Deloitte.

Datová kvalita nejen pro Solvency II

Seminář z aktuárských věd

Petr Dvořák

16 Května 2014

Agenda

- Požadavky Solvency II na datovou kvalitu
  - Definice datových standardů – přesnost, vhodnost, úplnost
  - Datový slovník a dokumentace
  - Odpovědnost za kvalitu dat

- Jak naplnit požadavky Solvency II a zefektivnit datové procesy?
  - Kontroly datové kvality (pro zajištění přesnosti, vhodnosti a úplnosti)
  - Automatizace a workflow
  - Datový slovník (a jeho souvislost s datovou architekturou)
  - Organizace a role v datových procesech

Data might lead to unexpected results…

Data quality is not only an IT problem – it is mainly about the understanding between different parts of the business

Common Data Quality Issues

Relevant persons for solving DQ issues could not be found	Processes are not documented enough and therefore source of inputs and calculation assumptions could not be found retrospectively	The calculations are too laborious
No responsibility for produced data	Low level of process auditability and traceability	A lot of manual work in SII processes
Data quality tests are not sufficient	Many data sources in use	Bad understanding of data
Data quality assessment do not ensure compliance with Solvency II requirements	Increase an operational risk; Different reports of the same indicator	Prolonging time required for data collection; Can lead to confusions and bad decisions

Solvency II Requirements

Data Quality Requirements

Solvency II defines **data quality standards:**

- Accuracy
- Completeness
- Appropriateness

The **key requirement** from the Solvency II Directive is very simple:

> Insurance undertakings should have **internal processes** and procedures in place to ensure the **appropriateness, completeness and accuracy** of the data used in the calculation of their technical provisions or entering their internal model.

The regulator will check whether these standards are met and how.
However, all of this makes sense even from a business point of view.
⇒ These standards may be used as the core for the enterprise wide data quality management – not just for Solvency II related processes and data.

- The requirements are on a **very general level** - it is left **up to the company** how to ensure they meet the standards
- The company needs to be able to **convince the regulator** – documentation

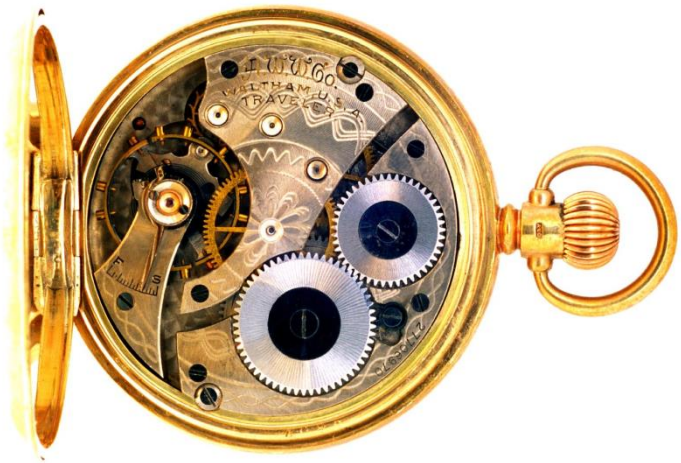
- Relevant for data entering
  - calculation of the technical provisions
  - internal model
  - calculation of undertaking specific parameters
- Not relevant for data for standard formula, MCR, QRT, ORSA etc.
- There are other DQ requirements e.g. connected with actuarial function

Accuracy

Solvency II definition of accuracy (*see Draft Delegated Acts, Article 14, par. 1*):

- the data are **free from material errors**
- data from different time periods **used** for the same estimation are **consistent**
- the data are **recorded** in a timely manner and **consistently** over time

- Typical causes of inaccuracy:
  - driven by bad inputs and inaccurate data transformations etc. caused by:
  - manual work or
  - problems with IT systems
    - many different IT sources,
    - data systems are outdated,
    - not general link data systems – technical/business areas



- Typical tests:
  - reconciliation test – reconciliation of technical reserves, premium, claims paid to the balance sheet or P&L
  - time consistency – e.g. data in a triangle are consistent in time
  - aggregate statistics – mean, standard deviation, quantiles, distributions

Completeness

Solvency II definition of completeness (*see Draft Delegated Acts, Article 14, par. 2*):

- the data include **sufficient historical information** to assess the characteristics of the underlying risks, such as to identify trends in the risks
- such data are **available** for each of the relevant **homogenous risk groups** used in the calculation of the technical provisions and no such relevant data is excluded from being used in the calculation of the technical provisions without justification

- Example:
  - MTPL triangle of claims paid based on 5-year history
    - not enough data for determine a trend
    - ⇒ could not be considered as complete data
  - MTPL triangle – 20-year history – together for property and bodily injury
    - different homogenous groups ⇒ level of granularity is not appropriate
    - ⇒ could not be considered as complete data
  - Lapses assumption in dimensions age and sex
    - may not be complete if significantly depends on other dimensions (e.g. a distributional channel)

Appropriateness I

Solvency II definition of appropriateness (*see Draft Delegated Acts, Article 14, par. 3*):

- the data are **consistent with the purposes** for which it will be used
- **the amount and nature of the data** ensure that the estimations made in the calculation of the technical provisions on the basis of the data **do not include a material estimation error**
- the data are **consistent with the assumptions** underlying the actuarial and statistical techniques that are applied to them in the calculation of the technical provisions
- the data appropriately **reflect the risks** to which the insurance or reinsurance undertaking is exposed with regard to its insurance and reinsurance obligations
- the data were **collected, processed and applied in a transparent and structured manner**, based on a given specification (see the next slide)

→ E.g. Do not use yearly assumptions, whereas a model requires monthly one

→ Simplification in modelling
Example: neglecting of a product because of its insignificance ⇒ monitor the significance of the product

→ Do data satisfy the model assumptions?
Example: triangle methods

→ Particular for historical data used for projection – historical data must be consistent with current risks

→ Transparency in line with the specification (“Data Policy”)

Appropriateness II

Appropriate data should be collected, processed and applied in a transparent and structured manner, based on a **specification** of at least the following areas (*see Draft Delegated Acts, Article 14, par. 3*):

- the **definition and assessment** of the quality of data, including specific qualitative and quantitative standards for different data sets;
- the use and setting of **assumptions** made in the collection, processing and application of data;
- the process for carrying out **data updates**, including the frequency of regular updates and the circumstances that trigger additional updates.

→ *process of DQ assessment*
→ *DQ matrices, definition of tests performed*
→ *DQ tests coverage*

→ *setting of the rules for collection, processing and application*

→ *frequency of regular data updates;*
→ *circumstances that trigger unscheduled data updates and the timeliness of their realisation*

Data Directory and Other Documentation Requirements

Insurance and reinsurance undertakings shall document the following processes (*Art. 197 – USP / 256 – TP*):

- the collection of data and analysis of its quality;
- the choice of assumptions used in the calculation/production of data;
- the selection and application of actuarial and statistical methods;
- the validation of the data.

