

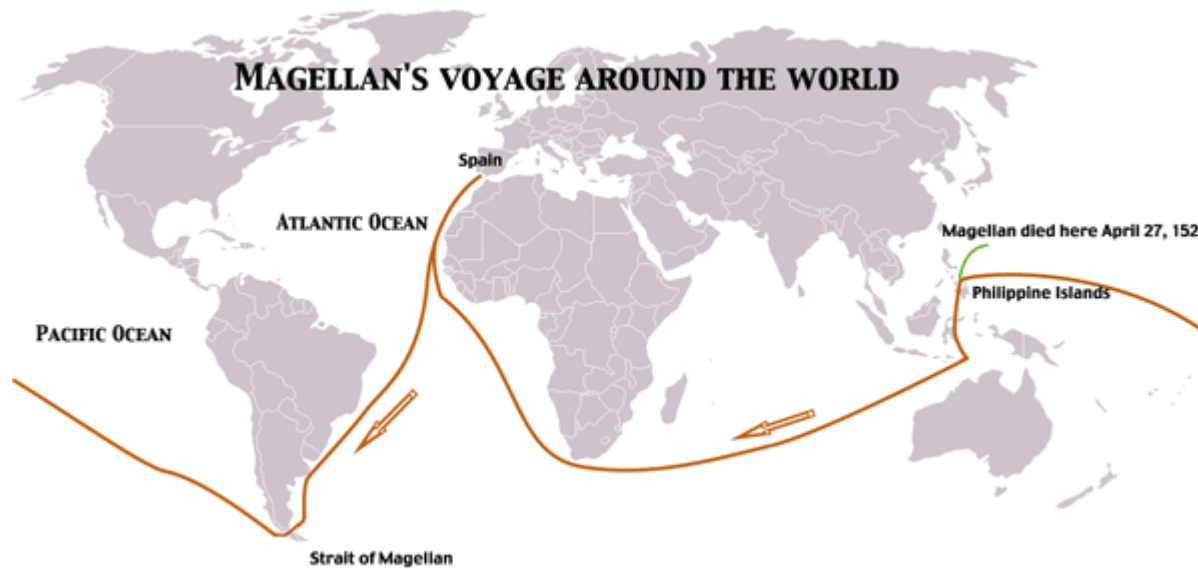
How to trace the audit trails in the big data jungle?

Drs. Wim Leeuwenkamp RA CIA
Strategic Advisor
Ministry of Finance - Tax Office

4 juni 2012



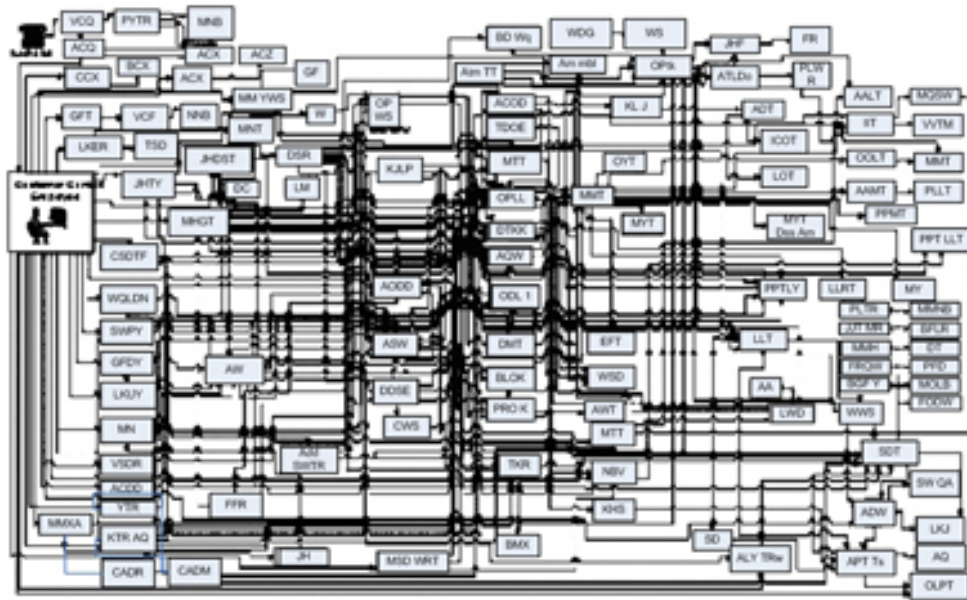
Pilot Project – First Expedition



Ferdinand Magellan (1480-1521) was a Portuguese explorer who led the first expedition that sailed around the Earth (1519-1522).



