

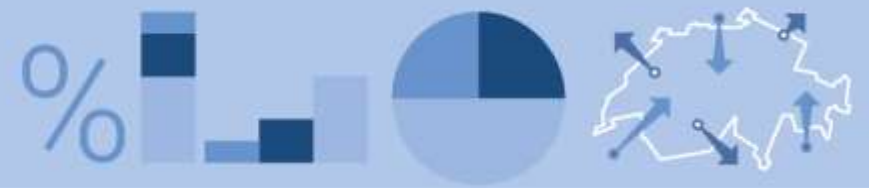
Are Current Frameworks in the Official Statistical Production Appropriate for the Usage of Big Data and Trusted Smart Statistics?

Prof. Dr. Bertrand Loison, Prof. Dr. Diego Kuonen

Vice-Director at Federal Statistical Office, CEO Statoo Consulting & Professor at UNIGE

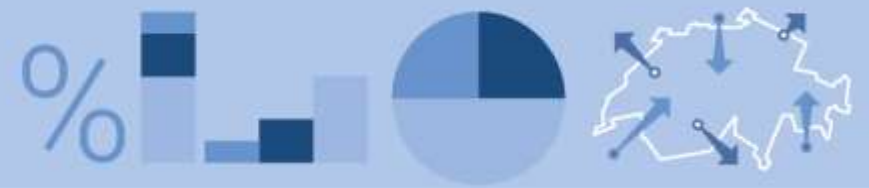
Directors General of the National Statistical Institutes (DGINS)

Bucharest, 10 - 11 October 2018



Agenda

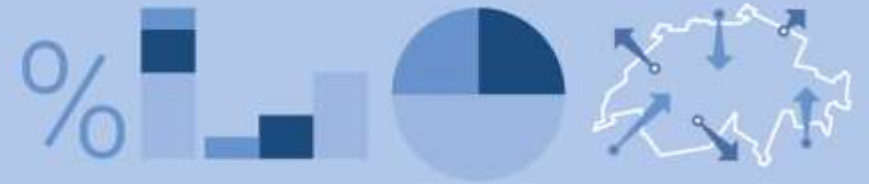
1. Introduction
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
4. **“Analytics of things” and “smart (official) statistics”**
5. Demystifying the two approaches of analytics
6. **Process models for continuous improvement**
7. Conclusion



Agenda

- 1. Introduction**
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
4. “Analytics of things” and “smart (official) statistics”
5. Demystifying the two approaches of analytics
6. Process models for continuous improvement
7. Conclusion



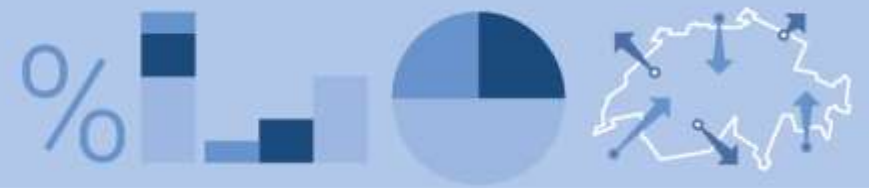


Introduction

“Just as haute cuisine must incessantly reinvent itself in order to stay at the forefront of gastronomy, official statistics is also confronted with a rapidly changing context and needs.”

Walter J. Radermacher

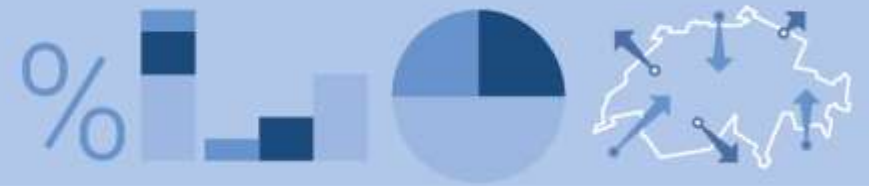
Radermacher, W. J. (2018). Official statistics in the era of big data opportunities and threats. International Journal of Data Science and Analytics (<https://doi.org/10.1007/s41060-018-0124-z33>).