The documentation of calculation of the technical provisions *(Art. 256)* / internal model *(Art. 232)* / USP *(Art. 197)* shall include

- **a directory of the data, specifying**
  - **the source,**
  - **characteristics and**
  - **usage and**
- the specification for the collection, processing and application of the data *(link to the "Data Policy")*
- where data are not used consistently over time in the calculation, a description of the inconsistent use and its justification (*note: not relevant to USP*)

The documentation of calculation of the technical provisions shall also include

- a **directory of all the relevant assumptions** that the calculation of technical provisions are based upon
- and others, see Art. 256, 232 and 197

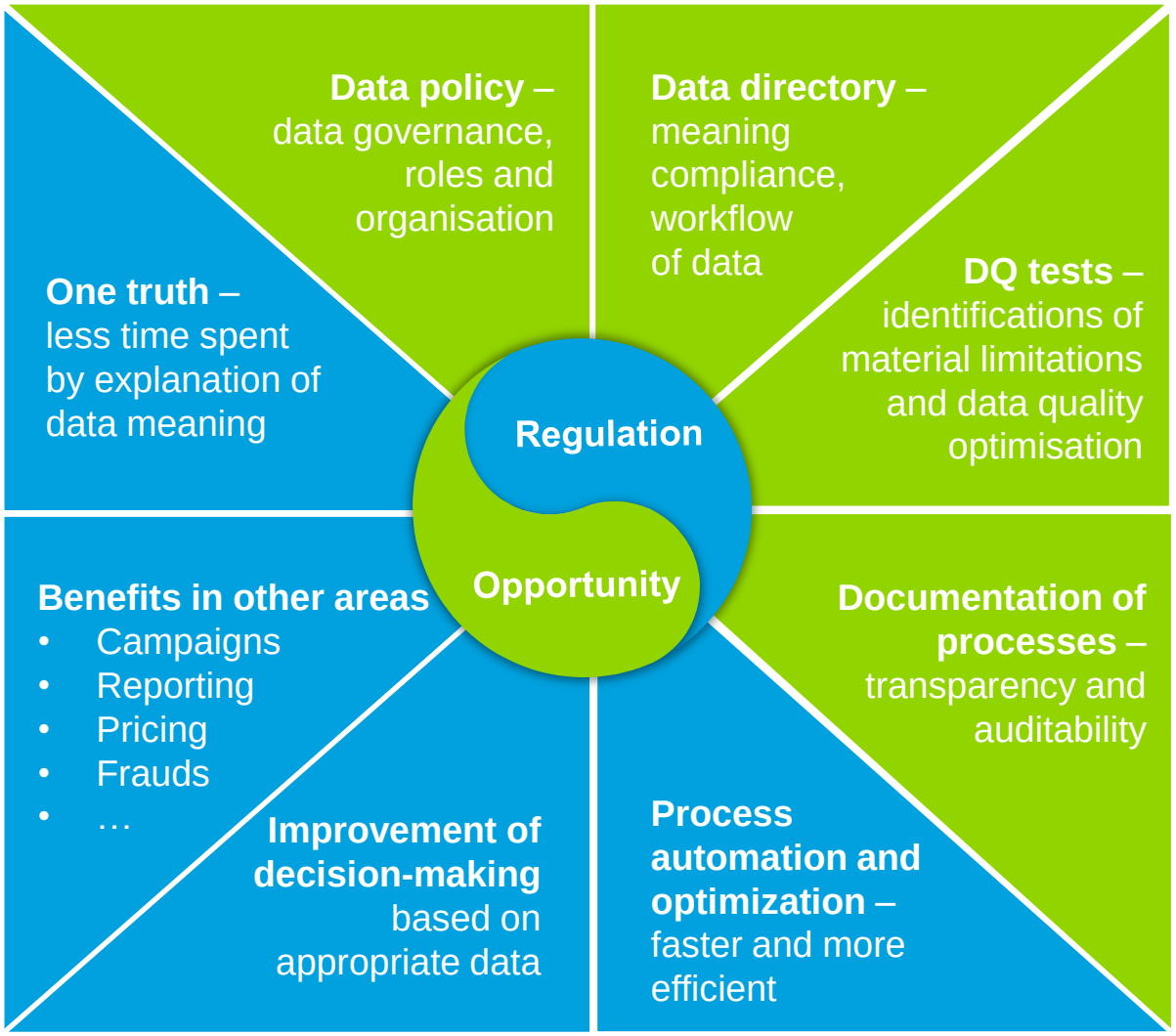
Actuarial and risk management function

According to the Solvency II directive and the draft Delegated Acts, the **actuarial function** shall among other

- **Assess the sufficiency and quality of the data** used in the calculation of technical provisions
- Ensure that any **limitations of data** used to calculate technical provisions are properly dealt with
- The comparison of best estimates against experience shall include comparisons between observed values and the estimates underlying the calculation of the best estimate, in order to draw conclusions on the appropriateness, accuracy and completeness of the data and assumptions used as well as on the methodologies applied in their calculation.

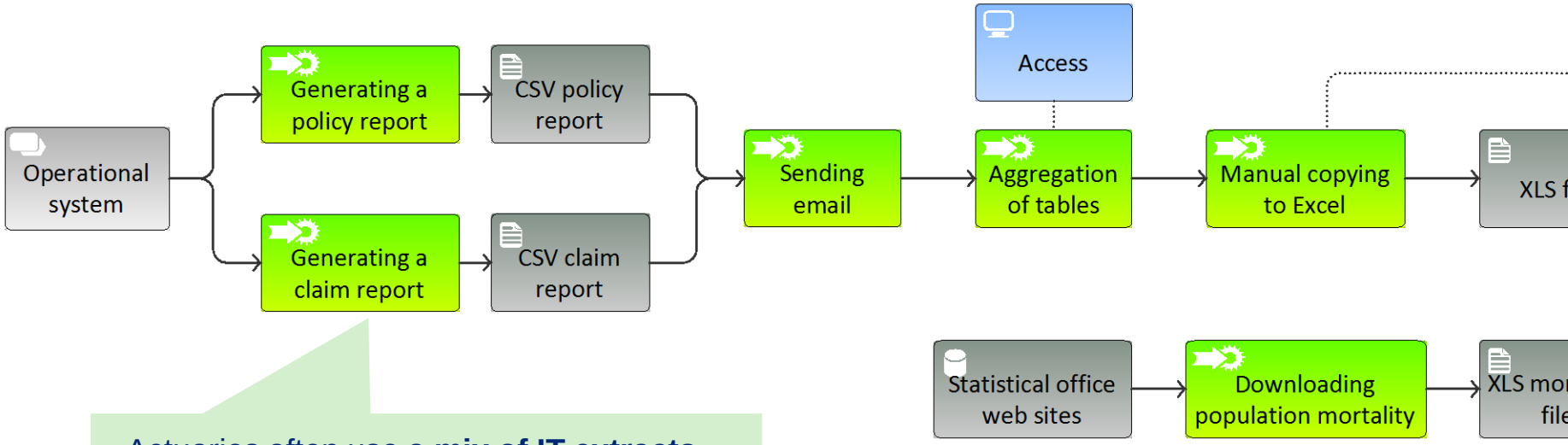
According to Article 44 the **risk-management function** shall test and validate the internal model. According to Article 124 the validation of internal model shall also include an **assessment of the accuracy, completeness and appropriateness** of the data used by the internal model.

