MOODY'S ANALYTICS

WHITEPAPER

APRIL 2013

Author
Brian Heale,
Senior Director

Contact Us
For further information please contact:
Brian.Heale@moodys.com

Alternatively, you may contact our customer service team:

Americas	+1.212.553.1653
Europe	+44.20.7772.5454
Asia-Pacific	+85.2.3551.3077
Japan	+81.3.5408.4100

Data governance best practice: smoothing the way to Solvency II compliance

In the first in a series of papers, which focus on key data topics, Brian Heale gives his view on the obstacles that insurers have to clear in order to get their data management and governance house in order.

1. Introduction

With successive waves of new financial regulation such as Solvency II, Sarbanes Oxley, Dodd-Frank and international financial reporting standards (IFRS), insurers can be forgiven for feeling overwhelmed by the attendant costly and complex compliance measures they have to establish. And, while Solvency II may until recently have put other regulation into the shade, its delayed implementation doesn't mean that insurers can safely postpone implementing their own data governance and management processes. The imperative for good practice, from both a decision making standpoint as well as the wider regulatory compliance perspective hasn't gone away.

With the delay in Solvency II, insurers might be tempted to postpone addressing some of the data management obligations of other regulatory regimes. However, a better strategy might be to take a more holistic approach to tackling all of the regulation in this area rather than one piece at a time: a sensible approach when one considers that insurers are already making extensive use of data, both for regulatory and business decision making purposes. Insurers would be wise to start aggregating their data, and ensuring its accuracy and consistency across lines of business now, rather than waiting for Solvency II implementation date to be announced. Indeed in March 2013 the European Insurance and Occupational Pensions Authority (EIOPA) issued a series of new consultation papers designed to reinvigorate the Solvency II initiative. In essence EIOPA are proposing a phased implementation approach to Solvency II based in the proportionality principle with larger insurers having to meet most requirements by 2015. Smaller insurers are to be given an extra year to comply.

In this paper we look at Solvency II as the prime example for scoping the requirements for good data management and governance, but the principles apply to all the analytical data an insurer requires. It is also worth mentioning that Solvency II type regimes are being implemented outside Europe—Japan, South Africa and Mexico are examples, and the North American Insurance Council (NAIC) in the US recently introduced own risk solvency assessment (ORSA) type legislation.

While many insurers are relatively comfortable with the quantitative aspects of Pillar I for Solvency II, the prospect of implementing Pillars II and III presents far greater challenges. A central requirement of all three pillars is the large amounts of analytical data needed for calculation and reporting purposes. Insurers have traditionally focused on operational and transactional data (policy, CRM and claims data) and they now need to focus on more analytical data (actuarial, finance, asset and risk data). EIOPA has set out additional base requirements detailing what a Data Quality Management Framework and Policy should look like within a system of governance. The Omnibus 2 directive will undoubtedly build on this.

The FSA recently stated that "Most firms underestimated the time required to embed data governance processes into business as usual (BAU). Dependence on complex IT implementations to support data governance also caused delays at some firms. Typically firms have had to recruit additional resources to manage BAU data governance activities".[1]

These multiple challenges in meeting the data requirements are compounded by the fact that analytical data typically resides in a vast array of systems. Those systems probably have no common data model, and so extracting and transforming the data to be consistent across all lines of business can be challenging. Equally the issue of the poor data quality is inherent in the core systems. Some systems, particularly in the actuarial arena, are desktop-based and are supported by a sea of unwieldy spread sheets.

Collating, aggregating and storing the data and making it accessible in a digestible form for regulatory reporting (such as the QRT templates) further add to the task at hand. So too does the requirement for the qualitative data needed for the Solvency Financial Condition Report (SFCR), Report to Supervisors Report (RSR) and ORSA, and these must also interact with the quantitative data.

Solving the analytical data problems are key to successful Solvency II and Integrated Risk and Finance programs but also to achieving a better understanding of risk and capital within the business, as this will lead to improved profitability.

2. Regulatory Requirements

We have already highlighted the fact that the insurance market has been subject to a whole range of regulatory initiatives including with Solvency II and IFRS having the greatest impact. Solvency II, in particular not only introduces a new risk-based capital culture but also a comprehensive regulatory reporting regime backed by demonstrable risk management practices. Consequently many insurers are looking for synergies between their Solvency II and IFRS programs and for ways that the two projects can integrate together. Indeed some insurers are already embarking on integrated risk and finance programs.

