

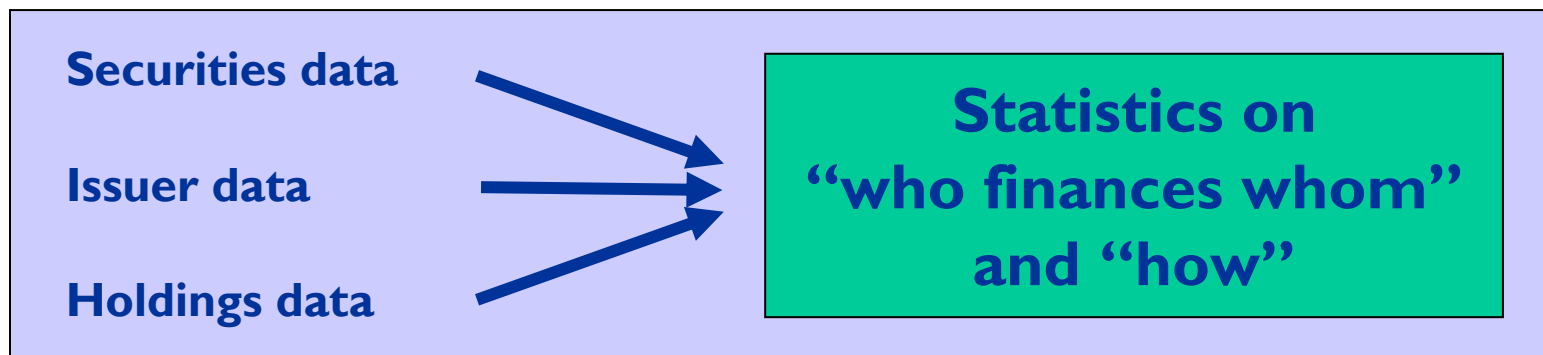
17 March 2009

Data Standards: the case must go to the boards

What can we offer there?

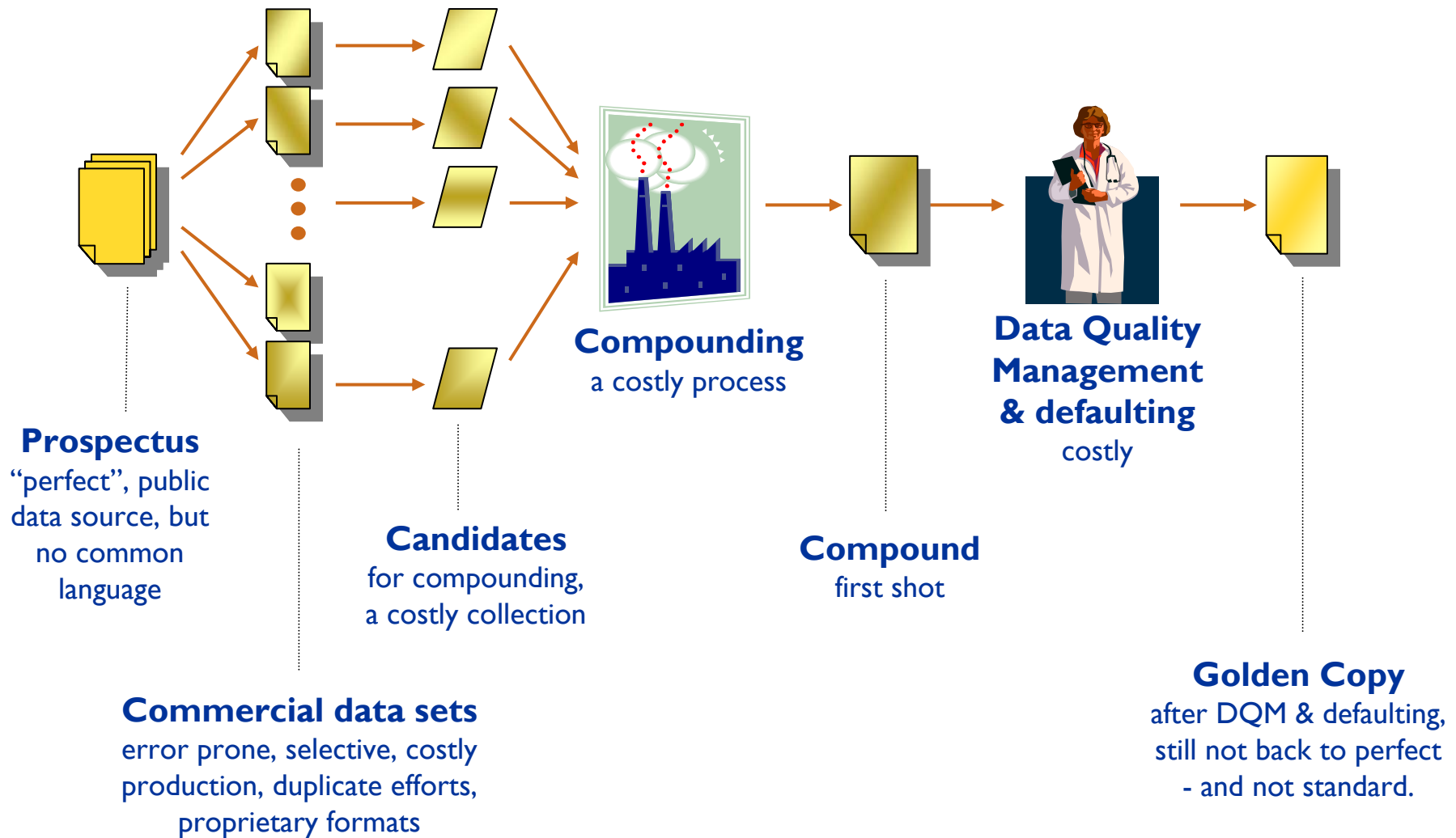
Francis Gross
DG Statistics, External Statistics Division