Regulation = Opportunity?



How to Fulfill the Requirements?

Example: Preparation of Mortality Tables

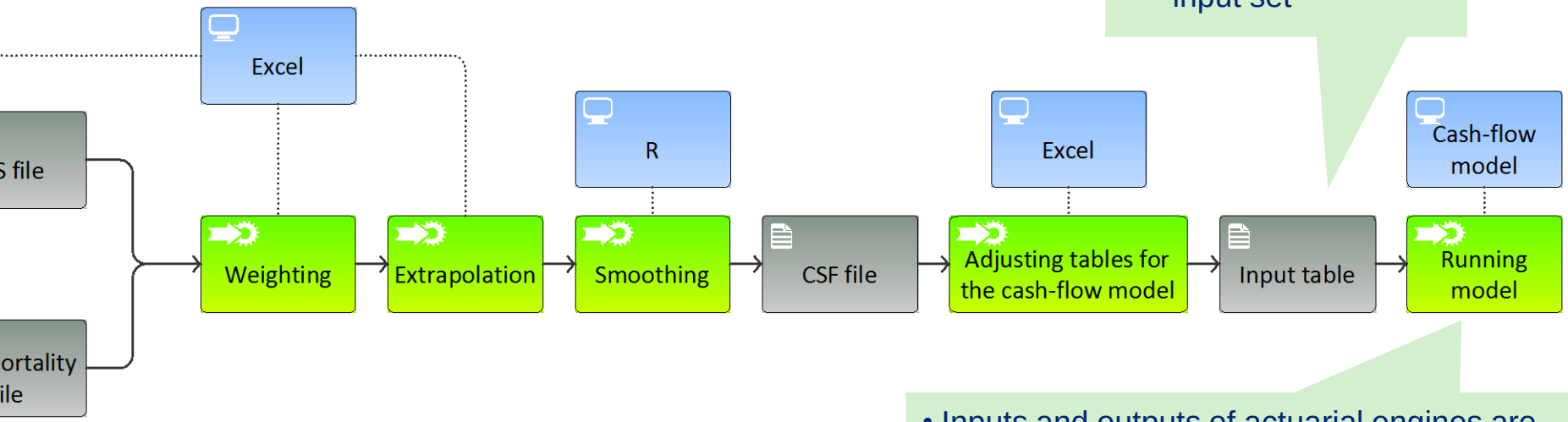


- Actuaries often use a **mix of IT extracts** from operational systems and various business managed files
- The **origin** of some data is sometimes **unknown**

Example: Preparation of Mortality Tables

- Process is highly **decentralized** and involves many tools
- Many **manual** adjustments are performed
- Data quality checks are performed but on an **informal** basis

- Data are not stored in consistent sets making difficult the reuse of a given input set



- Inputs and outputs of actuarial engines are not **systematically** defined
- **Traceability** is not an automated functionality
- No "**One single model**" approach

DQ Controls and Measurement

Data Quality Measurement – Process

Needed to fulfil data quality standards: accuracy, appropriateness and completeness

Appropriate data should be collected, processed and applied in a **transparent and structured manner**, based on **qualitative and quantitative standards** for different data sets. (“Data Policy”)

Data requirements definition

- Included in **data directory** for each data attribute
- Usage of the data in the calculation should be in enough detail
- Example: Model calculates lapses in the middle of a month ⇒ lapses assumption have to be consistent

Data attribute **preparation**

- **Documentation** of data preparation
- Example: Lapses were derived so that they refer to the middle of a month

Tests

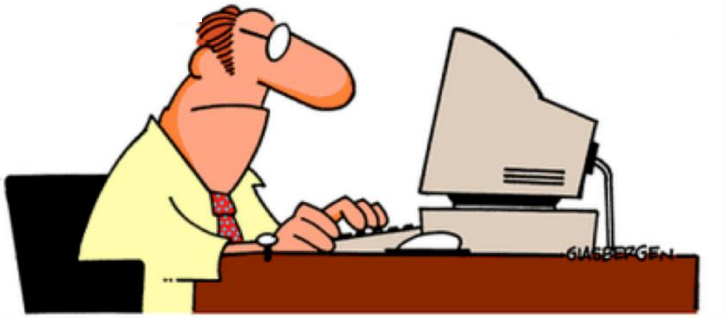
- **Tests** are performed according to the data requirements
  - automatic/manual
  - with/without knowledge of data usage
  - technical/business
- Example: **Documentation** of lapses assumption preparation should be checked – is probability of lapse related to the middle of a month?

Sign-off

- A sign-off is given after review of tests results

Data Quality Measurement – Examples
Accuracy

- Accuracy need to be controlled especially **at the level of data input**, where data enter the company's systems and when data are **transformed**
- Format of data attributes need to be limited to a logical set, but in a reasonable manner
- Actuaries should review the accuracy of data used for the calculations, but good quality of data need to be ensured already before – in the operational systems and data warehouses



Report date		Date of loss	
Last update		Risk	MTPL
Cause of loss		Cedant name	
Gross loss		Treaty	12929 - Treaty 1 94821 - Treaty 2 123345 - Treaty 3
Sum insured/PML			
Cash-call amount requirement			
Currency	EUR	Requirement on cash call	
RBNS		Loss desription	
Retention on loss			
Reinsurers' share on loss			
Present status of the File		Comments	
Original file	Browse		
Save			

Data Quality Measurement – Examples

Appropriateness and completeness

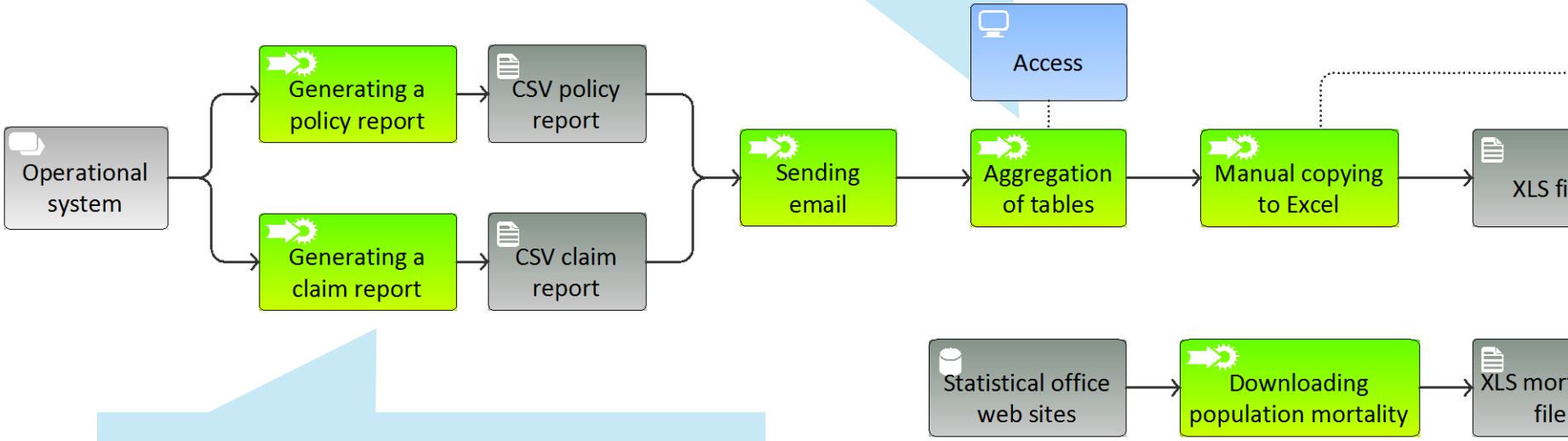
- Testing the completeness and appropriateness