The quantitative requirements of Pillar I, while challenging, are well known and within the capabilities of most insurers, and because they excel in producing numbers, calculating the actuarial, financial and capital metrics required for Solvency II and IFRS, this won't be a major concern. Pillars II and III of Solvency II will however present a much greater challenge. In particular, the data management and reporting provisions are extensive. These involve not only new data sets and reports, but data management standards and controls which must be transparent and fully auditable. Indeed EIOPA requires a Data Quality Management Framework and Policy[2] to be in place as part of the IMAP process and this is also relevant to the ORSA.

3. EIOPA Data Management Requirements

Good data governance and practice should already be in place for most insurers as part of their compliance program. However EIOPA also requires insurers to implement a dedicated Data Quality Management Framework and Policy. The purpose of which is to ensure that all data utilized for Solvency II purposes is accurate, complete and appropriate and sets down standards for data quality. A practical problem is that often users are not always able to define 'accurate', 'complete' or 'appropriate'. The FSA noted that this was a particular issue with catastrophe exposure data, where underwriting teams did not always have adequate understanding of the quality criteria or the point at which a data error could be considered material. The table below gives EIOPA's interpretation:

1 CP 43 and 46
2 Solvency II: internal model approval process data review findings Sept 2012

EIOPA Data Quality Requirements

Accurate	» Free from material mistakes, errors and omissions » Recording is accurate timely and consistent » High levels of confidence » Credibility demonstrated though usage in decision making process
Complete	» Allows recognition of the main homogeneous risk groups » Sufficient granularity to identify trends and the full understanding of the underlying risks » Sufficient level of historical detail is available
Appropriate	» Fit for purpose » No bias » Relevant to the portfolio of risks of the insurer

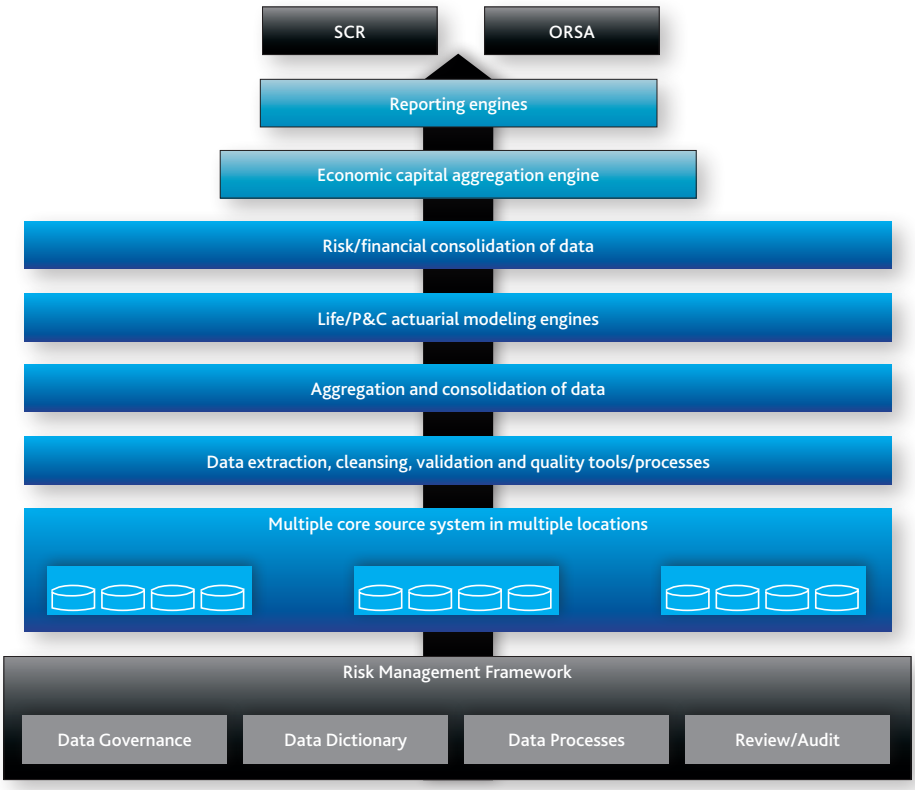
EIOPA will finalize the data management requirements in the Omnibus 2 Directive but sufficient principles are known for insurers to commence their data management projects now.

