorporate manually mastered site name changes
- ☐ Perform counts to identify overlaps
- ☐ Create Manually Mastered Site List to standardize Site names, e.g., 'Aphm - Hopital Nord' changed to 'Hopital Nord'

Insight

Sites in countries used by both entities which may be causing conflict

Country	Total Site Overlap	Study 1 Total Sites	Study 2 Total Sites
Country 1	12	22	18
Country 2	11	18	12
Country 3	8	17	14
Country 4	7	5	8
Country 5	5	5	18
Country 6	4	28	42

Analytics Insight: Site Initiation Potential Overlap



Potential overlap in Study 1 and Study 2 targets between organizations.

Helps the Business

Country	Site Name	Overlap Institution	Overlap PI
Canada	Institution 1	Yes	No
Canada	Institution 2	Yes	Yes
Canada	Institution 3	Yes	Yes
France	Institution 4	Yes	No
France	Institution 5	Yes	Yes
France	Institution 6	Yes	Yes
France	Institution 7	Yes	Yes
France	Institution 8	Yes	Yes
France	Institution 9	Yes	Yes
France	Institution 10	Yes	Yes
France	Institution 11	Yes	Yes
Germany	Institution 12	Yes	Yes
Germany	Institution 13	Yes	Yes
Spain	Institution 14	Yes	Yes
Spain	Institution 15	Yes	Yes
Spain	Institution 16	Yes	Yes
United Kingdom	Institution 17	Yes	No
United Kingdom	Institution 18	Yes	Yes

Country	Total Site Overlap	Study 1 Total Sites	Study 2 Total Sites
Country 1	12	22	18
Country 2	11	18	12
Country 3	8	17	14
Country 4	7	5	8
Country 5	5	5	18
Country 6	4	28	42
Country 7	3	4	4
Country 8	3	10	11
Country 9	3	8	20
Country 10	2	11	14
Country 11	2	8	4
Country 12	2	8	7
Country 13	1	6	4
Country 14	0	6	0
Country 15	0	5	0
Country 16	0	0	11
Country 17	0	4	0
Country 18	0	0	9
Country 19	0	8	7
Country 20	0	17	0
Country 21	0	16	13
Country 22	0	11	0
Country 23	0	11	0
Country 24	0	6	0
Country 25	0	5	0
Country 26	0	10	0
Country 27	0	12	5
Country 28	0	6	0
Country 29	0	7	0
Country 30	0	7	0

Data Governance Insights

- ❑ Enrichment: Add descriptive categorizations like “Lung”
- ❑ Country Standardization – Reference Data Management

CountryFrom	CountryTo
Russian Federation	Russia
Korea, Republic of	South Korea
Taiwan, Republic of China	Taiwan
USA	United States
United States of America	United States

Helps the Data
Governance Program

- ❑ Site Standardization – Master Data Management

Site Name	Site Name - Standard
"Medical Technologies" Ltd.	Medical Technologies Ltd.
1st Affiliated Hospital of Zhejiang University	The First Affiliated Hospital of College of Medicine, Zhejiang University
AOUP - U.O. Oncologia Medica 2 Univ.	Azienda Ospedaliero Universitaria Pisana
Antwerp University Hospital	Universitair Ziekenhuis Antwerpen
Azienda Ospedaliera Arcispedale Santa Maria Nuova - IRCCS	Arcispedale Santa Maria Nuova
Azienda Ospedaliera Universitaria Policlinico G Rodolico San Marco	Azienda Ospedaliero Universitaria Policlinico G.Rodolico - San Marco
Azienda Ospedaliera Universitaria Policlinico-OVE	Azienda Ospedaliero Universitaria Policlinico G.Rodolico - San Marco
Azienda Ospedaliero- Universitaria Policlinico S. Orsola-Malpighi	Azienda ospedaliero-universitaria Policlinico Sant'Orsola Malpighi)
Azienda Ospedaliero-Universitaria di Modena	Policlinico di Modena Azienda Ospedaliero-Universitaria di Modena, Via del Pozzo, 71, 41125