DQ standard	Problem
Completeness (sufficient information)	Do I have enough data to extrapolate them (e.g. mortality tables)?
Completeness (granularity)	Should I split mortality tables also according to other criteria (product, region, sum insured,...)?
Appropriateness (reflecting the risks)	Should I consider some trend? Should be considered prolonging of life expectancy?
Appropriateness (consistency with the purposes)	The model calculates deaths in the middle of a month – is the method of preparing probabilities consistent with this application?
Appropriateness (consistency with the purposes)	Products without mortality risk are recorded in IT systems in other way – should they be included in the derivation of mortality tables?
Completeness (sufficient information)	Deaths are often reported with delay – what impact does this implies? Should I adjust the data?

- Detailed knowledge of data collection, processing and application and models themselves is needed
- For completeness and accuracy we often need to know the usage of the data – link to a data dictionary

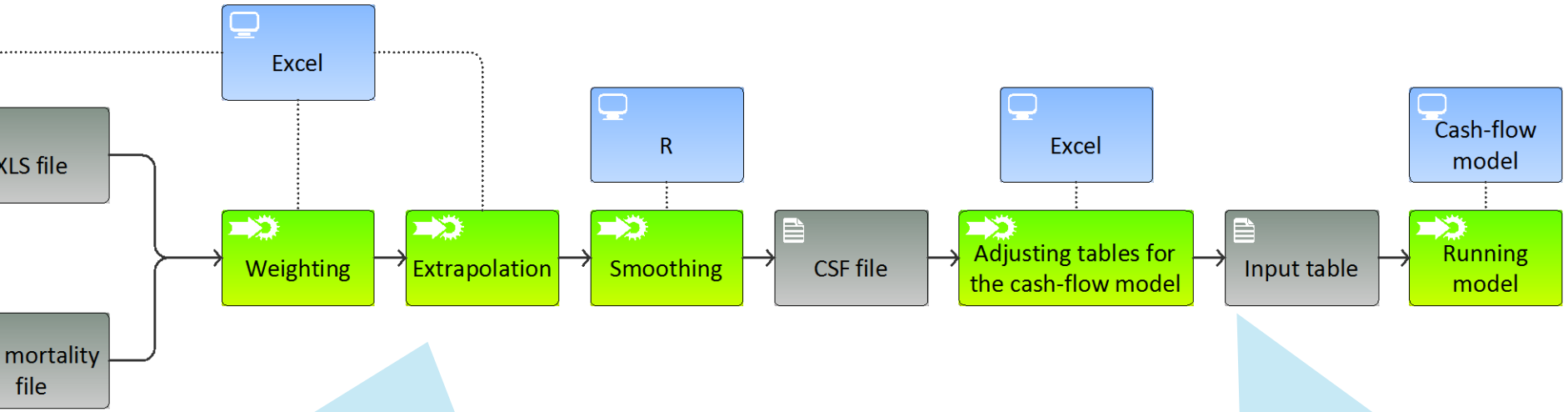
Example: Preparation of Mortality Tables

- Appropriateness/completeness: Mutual comparison, comparison to previous data, analysis of trends, comparison to portfolio statistics
- Accuracy: connecting errors?



- Individual accuracy tests (data formats, missing values...)
- Data requirements need to be defined in advanced by a data owner

Example: Preparation of Mortality Tables

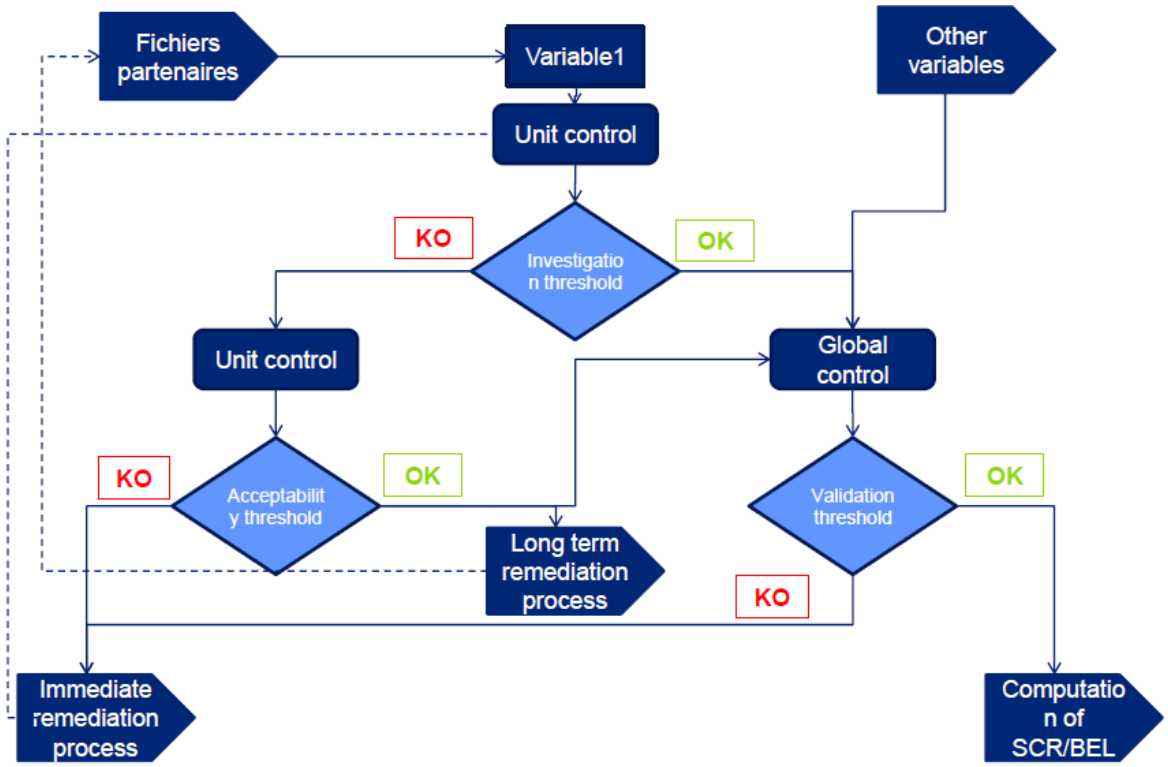


- Continuous checks of average mortality and analysis of change

- Appropriateness and completeness checks – are input data in line with model assumptions?