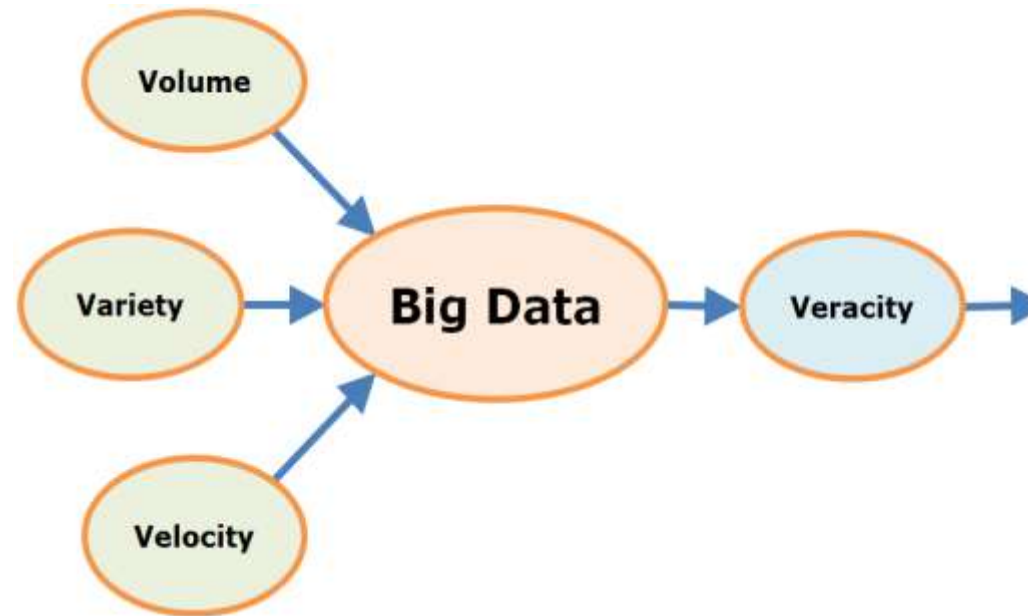
Agenda

1. Introduction
- 2. Demystifying the “big data” hype**
3. Demystifying the “Internet of things” hype
4. “Analytics of things” and “smart (official) statistics”
5. Demystifying the two approaches of analytics
6. Process models for continuous improvement
7. Conclusion

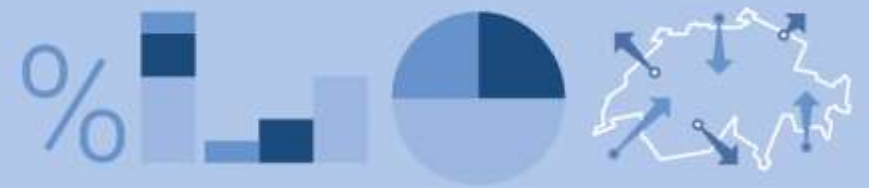




Demystifying the “big data” hype



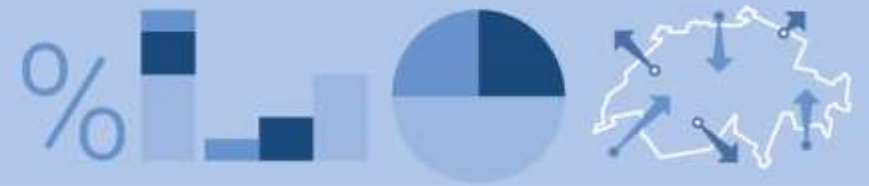
Kuonen D. (2017). Glocalised Smart Statistics and Analytics of Things: Core Challenges and Key Issues for Smart (Official) Statistics at the Edge. Invited presentation given on July 18, 2017 within Eurostat’s special topic session within Eurostat’s special topic session ‘From Big Data to Smart Statistics’.