Case: micro data based financial statistics



- **Further benefits of micro data in statistics:**
 - Flexibility in serving event-driven policy needs
 - Near-time oversight of complex markets
 - Ability to analyse and understand evolving systemic risk
 - Ability to drill down, linking macro- to specific micro-issues
- **The Centralised Securities Database (CSDB)**
 - Holds data on 5 million securities
 - Phase 2 now entering production

How the CSDB tries to clean reference data



The broader context

The broader context: a tale of progress...

- **Political progress gave us open markets and free trade.**
- **Progress in education has given us:**
 - **technical skills to support creativeness and**
 - **a worldwide “lingua franca” of business**
- **Technological progress has given us the means to:**
 - **overcome distance at low cost;**
 - **build the complex structures we can imagine;**
 - **compute and communicate them as fast as we can think;**
 - **process massive amounts of data at high speed.**

...towards more volume, complexity & speed...

- **There will be more data,**
- **From more sources,**
- **Processed faster,**
- **In more systems,**
- **Involving more people,**
- **Serving more complex processes,**
- **Driving more, increasingly interconnected decisions,**

**Complexity, volume & speed vs human limitations:
automation must be trusted ever more**

...with a limiting factor,...

- **Data is both raw material and product of processes.**
- **An organisation can control its own processes,**
- **But it must use data coming from many other organisations,**
- **Which are beyond its control.**
- **Each independent data source speaks a different "dialect";**
- **Differences can be subtle, but sufficient to defeat automation.**
- **Evolving markets bring more "data dialects" together;**
- **Mixing more "data dialects" gradually defeats known fixes.**

**Are we re-building the Tower of Babel,
but this time with data?**

...that needs recognition as a key risk ...

- **Weak data limits essential functions such as automation, market transparency, regulation and systemic risk diagnosis;**
- **Weak data certainly worsened the current crisis.**
- **Taken for granted, data is generally not a board issue, and data issues are not well understood at senior level; therefore**
- **Risks from weak data grow largely unnoticed by the board;**
- **With evolving markets and IT, weak data act as a slow poison.**
- **Moreover, a single organisation cannot solve it, hence a risk of inaction, even when data weakness is recognised as a risk.**

Weak data poses a key risk, in two ways: (i) it could grow into a danger unnoticed and (ii) be met with inaction.

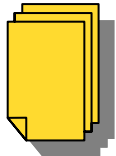
...and resolution through standardisation.

- **Good data requires standards, discipline and rigour, not IT.**
- **There must be less diversity in “data dialects”.**
- **Data will be good for all users or for none.**
- **The “human language / data” interface is the key challenge;**
- **That first layer of data is a shared strategic resource.**
- **The generation of good data for all requires top down action.**
- **Leaders must be helped to understand the importance of rigour for terms and definitions, and the mechanisms and toxicity of “dialect diversity” in data.**

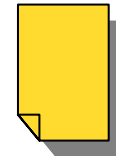
**Data standards must be driven from the top.
Technicians must convey the case to the top!**

What next?