Data Quality Metrics

- Data quality metrics
  - Definition of limits for each test
  - Set of aggregation way
  - Further limits on the aggregation level



Automation and Workflow

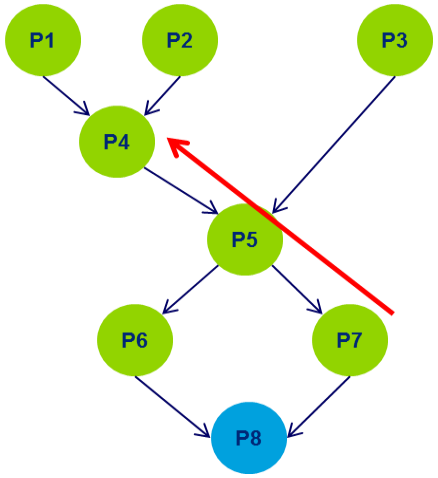
Automation

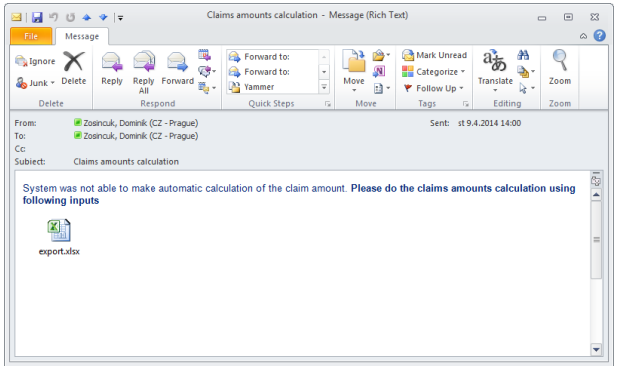
- Automation of selected processes aims to
  - Increase **transparency, auditability and reliability** -> appropriateness
  - Make the processes more **efficient**
  - **Speed up** the processes
  - **Decrease operational risk** connected with manual processing

- To be in the focus
  - Frequently used processes
  - Processes with higher risk exposure
  - Slow processes
  - Specialised technological solution

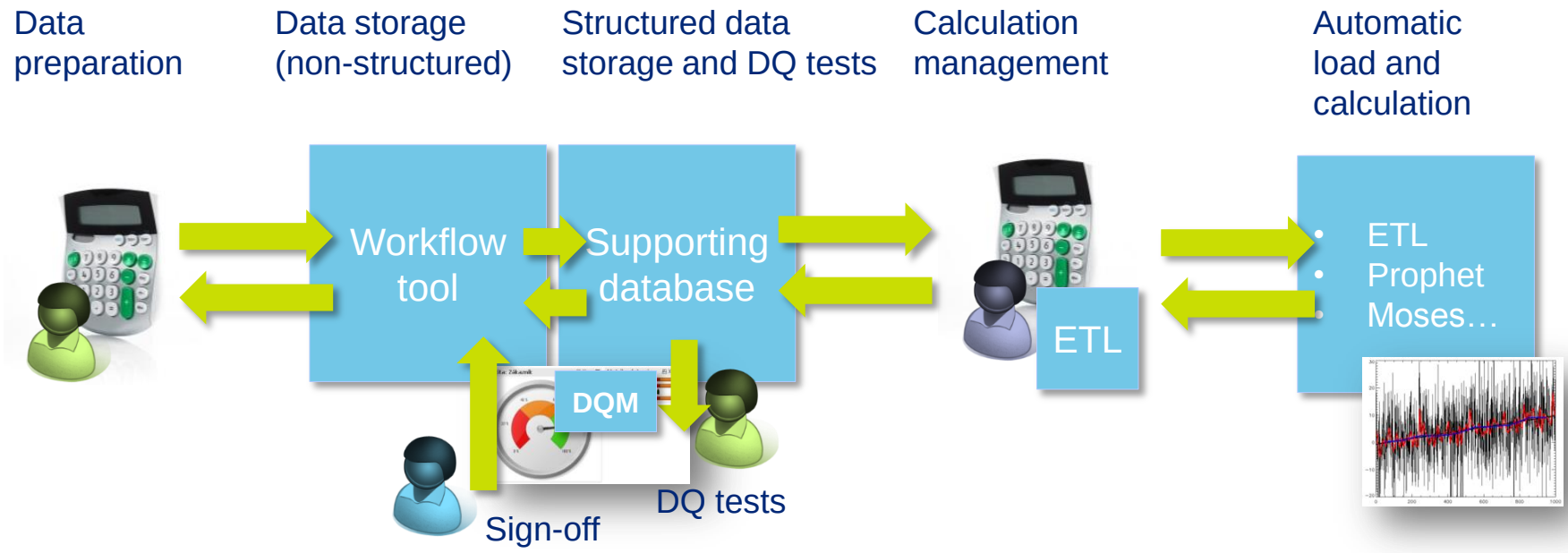
Workflow tool

- Process Definition Tool: A graphical or textual tool for **defining the business process**
- Task Initiation and Control: The business process defined above is initiated and the appropriate person are scheduled and/or engaged to **complete each activity** as the process progresses
- Document Routing: In simple systems, this might be accomplished by **passing a file** or folder from one recipient to another (e.g., an email attachment). In more sophisticated systems, it would be accomplished by checking the documents in and out of a central repository.
- Process documentation evidence: the process progress is **documented** – what data/information was processed, who is the responsible person etc. This ensures the auditability of processes.
- Work-lists: These allow each worker to quickly identify their **current tasks** along with such things as due date, goal date, priority, etc.
- **Task Automation**: Computerised tasks can be automatically invoked. This might include such things as letter writing, email notices, or execution of production applications.





Process Auditability and Automation



Workflow tool
- manages the calculation process
- manipulates inputs and outputs of each calculation step
- documents the calculation itself (e.g. stores model)

Supporting database
- one place for storing SII data
- enables performing DQ tests

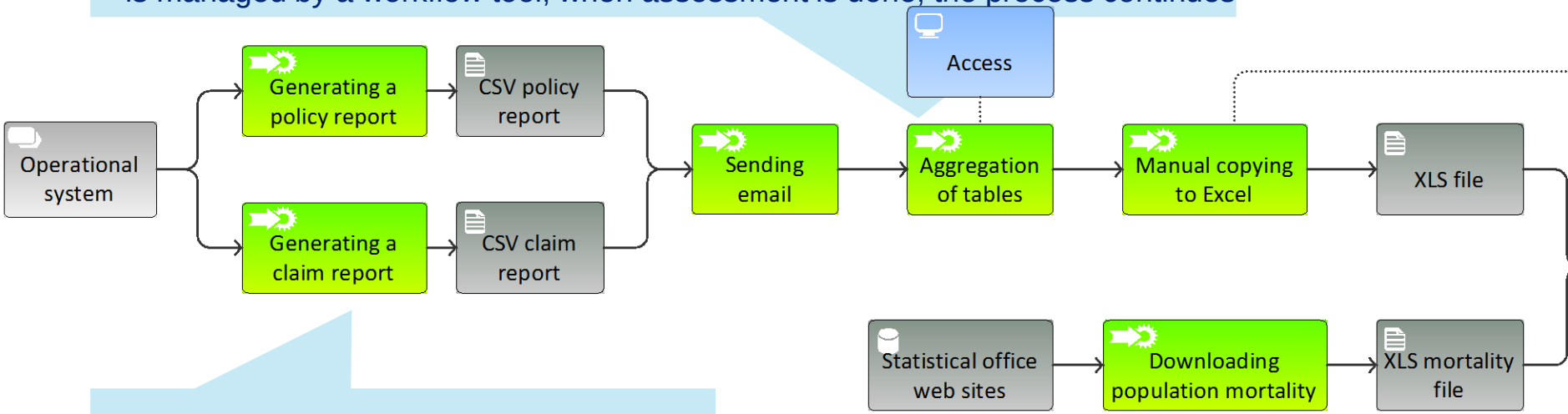
ETL
- extract, transform and load
- automatic data procedures