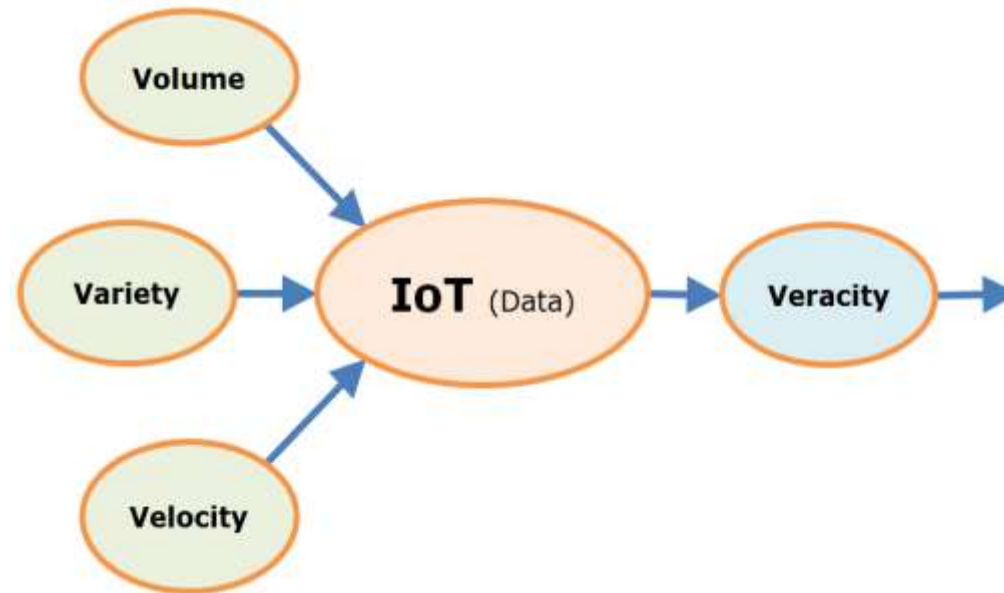
Agenda

1. Introduction
2. Demystifying the “big data” hype
- 3. Demystifying the “Internet of things” hype**
4. “Analytics of things” and “smart (official) statistics”
5. Demystifying the two approaches of analytics
6. Process models for continuous improvement
7. Conclusion

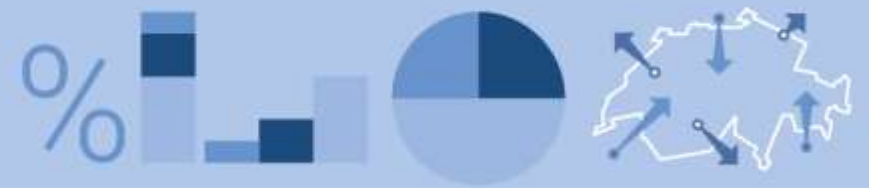




Demystifying the “Internet of things” hype



Kuonen D. (2017). Glocalised Smart Statistics and Analytics of Things: Core Challenges and Key Issues for Smart (Official) Statistics at the Edge. Invited presentation given on July 18, 2017 within Eurostat’s special topic session within Eurostat’s special topic session ‘From Big Data to Smart Statistics’.



Agenda

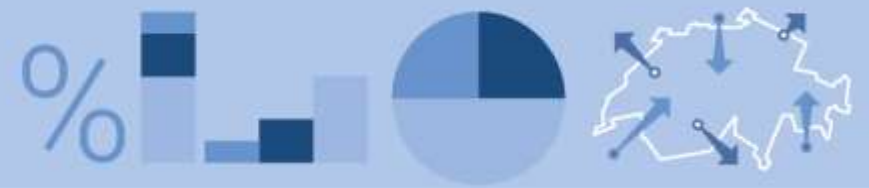
1. Introduction
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
- 4. “Analytics of things” and “smart (official) statistics”**
5. Demystifying the two approaches of analytics
6. Process models for continuous improvement
7. Conclusion





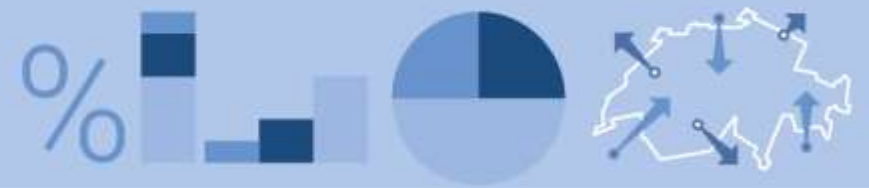
“Analytics of things” and “smart (official) statistics” - #1

1. The “Analytics of Things” (AoT) corresponds to the “analytics layer” that occurs with the IoT devices and their generated data.
2. This paradigm is referred to analytics at the edge or **edge analytics** (based on a distributed “IT architecture layer” called “edge computing”).
3. **It is about moving the analytics and the data quality frameworks to the data and not the data to the (centralised) analytics and (centralised) data quality frameworks.**
4. The usage of AoT within official statistics clearly illustrates another paradigm shift, along with related considerations of transparency and glocalisation.
5. Saying this, standardisation efforts are urgently needed. Could official statistics play a key role in their development and governance?