IT infrastructure – The Jungle



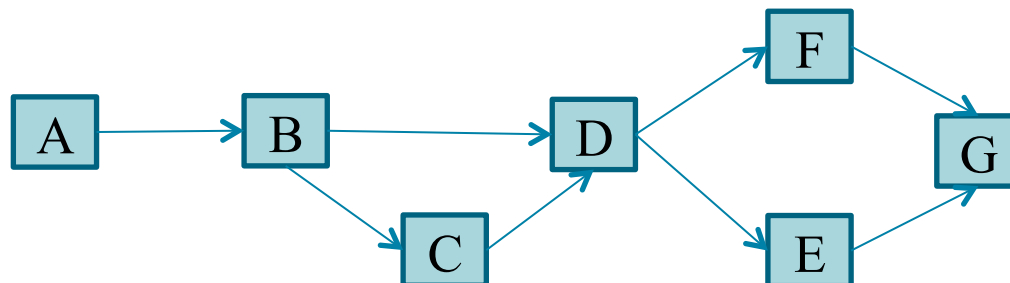
Legacy-systems

“Home made” IT

Complexity
Reduction Projects

Legacy Integration
Projects

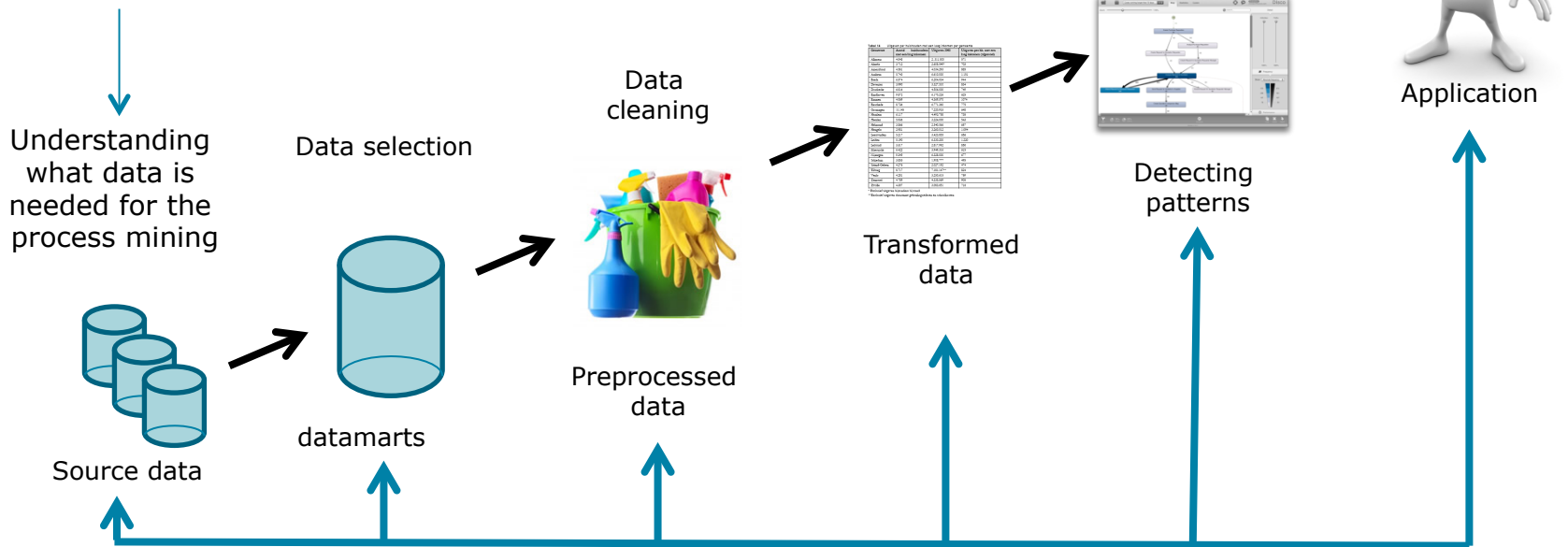
Data quality
Projects





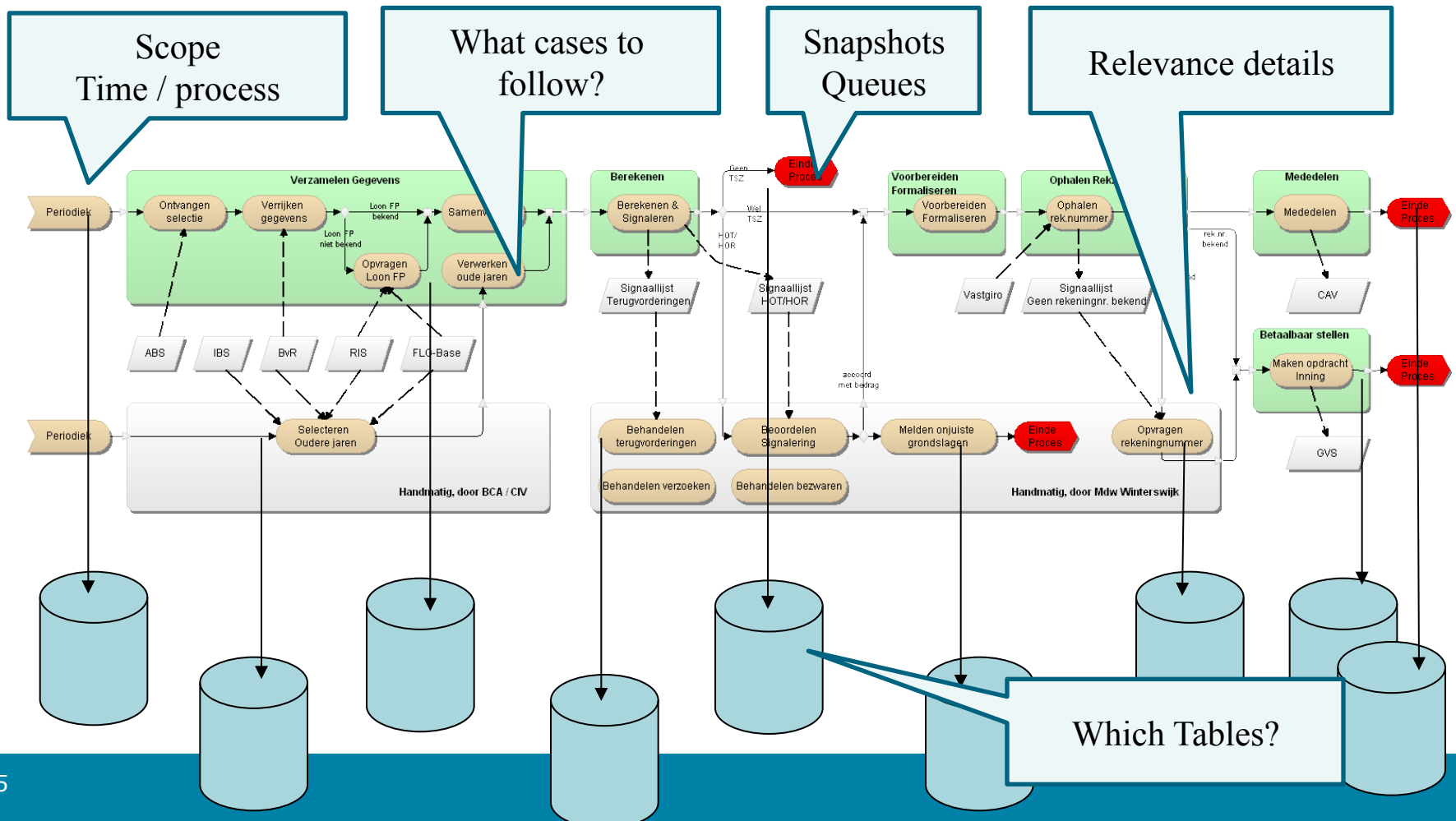
How to cultivate the Jungle?