DQM tool
- automatically performs technical tests

Example: Preparation of Mortality Tables

Fully automatize by a system

- Automatic aggregation of tables using defined rules
- DQ tests: automatic preparation of reports for proving appropriateness and completeness, which is assessed by a responsible person – the report delivery is managed by a workflow tool, when assessment is done, the process continues

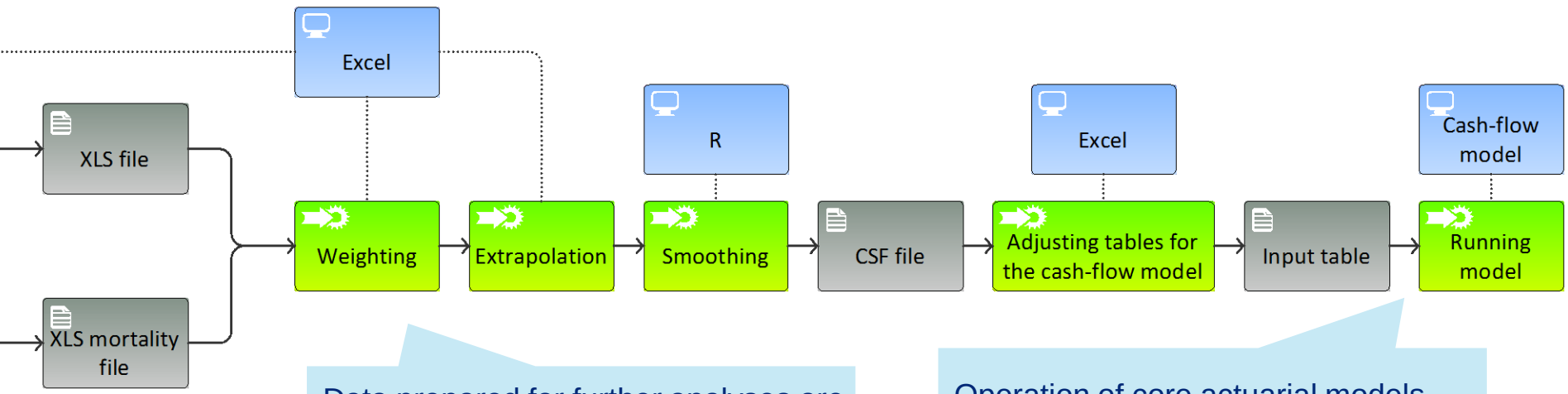


- Policy and claim reports generated automatically,
- Predefined DQ tests for accuracy performed manually

- Automatic download of population mortality tables

Example: Preparation of Mortality Tables

Some analyses could not be automatize, but can be processed more easily – e.g. to perform all analyses in **one system**?



Data prepared for further analyses are handed over via a workflow tool

Operation of core actuarial models should be automated as well to ensure process auditability (and to prove appropriateness of model outputs)

Industrialisation of actuarial models

Inputs

- **Removal of manual** interventions (adjustments, scaling, and grouping) to facilitate automation and control
- Harmonise file formats
- Automate **validation** and fix errors at source
- **Assumptions should be externalised** from the model to facilitate easy updates and sensitivity testing
- **Automate the control** and governance processes – e.g. assumption for SCR shocks (standard formula)