“Analytics of things” and “smart (official) statistics” - #2

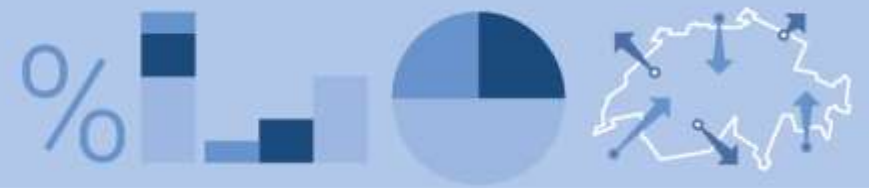
1. Eurostat defines “**smart (official) statistics**” as being the future system of producing official statistics where essentially, data capturing and data processing capabilities coupled with analytical and statistical capabilities will be **embedded** in the smart systems themselves.
2. The traditional model of pulling data in – from data sources to NSIs – **will not fit** in the new scenario.
3. Eurostat envisions a model based **on pushing computation out** – from NSIs to the data acquisition systems. This shift of focus from sources to systems lies at the core of what Eurostat call smart statistics, and corresponds AoT using edge analytics, in the context of official statistics.



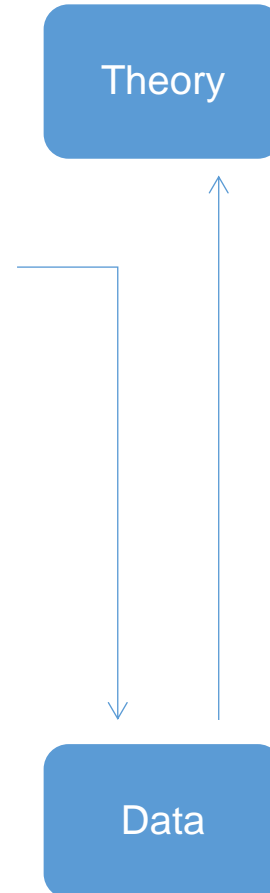
Agenda

1. Introduction
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
4. “Analytics of things” and “smart (official) statistics”
- 5. Demystifying the two approaches of analytics**
6. Process models for continuous improvement
7. Conclusion

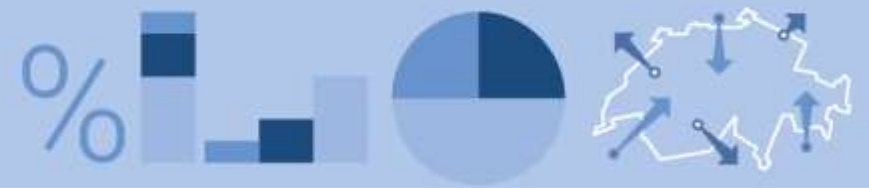




Demystifying the two approaches of analytics



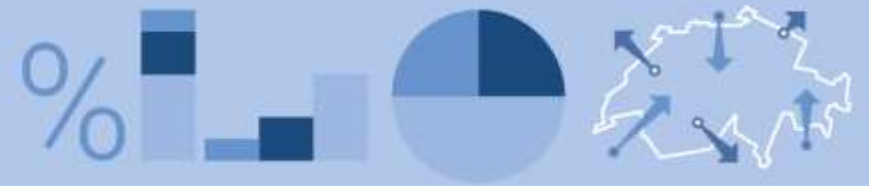
Loison B. (2018). 26th Wiesbaden Group on Business Registers, Neuchâtel, 19 – 21 September 2018



Agenda

1. Introduction
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
4. “Analytics of things” and “smart (official) statistics”
5. Demystifying the two approaches of analytics
- 6. Process models for continuous improvement**
7. Conclusion

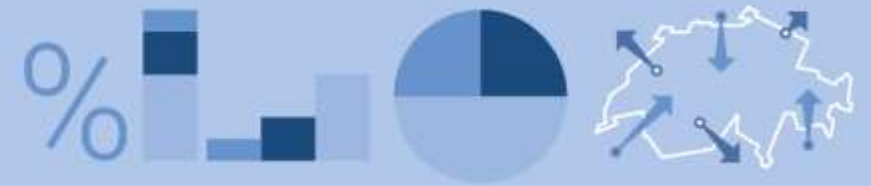




Process models for continuous improvement - #1

“If you can not describe what you are doing as a process, you do not know what you are doing.”