A simple idea: clean reference data on (public) tap



Prospectus
"perfect" public
data source



"Golden Source"
"perfect" electronic data
sheet, from the source

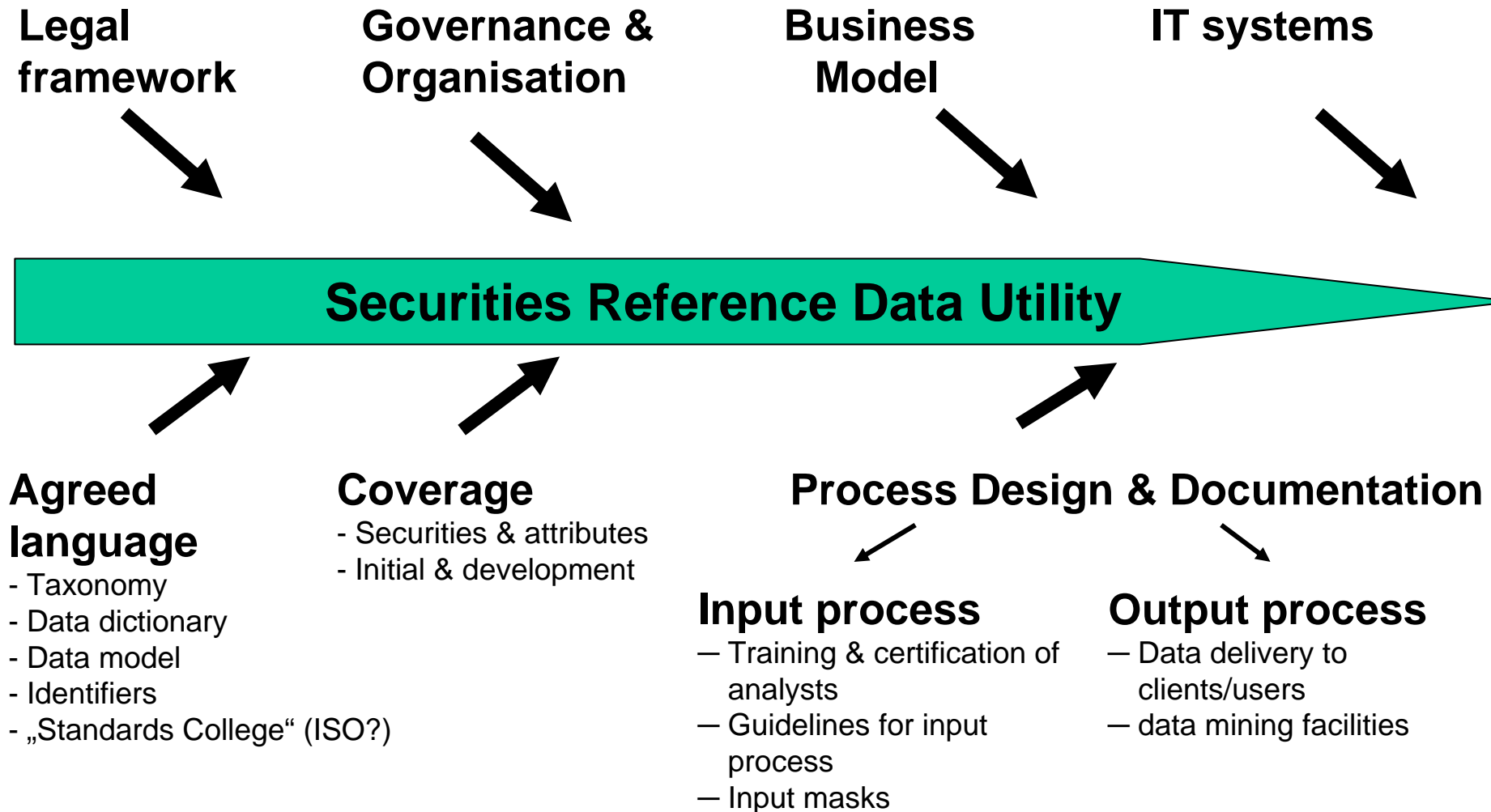
- **Delivery of electronic reference data sheet part of issuance process**
- **Sheet covers a selected set of reference attributes**
- **Sheets are collected and maintained in a central database**
- **If the source is public, data is public: the database is a public good**
- **The database offers a source for official, private & commercial use**