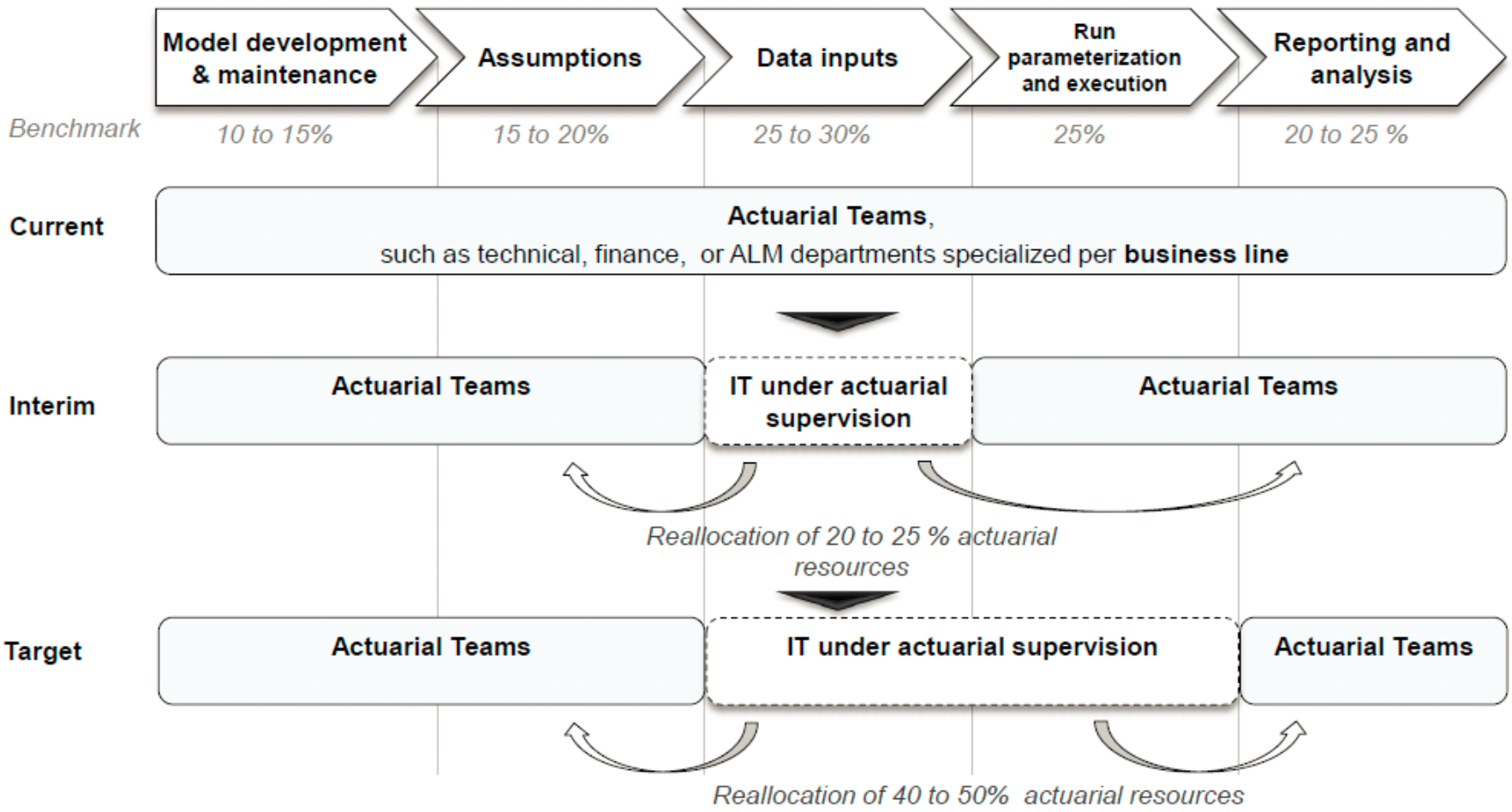
Models

- Build a **single model** to be used for all applications and all lines of business
- Utilize workflow tools to **automate and schedule** all projections to meet full range of reporting requirements
- Consistency of calculations
- Move external processes/calculations in to the model to reduce manual adjustments
- Reduced maintenance costs

Reporting

- **Identification** of all reporting requirements to centralize and automate process
- **Elimination of manual manipulation** of data through automated transformation of the data – automatic processing of model outputs
- Extend coverage by the existing tools (actuarial model / reporting tool) to cover all required adjustments

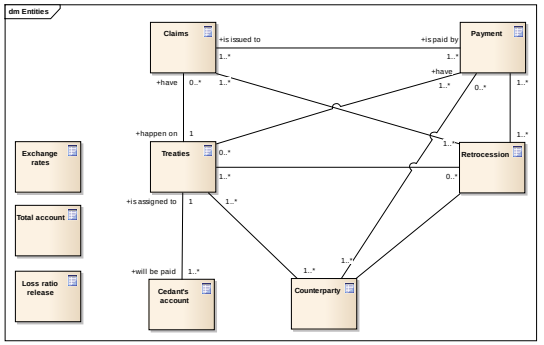
Organisational changes



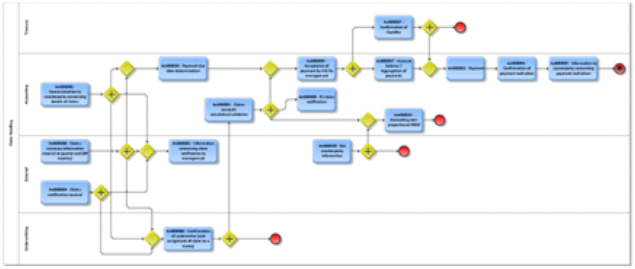
Data Directory

Data Directory

- Either specifically prepare just for Solvency II purposes or use (and broaden) techniques of data management:
- **Data dictionary** (data model)
  - describes data attributes (smallest data pieces) within data entities (logical groups) typically in s given system
  - metadata (information about data) for each data attributes
  - includes data type, data owner, attribute ID, IT system, archiving requirements, access limitations etc.



- **Process flow**
  - describes dependency of data attributes

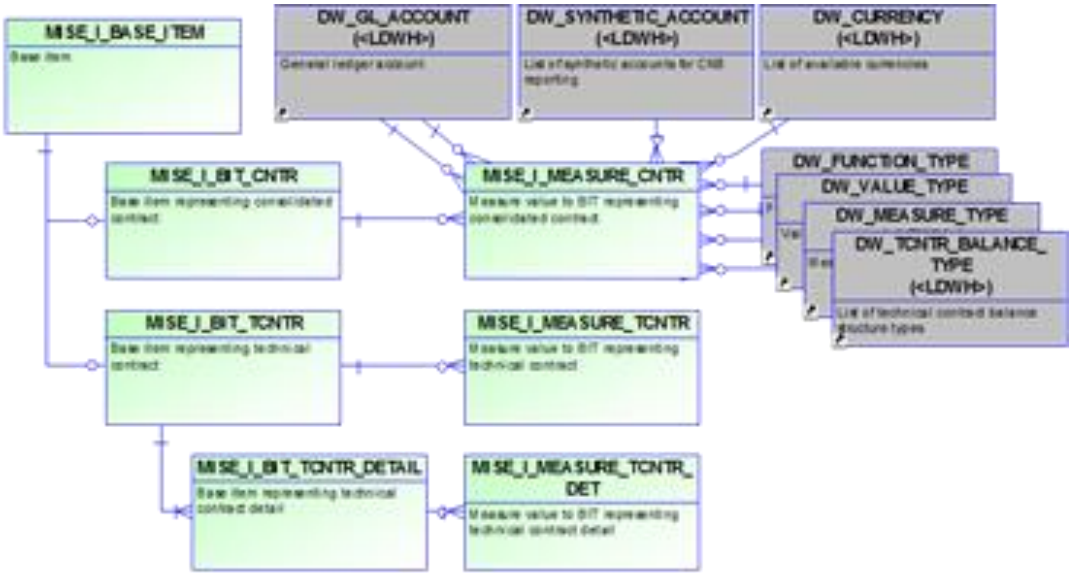


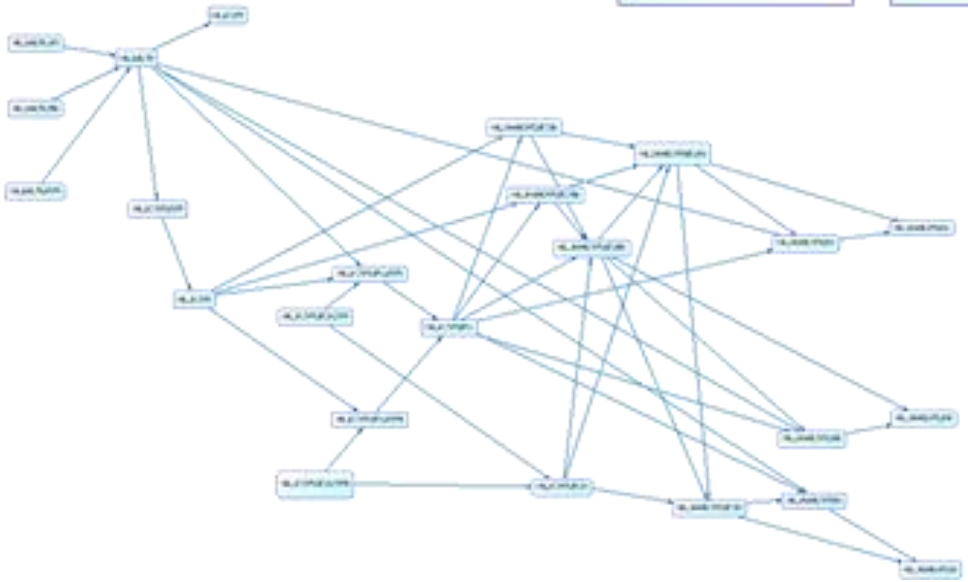
- **Business dictionary**
  - defines business meaning of often used terms
  - e.g. underwriting year, earned premium, claim ratio etc.
  - prevents confusions

