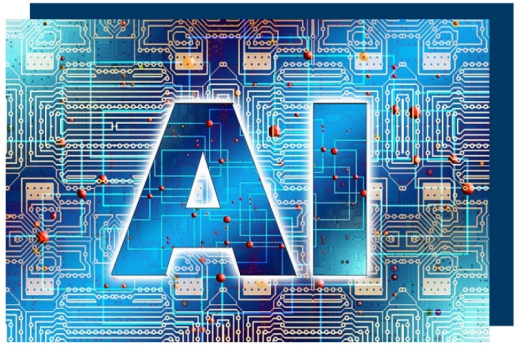
ISO/IEC *AI workshop sessions*
on 29 and 30 November



ISO/IEC 24668

Process Management Framework - Big Data Analytics

GAUTAM BANERJEE

BUSINESS BRIO

Session 2
30th November 2022

ENTERPRISE TECH

The Data Science Struggle: Why Organizations Routinely Fail To Realize The Full Potential From Their Data Science Efforts

Hannah M. Mayer Contributor
Hannah is a former HBS & HSG researcher, and a Wiley-

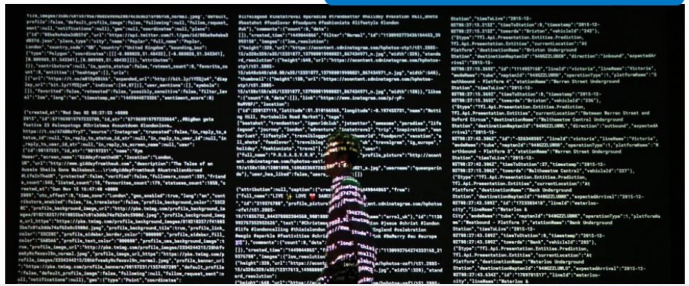
Follow

Nov 7, 2022, 03:30am EST

0

New!
Click on the conversation bubble to join the conversation
Got it!

By Hannah M. Mayer, Luca DeStefano



OPINION

Data-driven decision making will fail – and here is why

Marc Warner, CEO at Faculty, contends that corporate organisations need to go through a cycle similar to the development of scientific thought – from divination, via data deluge to theoretical understanding

By **Marc Warner**

Published: **07 Nov 2022**

Most people have heard of data-driven decision making – the admirable desire to make better decisions through data. The conventional justification is that given the world's complexity, organisations need data to make decisions.

Such ideas exist across organisations from governments to companies to hospitals. In fact, they are wrong in principle and misleading in practice.

Data-driven decision making puts the emphasis in the wrong place, and means organisations focus on the wrong thing. Let's look at why.

Latest News

British broadband customers blast proposed mid-contract bill rises

How to Prevent Big Data Analytics Failures

December 18, 2015

Contributor: Susan Moore

Big data analytics projects don't fail for a single reason, nor due to technology alone.

CIO DIVE Deep Dive Opinion Library Events Topics

4 reasons why most data science projects fail

Published May 2, 2017

By James Roberts

Pixabay

The following is a guest article from James Roberts, chief data scientist

"To succeed, you must develop a viable strategy to deliver business value from a big data initiative. Then map out and acquire or develop the missing and specialized skills that are needed. Once strategy and skill priorities are addressed, then you can move on to big data analytics."

Gartner research director **Svetlana Sicular**

https://www.gartner.com/smarterwithgartner/how-to-prevent-big-data-analytics-failures