Dumps of Operational Data



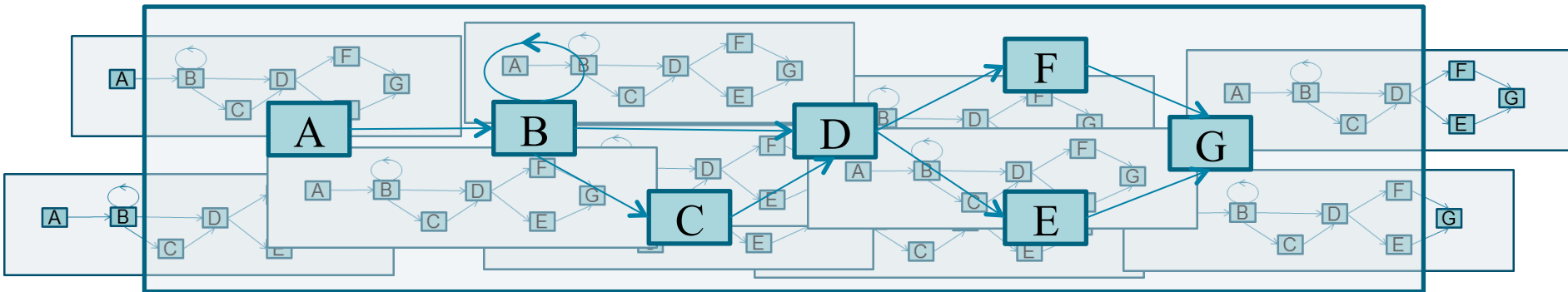


The process – The Data – The Log → Pitfalls...