Data Dictionary Structure – Example

Data dictionary item	Item description	Example
Entity ID	Entity identifier	E000001
Entity	Logical group of data attributes, typically a database table with a given structure/dimensions	Claim
Attribute ID	Data attribute identifier	A000001
Attribute	Data attribute is an elementary piece of information, typically a column in a table	Date_loss_occ
Attribute description	Business description of data	Date of loss occurrence reported by the policyholder in a claim notification
Source process ID	Source process identifier	P000001
Source process	The process, where the attribute is generated	Claim reporting
System ID	Storage system identifier	DB0001
System	Storage system	Database XY
Data owner	Person responsible for data	Person A
Minimum value	Minimum available value	1.1.2000
Maximum value	Maximum available value	TODAY
List of values	List of available values	n/a
Data type	Number, integer, text, date, etc.	date

Data Dictionary Structure – Example





Process Flow Structure – Example

Process ID	Process	Attribute ID	Attribute name	IN/OUT	Usage in calculation
Source process identifier	Part of a continuous process generating a new information or passing information	Data attribute identifier	Data attribute is an elementary piece of information, typically a column in a table	IN - data attribute entering the activity; OUT - data attribute resulting from the activity	Description of data used in the calculation
P000356	IBNR triangular calculation	A000052	IBNR	OUT	
P000356	IBNR triangular calculation	A000001	Date_loss_occ	IN	Date of accident is used as a dimension in the triangulation calculation in line with assumptions of method M0001
P000356	IBNR triangular calculation	A000002	Loss_amount	IN	Based on loss amount already reported, the not reported losses are projected using a triangulation method

- Farther information about:
  - Process owner / responsibility for the process
  - Process description

Organisation and DQ Roles

Data Quality Roles

Vision

Sponsor

Data Quality Manager

Coordination
Translate the vision to actions

Consistent business meaning and usage rules

Data Owner

Data entity 1

Data Owner

Data entity 2

BI Analyst

Data delivery

Sign-off

Data handling

Department 1

Process Owner

Process Owner

Process Owner

Data steward

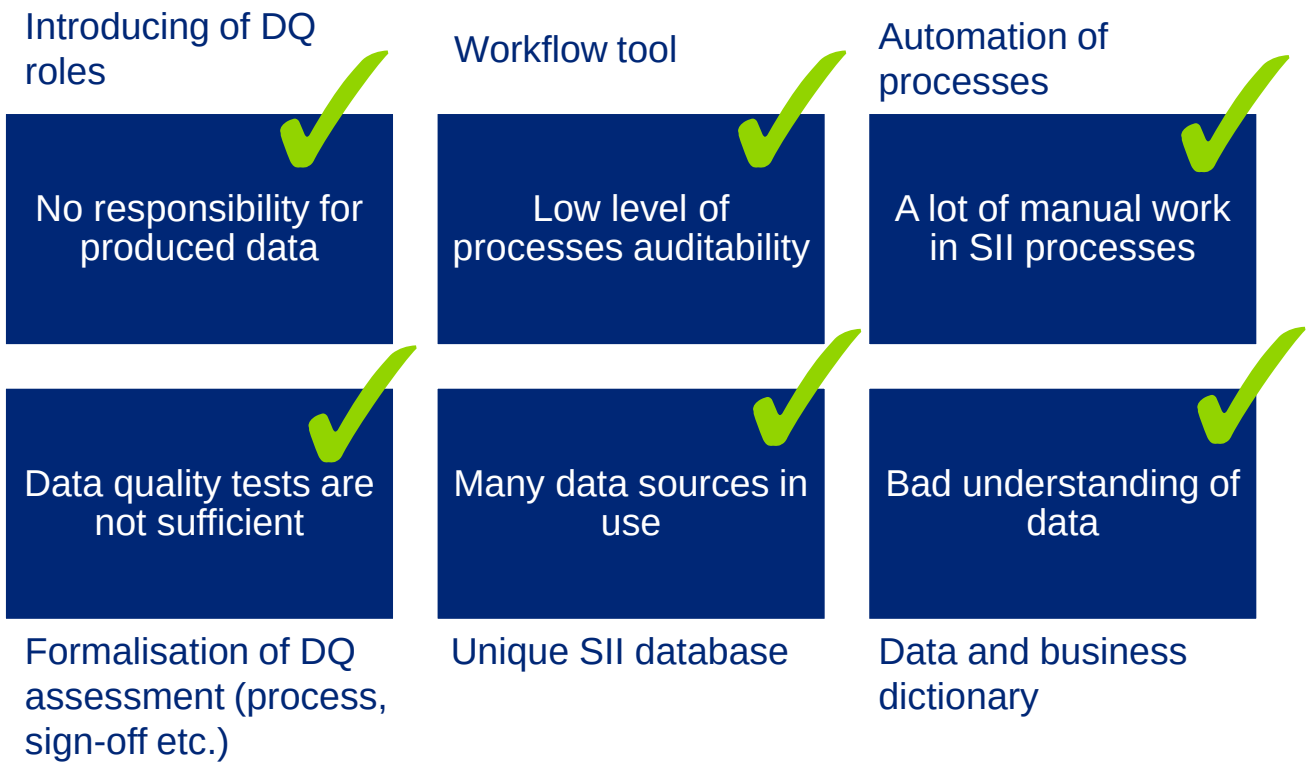
Department 2

Process Owner

Process Owner

Data steward

Common Data Quality Issues



Sources

Sources of Solvency II Requirements

Level 1	Level 2	Level 3
Directive 2009/138/EC, 25 November 2009	Draft Delegated Acts Solvency II, EC, Brussels, 10 January 2014	Guidelines (pre-consultation process) – binding on comply or explain basis

Sources of information for the development of Level 2 Implementing measures (preceding Draft Delegated Acts):

1) CEIOPS, CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: Technical Provisions – Article 86 f, Standards for Data Quality, Frankfurt, 2009
2) CEIOPS, Draft CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: Articles 118 to 124 Tests and Standards for Internal Model Approval, Frankfurt, 2009

Guidelines on preparing for Solvency II

3) EIOPA, EIOPA Final Report on Public Consultation No. 13/008 on the Proposal for Guidelines on the System of Governance, 2013
4) EIOPA, EIOPA Final Report on Public Consultation No. 13/009 on the Proposal for Guidelines on Forward Looking Assessment of Own Risks (based on the ORSA principles), 2013

Other relevant guidance:

5) EIOPA, Proposal for Guidelines on Solvency II, Frankfurt, 2012
6) EIOPA, Draft proposal for Level 3 Guidelines on External Model and Data, 2011
7) Groupe Consultatif Actuariel Europeen, Exposure Draft of Groupe Consultatif Actuarial Standard of Practice 2 (GCASP 2) – Actuarial Function Report under Directive 2009/138/EC, 2012