Standards

Need

Sharing; Re-use; Consent; Collaborate; Interoperability; Adoption; Trust

Forms

Technology ; Frameworks; Methodology; Use Cases

Means

Published papers; Patents; Standards; Certification; Regulation; Legislation

Source

Forums; Bodies; National; International; Unions and groups of countries

4 Standards

From ISO/IEC SC 42

B 5259 Series

Data Quality

C 5338

AI LifeCycle Methodology

C 38507

Governance implications for AI applications

A 24468

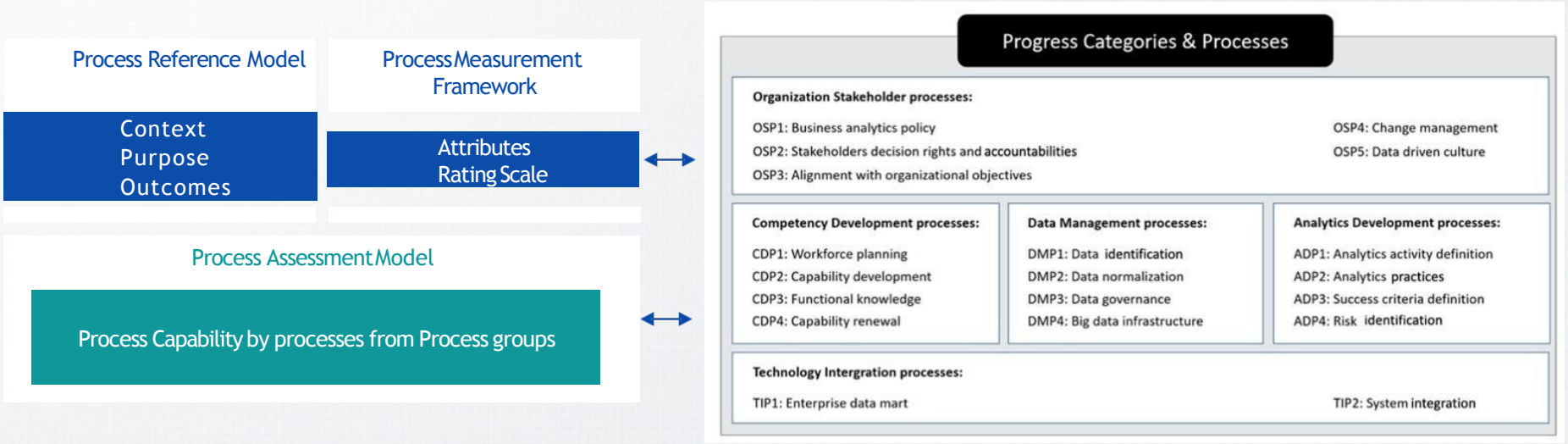
Process Management Framework -Big DataAnalytics

NWIP | WD | CD | DIS | FDIS | IS

24668

Process Management Framework - Big Data Analytics

The proposed standard provides a framework for developing processes to effectively leverage big data analytics across the organization irrespective of the industries/sectors. It specifies process management for big data analytics with the five process categories as listed below with help of Process Reference Model, Measurement Framework and Assessment Model:



Sample process Reference Model

A sample of Data Identification process (refer DMP1 from previous figure) from Data Management Process group is given below for understanding

ID	DMP1
Name	Data identification
Context	This process covers the key step of identifying data elements. This involves identification of data elements so that the team will not lose sight of any important dimension or factor that plays a key role in the outcome analysis.
Purpose	The purpose of DMP1 process is to identify, define, classify, and collect data for all data elements available for the information flow in the context of the project or department or function.
Outcomes	The outcomes of this process include: a) data elements relevant to the process or function or department are identified; b) data elements can be classified into categories, such as unstructured, transactional, hierarchical, reference; c) meta data (operational definition) such as units, frequency, source or sources, functional definition, range, possible functional correlation, producer or consumer or ownership or steward is derived; d) new data sources are identified; e) data collection is performed.

Sample process Base Practices and Information Products

A sample structure for defined base practices for Data Identification process is given below

ID	DMP1
Base practices	BP1 identify the key data elements: such as features, co-variates, factors [outcome (a,d)]. BP2 classify the data elements: [outcome (b)]. BP3 provide the operational definition of data elements: [outcome (c, d)]. BP4 collect the data with rigor and compliance: [outcome (e)].

A sample structure of Information Products (IPs) for Data Identification process is given below

Information products					
Inputs			Outputs		
IP ID	Name	Outcome	IP ID	Name	Outcome
IP_27	SIPOC	a), b)	IP_32	Data collection plan	a), c), d), e)
IP_28	Process flow	b), c)	IP_33	Data operational definition	a), b), c), d), e)
IP_29	CTQ matrix	a), b), c)	IP_34	Data sample	e)
IP_30	Use case studies	c), d)	-	-	-
IP_31	Domain documents	c), d)	-	-	-

Assessment

For each process, the process assessment model further provides:

- A set of base practices (BPs) defining the tasks and activities needed to accomplish the process purpose and fulfil the process outcomes; each base practice is explicitly associated to a process outcome.
- Number of information products (IPs) associated with each process and related to one or more of its outcomes.
- Characteristics associated with each information products.

The base practices and information products constitute the set of indicators to assess process performance for a particular process. The capability levels of each process area can be determined accordingly and then action items can be drafted to address the gaps.

More importantly, the assessment should lead to a roadmap to implement Data Science/AI or Big Data Analytics in terms of use cases for different functions.