Analytics Example 4

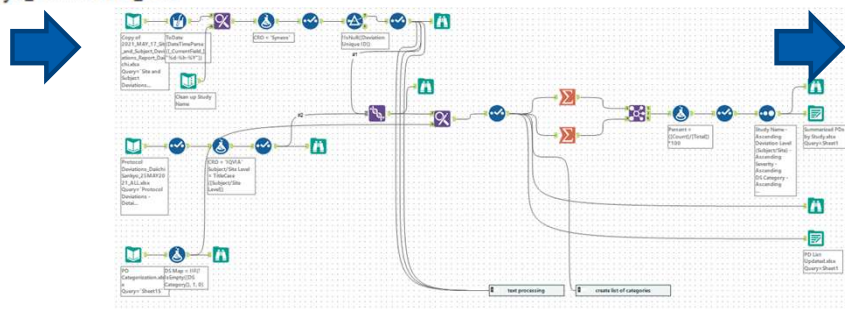
Protocol Deviations: Studies to date have categorized protocol deviations using different categorization schemes. What insight can be drawn from an aggregated protocol deviation file across studies and CROs?

PDs from CRO1 & CRO2

Aggregate, Filter and Cleanse Data

Copy of 2021_MAY_17_Site_and_Subject_Deviations_Report_Daiichi

Protocol Deviations_Daiichi Sankyo_25MAY2021_ALL



Aggregated file

Insight

- ❑ Categories are wildly different between studies preventing detailed comparison and trending
- ❑ New consistent categorization needed to be rolled out to support analytics and insight

- ❑ Standardize Column names
- ❑ Integrate manually created mapping list for PD Categories
- ❑ Summary statistics by category

- ❑ Task is initially 1-time to investigate and providing a data-focused reason to change

Analytics Example 4 – Protocol Deviation Results

- ❑ Protocol Deviations reported between October 2016 to April 2021 between active or completed studies placed with CRO1 and CRO2 were analyzed
- ❑ Manual mapping to the DS standard PD categories was required in order to normalize the categorization
- ❑ Conclusions
 - ❑ Individual Studies following SOPs explicitly and assigning categories for PDs as needed
 - ❑ Different choices by studies impede aggregate analysis and insights to find program level and CRO level trends
 - ❑ **Standardized categorization structure needed**

Drive agreement to a standardized categorization structure

Category	Sub-Category	Deviation Code	Important/Non-Important
	Prohibited non-drugtherapy administered/taken	CM03	Important/Non-Important
	Concomitant non- drug therapy not used as per protocol	CM04	Important/Non-Important
Study Procedures	Did not follow withdrawal criteriaas per protocol	SP01	Important/Non-Important
	Visits, assessments, or procedures not performed per protocol	SP02	Important/Non-Important
	IMP Management and Documentation	SP03	Important/Non-Important
	(Non-IP related)randomization procedure not conducted as perprotocol	SP04	Important/Non-Important

Dramatically different names across existing studies for proposed unified category “Study Procedures”

Study Procedures
Study Procedures
Visit/Procedure Required
Laboratory Assessment
Visit Schedule
Study Procedures_SP02
Laboratory
Study Procedures_SP05
Study Procedures/Visits, assessments, or procedures not performed per protocol
Study Procedures_SP01
Study Procedures_SP06
Study Procedures_SP10
Study Procedures_SP11
Study Procedure (non-subject)
Randomisation
Study Procedures/(Non-IP related) randomization procedure not conducted as per protocol
Study Procedures_SP04
Study Procedures/IP compliance not checked per protocol
Study Procedures_SP03
Randomization (non-subject)
Study Procedures/Did not follow withdrawal criteria per protocol
Study Procedures_SP07
Study Procedures_SP09

Illustrative Interaction Between Adhoc Analytics and Data Governance Issues

Week	1	2	3	4	5	6	7	8	9	10	11	12
Adhoc Analytics Problems												
Use Case 1	I1	I2							Small, fast, adhoc analytics problems with weekly turn-around times			
Use Case 2		I1	I2	I3	I4				Quick iterative analytics requirements definition to feed longer time scale data governance process			
Use Case 3				I1	I2			I3	I4	Pick up and put down quickly as the business needs or learns more		
Use Case 4	I1	I2			M				M	Monthly “maintenance” runs to update the adhoc analysis as needed		
Data Governance Problems												
Project 1								Data Governance Issue 1				
Project 2	Data Governance Issue 2											



Solving Adhoc Analytics issues lead to uncovering Data Governance Issues

Summary



- ❑ Starting a Data Governance Program and not getting the business-led issues you know are there?
- ❑ Offer to solve Adhoc data analytics problems for the business
 - ❑ Generates good will for your analytics program
 - ❑ Gives you a direct line into clear Data Governance issues that would have not naturally been recognized or brought forward

Thank You