W. Edwards Deming



Process models for continuous improvement - #2

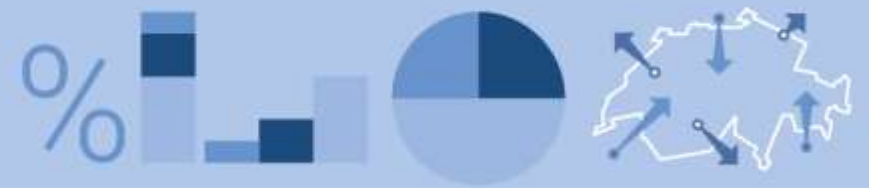
1. “Plan–Do–Check–Act” (PDCA)



2. “Generic Statistical Business Process Model” (GSBPM)



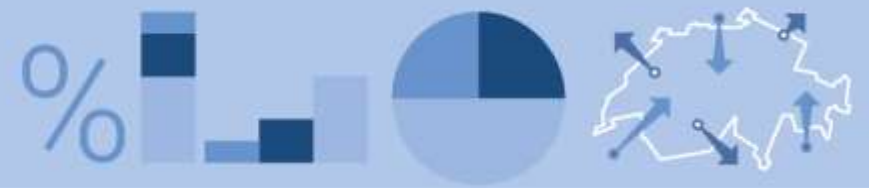
Quality Management / Metadata Management								
1	2	3	4	5	6	7	8	9
Specify	Design	Build	Implement	Process	Analyze	Evaluate	Review	Improve
1.1 Identify requirements	2.1 Design model	3.1 Build data collection infrastructure	4.1 Implement infrastructure	5.1 Operate infrastructure	6.1 Monitor data quality	7.1 Analyze data quality	8.1 Review data quality	9.1 Improve data quality
1.2 Identify data sources	2.2 Design data model	3.2 Build data collection infrastructure	4.2 Implement infrastructure	5.2 Operate infrastructure	6.2 Monitor data quality	7.2 Analyze data quality	8.2 Review data quality	9.2 Improve data quality
1.3 Identify data sources	2.3 Design data model	3.3 Build data collection infrastructure	4.3 Implement infrastructure	5.3 Operate infrastructure	6.3 Monitor data quality	7.3 Analyze data quality	8.3 Review data quality	9.3 Improve data quality
1.4 Identify data sources	2.4 Design data model	3.4 Build data collection infrastructure	4.4 Implement infrastructure	5.4 Operate infrastructure	6.4 Monitor data quality	7.4 Analyze data quality	8.4 Review data quality	9.4 Improve data quality
1.5 Identify data sources	2.5 Design data model	3.5 Build data collection infrastructure	4.5 Implement infrastructure	5.5 Operate infrastructure	6.5 Monitor data quality	7.5 Analyze data quality	8.5 Review data quality	9.5 Improve data quality
1.6 Identify data sources	2.6 Design data model	3.6 Build data collection infrastructure	4.6 Implement infrastructure	5.6 Operate infrastructure	6.6 Monitor data quality	7.6 Analyze data quality	8.6 Review data quality	9.6 Improve data quality
1.7 Identify data sources	2.7 Design data model	3.7 Build data collection infrastructure	4.7 Implement infrastructure	5.7 Operate infrastructure	6.7 Monitor data quality	7.7 Analyze data quality	8.7 Review data quality	9.7 Improve data quality
1.8 Identify data sources	2.8 Design data model	3.8 Build data collection infrastructure	4.8 Implement infrastructure	5.8 Operate infrastructure	6.8 Monitor data quality	7.8 Analyze data quality	8.8 Review data quality	9.8 Improve data quality
1.9 Identify data sources	2.9 Design data model	3.9 Build data collection infrastructure	4.9 Implement infrastructure	5.9 Operate infrastructure	6.9 Monitor data quality	7.9 Analyze data quality	8.9 Review data quality	9.9 Improve data quality



Process models for continuous improvement - #3

“The reason [operational] cost [of different parts of the statistical business process] has not been a central focus is a difference between NSIs focus on measuring quality of their products and services, rather than continuous improvement of quality.”