Thank You

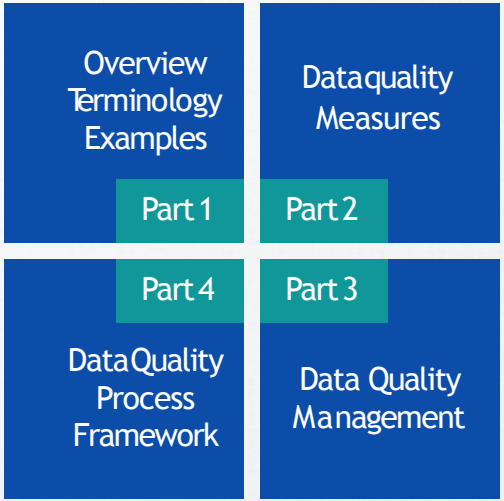
Gautam Banerjee

https://www.linkedin.com/in/gautambanerjee/

Appendix

5259 Series

Data Quality



MEASURES

- Accuracy
- Precision
- Uniqueness
- Completeness
- Consistency

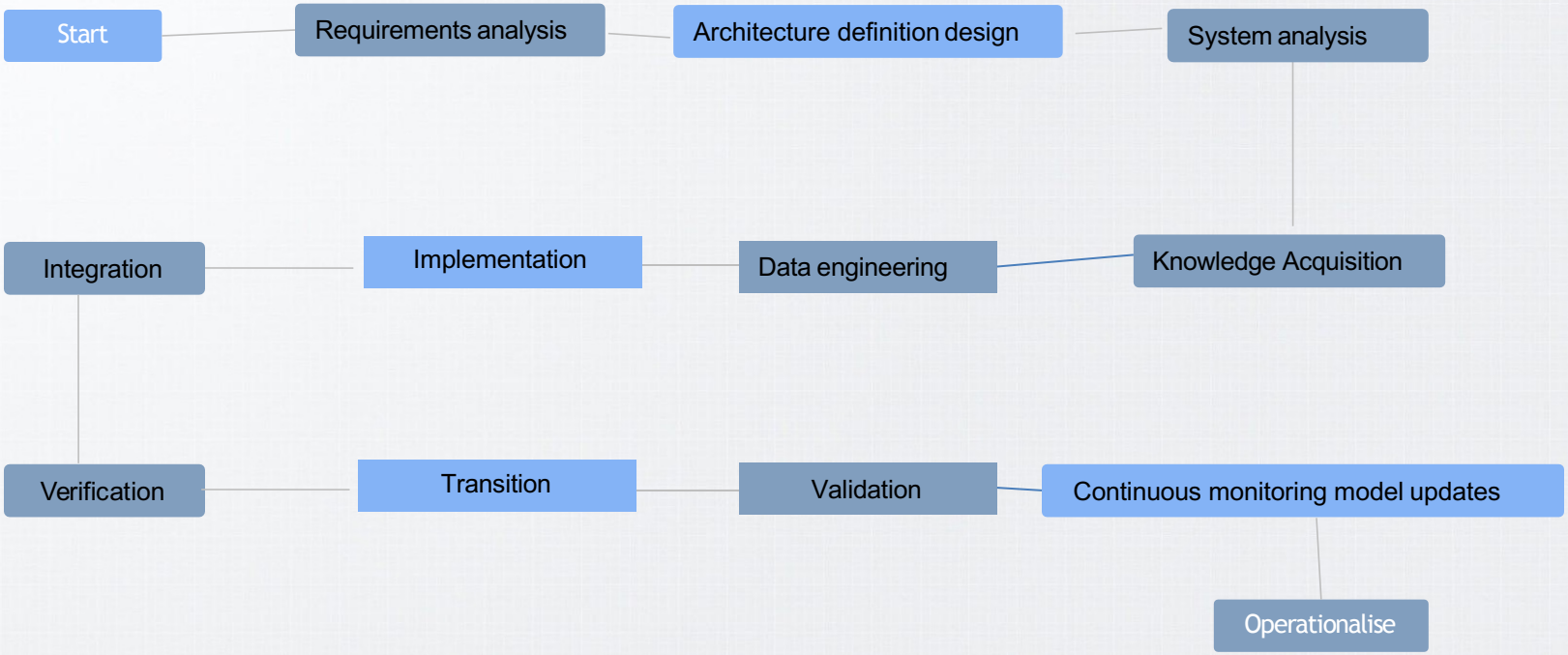
CHALLENGES

- Structured Data vs BigData
- ML vs AI
- Part 4 vs Data Governance

* These tenets are WIP. They may not be accurate and meant to give only a high level idea

5338

AI Lifecycle Methodology



* These tenets are WIP. They may not be accurate and meant to give only a high level idea

38507

Governance implications for AI applications

Decision Oversight

Data Use

Culture & Values

Compliance & Risk

* These tenets are WIP. They may not be accurate and meant to give only a high level idea