So what might a Data Quality Management Framework and Policy look like in practice? The following table illustrates a practical example:

Data Quality Framework

1. Policy	Establish a data management policy which details the approach to managing data quality and define realistic and practical metrics to measure the completeness, accuracy and appropriateness of data. Educating users will be important in this respect. Due to the movement and transformation of data from one system to another, it is important to assign data ownership as part of framework.
2. Data Quality Management	Embed a system of data quality management across the enterprise and undertake periodic data quality assessments. This should include a process for identifying and resolving data errors and deficiencies. Cases where data quality has been compromised, including the implications and mitigating actions should be fully documented.
3. Data Dictionary	Formulate a "data dictionary" (as opposed to a data directory) containing all the data entities (and their attributes) used in the internal model/Solvency II processes, defining each attribute's source, owner, characteristics and usage.
4. Monitor Processes	Define and monitor processes for identification, extraction , transformation, processing, and storage of data and how it can be accessed.
5. Lineage	Ensure that data processing and lineage from the underlying source systems all the way though the various layers up to the models is transparent, auditable and demonstrable.
6. Auditor's role	Agree and document the role that the internal and external auditors play in assessing data quality.
7. Audit Trails	Provide audit trails and rationale for data updates particularly when applying expert judgment in lieu of reliable internal or external data.
8. Change Management	Manage changes or data updates which materially impact model outputs.
9. Communication	An often overlooked aspect in data quality is communication across the enterprise. Everybody involved in the processes needs to understand their role and the importance of data quality

One of the most complex areas is auditabilty and lineage. In principle there should be a look-through capability starting from the individual policies in the administration systems through to the various process layers and ultimately through to the standard model/internal model/ORSA calculations. In practice this can be quite convoluted particularly in terms of data trail in relation to actuarial engines (e.g. model points and policy aggregation where the link to individual policies is lost). The diagram below highlights the possible data tree from source systems through to the SCR and ORSA:



For all but the smallest of insurers the data management process will be heavily dependent on technology and in particular extract, transform and load tools (ETL), underlying database technologies and reporting engines.

4. The Real Data Challenges

Practical data management challenges can be split into a number key areas:

4.1 Understanding what data is actually required

Solvency II and IFRS reporting requires a vast array of analytical data to be generated together with high degree of granularity. While some of the data will not be new, particularly from the finance perspective, much of it will be, for instance, asset transaction data and look-through requirements for the Asset QRT templates. In practice the required data will come from a myriad of systems (as well as some external feeds e.g. from investment managers). Some of the data will even have to be "downloaded" from the heads of experts within the organization and manually input. Thus it must be accepted that there will always be an element of manual data input and expert judgment which must be catered for.

The situation is further complicated by the fact that certain data from solo entity and business unit operations has to be consolidated up to group level. Each entity will typically have its own unique systems and technologies. There will typically be an absence of a common data (or metadata) model, with each entity having their own data and governances processes. Group and solos will have different data and reporting requirements.

The problems above are well illustrated by the QRT templates – these require, literally thousands of data items, a high degree of granularity (e.g. the technical provisions and asset templates) and both solo and group perspectives that have to be aggregated and consolidated (for example, equity elimination) at various levels. Complex data transformations between local entity and group require a consistent interpretation and application of group templates, assumptions and standards.

Having defined the requisite data the insurer must proceed down the route of developing the data dictionary[3] which describes the attributes, ownership and usage. This may sound simple but the sheer volume of data required, the number of business entities and plethora of systems involved can make this a very complex exercise for large insurers.

The next step is to understand where the data actually comes from. A key problem here is that in many instances an insurer may have a significant number of systems that contribute data. Many of these systems can be quite old (so called legacy systems) which are typically poorly documented with few people who understand them. Thus extracting the data can be problematic in its own right and even then there is the issue of addressing the poor data quality often inherent in legacy systems.

A further problem presents itself in the actuarial arena where models are often effectively hand coded. Consequently each model may be different in format, structure and results output. In theory this should not be a major problem, but in reality an insurer may have tens or even hundreds of actuarial models to extract and collate data from (e.g. for the Technical Provisions QRT templates).

Control, lineage and auditabilty of datasets and assumption sets that are used by actuarial models must also be considered. These often take the form of spreadsheets which presents another control issue. At a minimum insurers will need an inventory of critical spreadsheets classified by use, by the impact on the internal model, and by complexity.

Some insurers automate their spreadsheet process and this reduces the risk of manual error, but it can also introduce different problems such as reduced oversight, inadequate transparency about the extent of linking and proliferation of nested linked spreadsheets. Linked spreadsheets typically pass only single numerical values, without an indication of the date of last update, creating the risk of relying on out of date data.