Explanation of 3 Pitfalls –Scoping Process

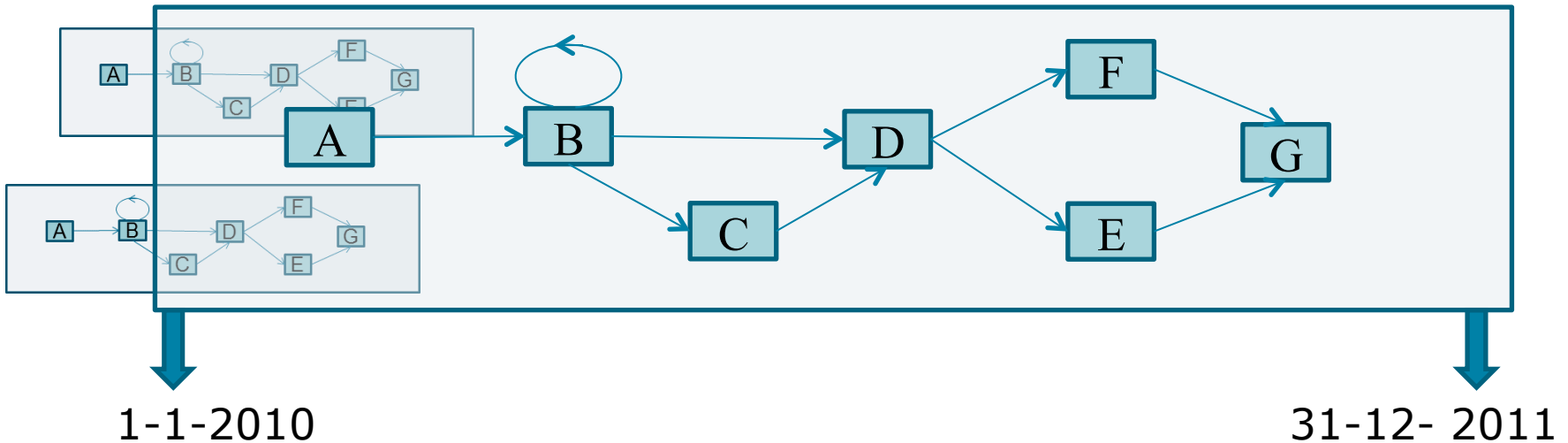


Don't use all available cases or activities in your first log-import-file

- Pre-select the cases with start-end timestamps (filter the missing values).
- When you put in cases that didn't start in a right way, probably the process has more input-options than you expected.
- When you put in cases that didn't end in a proper way, there can be a lot of queuing or the GIGO rule appears to be true!



Explanation of 3 Pitfalls – Scoping Time-periods

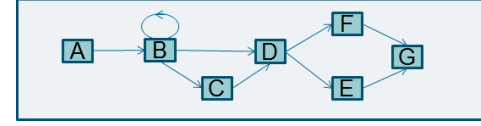
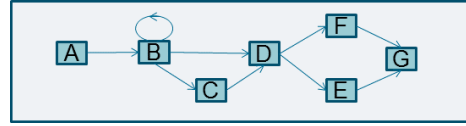
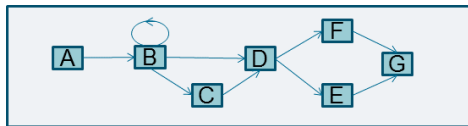
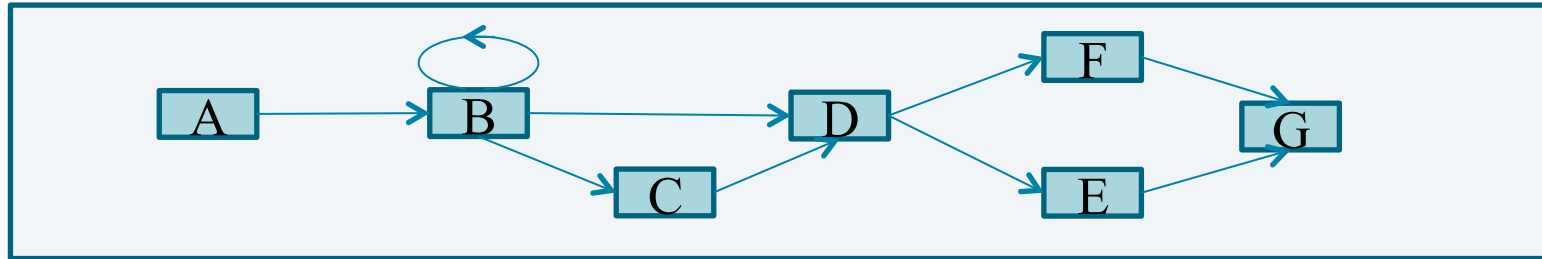


Don't cut-off your input data by filtering on time only (for example year endings)

Pre-select cases that are available in a certain time period and put all the history of these cases in your log-import-file



Explanation of 3 Pitfalls –Analyzing the process



Don't try to analyze all the available data!
Be careful with adding manual sub-processes
into your log-file

It is difficult to stick together processes with different case-identification.
It is difficult to do a conversion of dates to timestamps in a high-speed process



Analyzing & Tooling → The Gamekeeper ...



Process Mining Experts

Business Analysts

Database Managers

Domain Experts

Don't forget your
tooling!



The Audit Trail – Remember!

- Balance Relevance - Availability
- Reliable data!
 - Data quality issues, do not underestimate!
 - 1e solve your data issues & 2e start to analyze
- Exploitability
 - What are the cost to implement logging
 - Use IT-system updates to improve logging
- Reporting & Referring (keep it simple)



Final Remarks - Fireside Chat

The “new kind of auditing” by analyzing the process with a process mining tool gives the auditor a complete view on all the weaknesses in a process. The auditor has more detailed information about the quality of the processing of transactions and can give a better advice to improve processes or organizations.



Questions