**The idea:
a Securities Reference Data Utility**

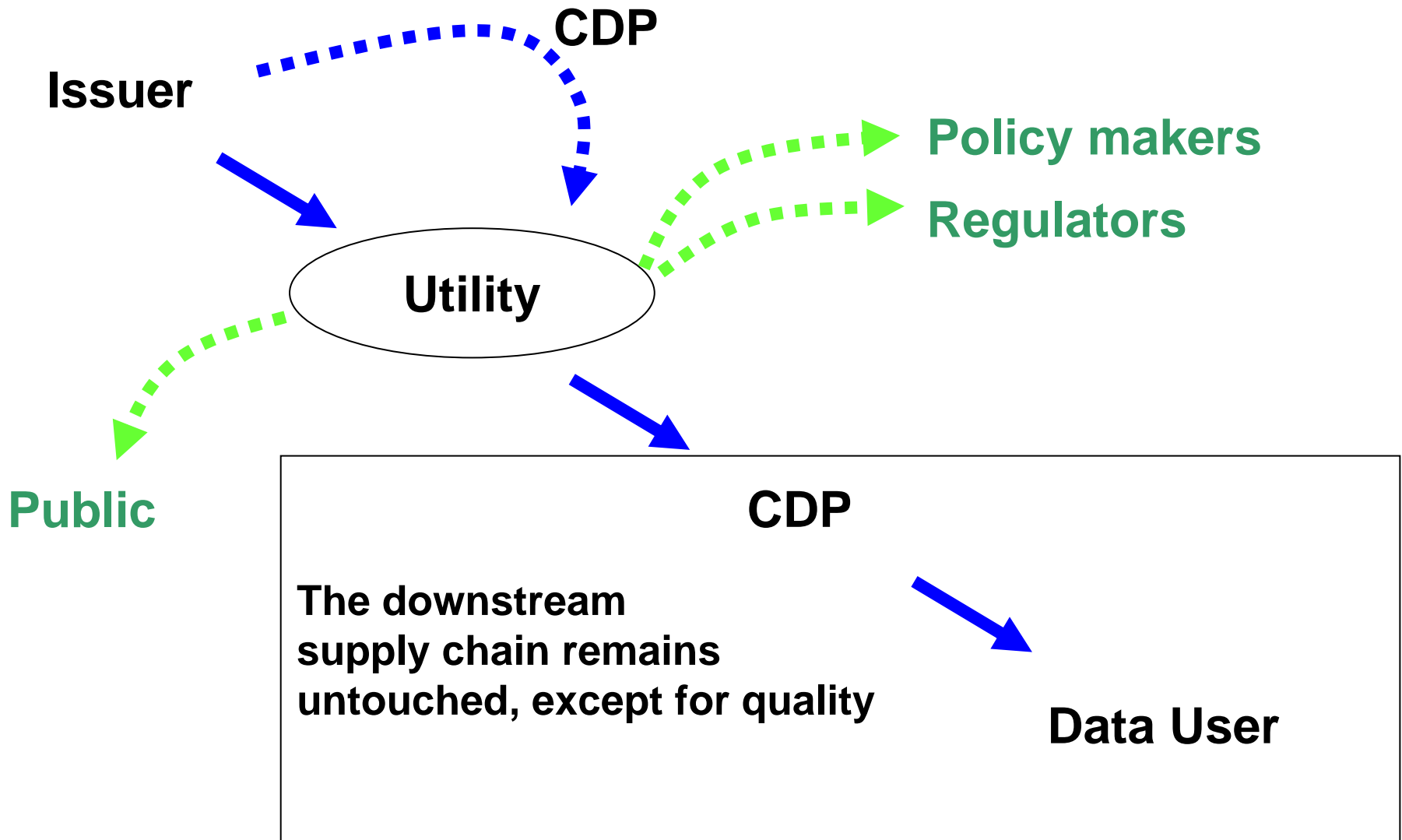
How might it work?

- **Standard taxonomy**
- **A competition-neutral data model**
- **Issuer responsible for data production and maintenance**
- **Trained & certified analysts perform the service**
- **CDPs offer production and maintenance services for data**
- **CDPs can source reference data from the Utility**
- **Users can continue using CDPs as so far**
- **Quality management the internet way: „any to one“**

Components needed to make it work



Positioning in the data supply chain



Alternative: Multilateral self-organisation

- Indispensable for agreeing on high-level principles and initiatives
- Hard to insulate from the competitive game
- Less well suited for converging rapidly on technical solutions

2 July 2008

Ten Industry Initiatives to Increase Transparency in the Securitisation Market
Executive Summary

Costs of development and implementation:

- **Data model:** existing, state-of-the-art
 - **IT:** a simple, „flat“ database; online access
 - **Input templates:** designed once, can evolve over time
 - **People:** training & certification of analysts
 - **Lawmaking:** might cost time
- **Running costs:**
 - **Filling the sheets:** a fraction of emission costs
 - **Quality mgmt:** no duplication anymore, users provide input
 - **Taking out duplication alone reduces costs significantly**

Benefits for the financial industry

- **Commercial data providers**
 - Improved quality of reference data offered to clients
 - Lower costs for data collection
 - A new service to issuers, producing & maintaining their data
- **Businesses in the financial industry**
 - Improved information on markets and on own position
 - Safer and freer markets through higher transparency
 - Lower costs and risks in business processes
 - Cost of collection (more than) offset by savings in operations

Benefits for central banks & regulators

- **Transparency of financial markets**
 - Improved technical infrastructure for transparency
 - **Public impact: concrete and credible progress**
- **Financial Market Integration**
 - A shared infrastructure for official, public and industry uses
 - Across Europe or, better, worldwide
- **Operational benefits for central banks and regulators**
 - In Statistics, Collateral Management, Financial Stability
 - Better data quality in the CSDB and (over time, much) less DQM