David A. Marker

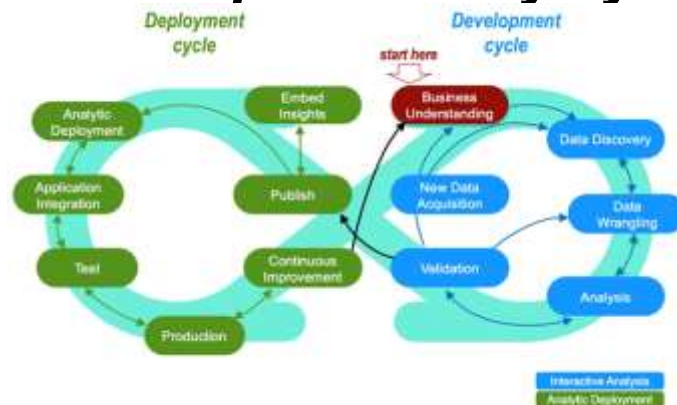


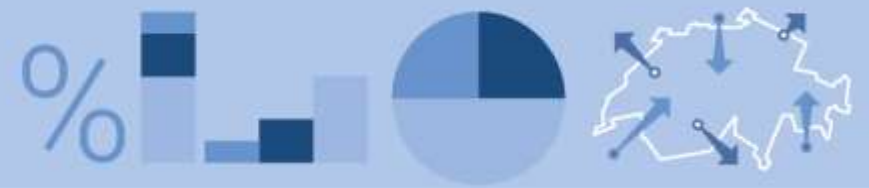
Process models for continuous improvement - #3

3. “CRoss Industry Standard Process for Data Mining”



4. *The complementary cycles of developing and deploying “analytical assets”.*

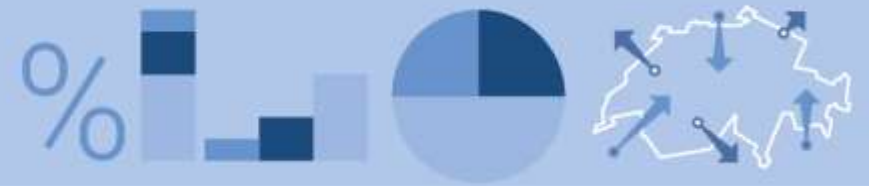




Agenda

1. Introduction
2. Demystifying the “big data” hype
3. Demystifying the “Internet of things” hype
4. “Analytics of things” and “smart (official) statistics”
5. Demystifying the two approaches of analytics
6. Process models for continuous improvement
- 7. Conclusion**

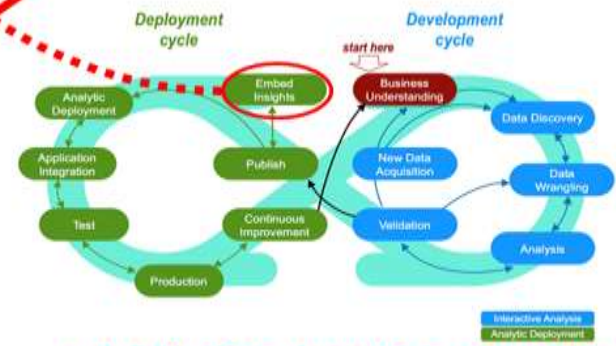




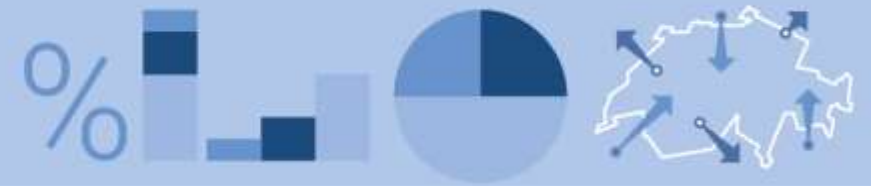
Conclusion

GSBPM («current statistical production»)

Quality Management / Metadata Management							
Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Build collection instrument	4.1 Create data sets & schema	5.1 Integrate data	6.1 Prepare data outputs	7.1 Update output systems	8.1 Initial evaluation inputs
1.2 Consult & resolve needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Estimate output requirements	2.3 Design content	3.3 Build or enhance dissemination components	4.3 Run collection	5.3 Review & calibrate	6.3 Interpret & explain outputs	7.3 Manage release of dissemination products	8.3 Agree on action plan
1.4 Identify concepts	2.4 Design frame & schema	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit & input	6.4 Apply processes & controls	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing & analysis	3.5 Test production system	4.5	5.5	6.5	7.5	
1.6 Prepare business case	2.6 Design production systems & workflow	3.6 Test operational business process	4.6	5.6	6.6	7.6	
		3.7 Finalise production system	4.7	5.7	6.7	7.7	
			4.8	5.8	6.8	7.8	
			4.9	5.9	6.9	7.9	
			4.10	5.10	6.10	7.10	



«Analytics process model»



Questions & Answers