One of the key requirements of Solvency II in relation to models is the ability to re-run a model at a future date and produce the same results. This means that the model itself must be stored together with all the assumptions and datasets pertaining to that model, even down to the seed number in a random number generator. Additionally the actual technology that runs a particular model will change over time and may, by the time it comes to execute a re-run have been replaced or upgraded.

4.2 Data Transformation Challenges

The next step in the process is to transform the extracted raw data. This requires knowledge of the systems or source of the data, the end system that the data is being posted to and the purpose for which it is being used. This is effectively a mapping exercise which is generally undertaken using one of the extract, transform and load (ETL) tools - such as Informatica or ODI - or tools built into the database technology. This process can be automated once the initial mappings have been done, and it will become increasingly important to do so as reporting deadlines become progressively shorter.

3 In the context of Solvency II a data dictionary is a list containing a description of all data items together with their attributes and owners that are used for Solvency II and Internal models. The data directory is meant to be a documented repository where different users can go to understand which data is being used in the model, the source it comes from, how it is used, and its specific characteristics.

4.3. Data Quality Challenges

Unfortunately data extraction and transformation is but the first of many stages: data then has to be run though a series of data quality checks to ensure that it is fit for purpose. Data quality checks can either be undertaken in the underlying source systems, in the ETL process or within the data repository itself. Going back to all the source systems and correcting all the data is, for most insurers, impractical due to sheer scale and complexity and age of the systems.

Perhaps the most effective way of executing data checks for analytical data is to execute them within the data repository itself. Executing checks after the data loading allows business users to have direct access to the data quality checks (even if they are not directly responsible for the data quality). This means that they, in their ordinary activity of reviewing, modeling and testing, can have a view of the data quality assessment and validate each single data check. Furthermore, embedding quality checks inside the database enables the maintenance of full traceability.

An important aspect of Solvency II is that it requires a reconciliation against accounting data. Executing data quality checks after data loading facilitates easier general ledger reconciliation—IFRS and local GAAP against Solvency II balance sheet.

A final step in the process is for the data to be approved or signed-off by the relevant qualified person and the approval process tracked and logged for audit purposes. This is particularly important as analytical data is a key constituent of an internal model and SCR calculations. Once approved, the final version of the data needs to be locked down in the repository.

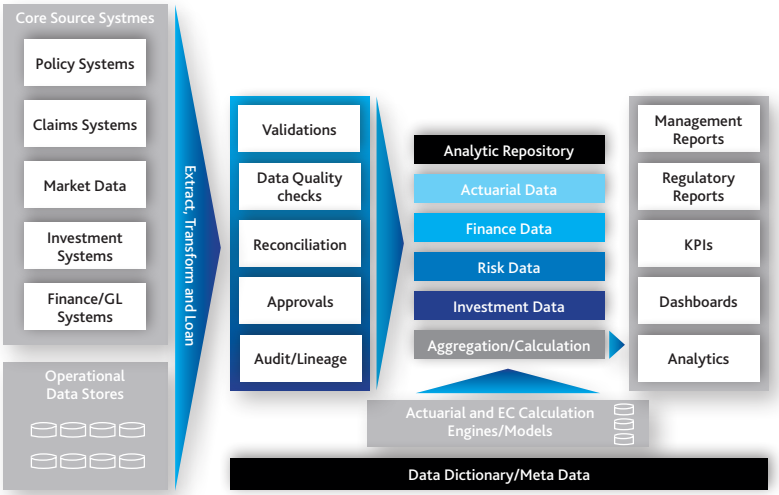
The data quality checks are designed to protect the integrity of the repository. As the process checks the data quality, it attempts to correct the problems. In practice the process has to cleanse the data, remove duplicate fields and finally run a series of contextual "rules" that check the format and content of the data against business logic. An important aspect that is often forgotten is that it is not sufficient merely to have data quality checks in place—you must be able to demonstrate the effective operation of data quality checks to both auditors and the regulator. Thus, evidence of controls and the reporting of issues identified as a result of conducting the check are essential.

Sourcing of third-party data can also be problematic in that sometimes insurers may rely on controls by third parties (for example, investment managers, market data feeds such as Bloomberg and Reuters) without any mechanism to obtain assurance over the control environment and without independently validating the quality of external data received.

Specialist data profiling tools are available and these can also be used to improve data quality and are used in the analysis of data sets for accuracy, completeness and appropriateness using known properties of that data (e.g. outliers, unusual trends and patterns, statistical distribution, consistency with historical data, etc).

4.4 Challenges of Storing Analytical Data

Storage of analytical data can pose a problem as its characteristics differ from operational data which insurers traditionally store in warehouses. While current storage arrangements can be adapted, many insurers are looking for a dedicated solution designed not only to store analytical data but also to undertake additional functions within the database itself. These additional functions might include data being shocked, stressed and aggregated within the database. Further, there might be capacity to store cash flows and loss triangles produced by actuarial modeling engines with a sufficiently high degree of granularity—this is very different from storing claims and CRM data. Except for very small insurers, the solution is to build out a dedicated *analytical data store* as the central source of the truth to drive both regulatory and management reporting. The diagram below illustrates the role of an analytical data store and its integration with existing systems.



4.5. Reporting Challenges

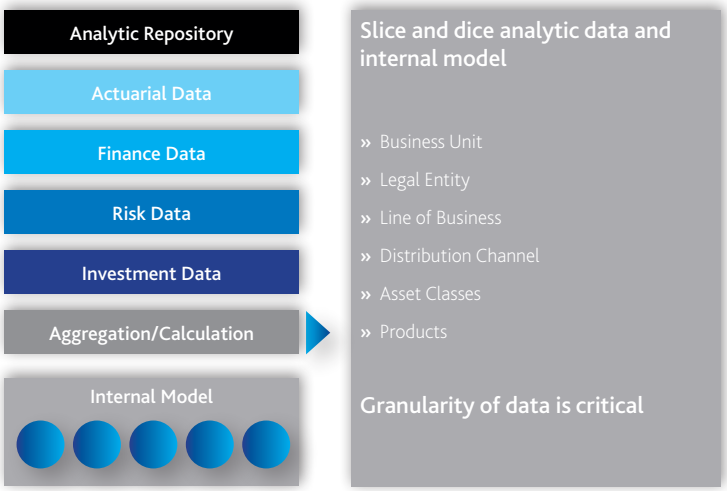
The whole purpose of storing analytical data is to use it to provide information to the business and regulatory reporting processes. Clear regulatory reporting is a major driver. Insurers now need to:

» Have in place by the submission deadlines a series of annual and quarterly regulatory reports including the new QRTs, SFCR and RSR.

» Produce an ORSA document that basically sets out an insurers risk appetite, their risk management practices, and framework, stress test approach and a projection of the balance sheet over the planning horizon.

These reports are defined by EIOPA but local regulators may also specify additional requirements and templates. In parallel with the solvency requirements an insurer will also have to integrate solvency reporting with their IFRS and annual reporting regimes.

While it is important for the quantitative data to be stored with a sufficient level of granularity it must also be possible to slice and dice it to accommodate various reporting formats. The diagram below illustrates the levels of granularity that may be required.



5. Beyond Solvency II and IFRS

Perhaps the greatest challenge of both Solvency II and IFRS is to actually drive tangible business benefits from the respective programs. In practice this boils down to a better understanding of risk capital within key aspects of the business. Additionally insurers are increasingly looking to risk adjusted measures for performance analysis. Thus the calculation and storage of risk and capital data is strategically critical.

The following table is not an exhaustive list but is indicative of the areas where enhanced capital/risk data can support critical business decision making.

Business Benefits

Business Planning	Capital Allocation based on Risk Diversification
Product and Pricing Strategy	Business Transformation and Expansion
Investment Strategy	Profitable Capital Allocation linked high ROE
Optimal use of Reinsurance	Capital better aligned to Risk
Risk Optimization	Mergers and Acquisitions
Alternative Risk Transfer Mechanisms	Maintaining adequate Rating Agency status

6. Conclusion

There is little doubt that data and the associated data management technologies play a major role in regulatory reporting. Furthermore, they can also can add value in the form of supporting the internal business decision-making process. Insurers are compelled to invest in data management to meet the Solvency II requirements and it therefore seems logical to maximize the benefit of the investment being made by extending the project to cover both IFRS and inform better business decision making.

The underlying theme here centers on better understanding the risks within a business and adopting capital allocation strategies that optimize return on capital in line with a firm's stated risk appetite.

Defining this data and implementing tools that manage it certainly requires extra effort and resources but in the long run this is essential for gaining competitive advantage.

At the heart of a successful data management project will be the implementation of an analytical data store that not only holds all the risk/capital data required but also provides the lineage and auditabilty required and tools that extract and improve the inherent quality of the data.

Ultimately, insurers who step beyond the minimum regulatory requirements will be successful – they are the ones who are investing in data today.

SP22812