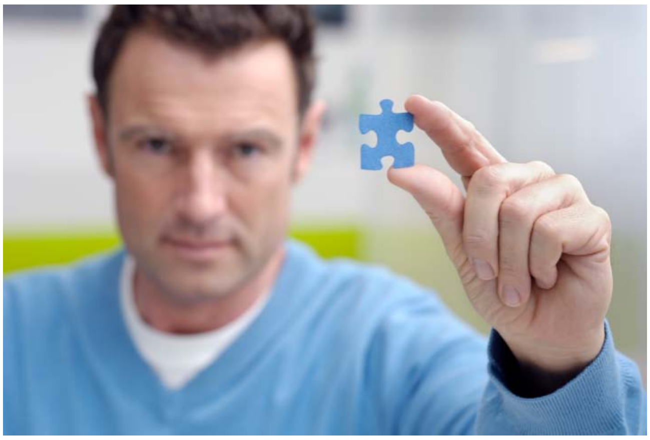
IPL Banking & Finance

An Information Framework for Financial Services Supervision



1 Background to this Document

Modern financial systems have reached dizzying heights of complexity and speed. A phenomenal amount of information is present in these systems, but in its raw form this information is unintelligible to human users with our limited input channels and processing mechanisms.

IPL has an extremely strong capability in Information Management, with many of our staff being industry recognised experts and published authors.

Our stated objective is to use our Information Management expertise and proven technical delivery capability to help customers make sense of this information, meet their regulatory requirements and achieve compliance – an end-to-end approach we have branded 'Intelligent Business'. Our market leading Information Management (IM) and Business Intelligence (BI) capabilities mean we are able to present the most important information to business users in a format that they can readily understand and, more importantly, use to make informed business decsions.

IPL[i] also understands the pressures and challenges faced by modern financial institutions. The ability to identify appropriate solutions and successfully integrate them into our customers' environments is one of our key strengths. We have a long history of delivering the right solution for the problem, whether that be developing a bespoke system, working with subject matter experts and/or integrating third-party products.

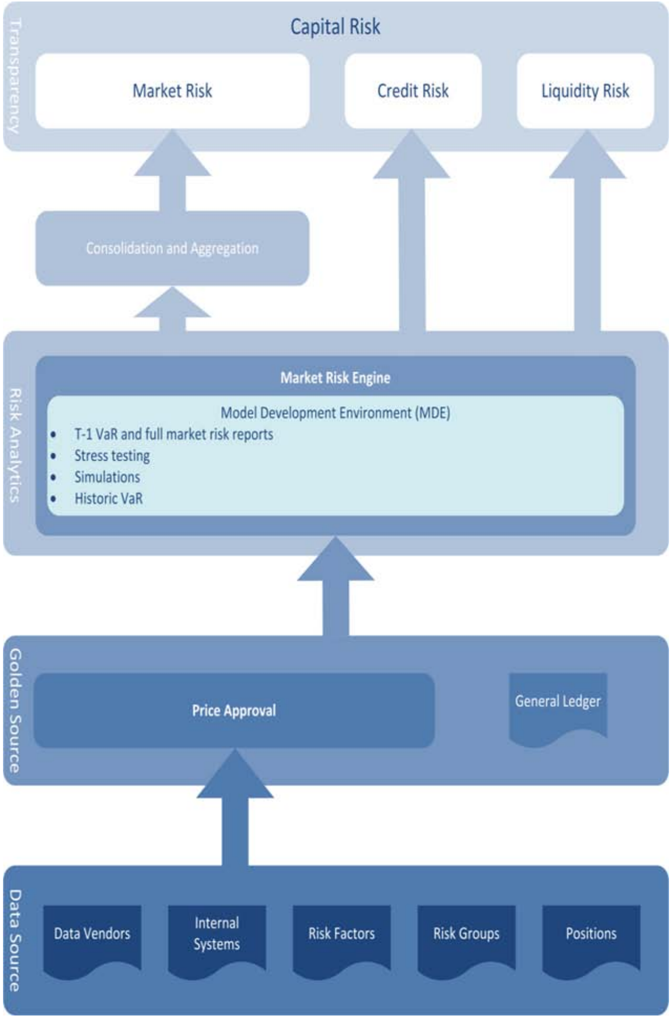
In order to reinforce IPL's expertise in regulatory reporting and compliance, we have developed a strategic relationship with Asymptotix[ii], a highly reputed independent consultancy.

Asymptotix has a unique background in advising financial institutions on how best to tackle risk management and its constituent components. The commercial viability of large complex IT implementations within framework contracts increases in line with Project Risk, which itself is an inverse function of capability or expertise. Asymptotix offers its clients "assurance"; formally in other words, "domain knowledge". Asymptotix, alongside IPL, can apply such expertise to remove risk from your highly complex regulatory reporting and compliance challenges.

Working together with Asymptotix, IPL has developed a vision for an Information Management Framework to achieve 'Transparency' and to meet the tough conditions of Supervision in Financial Services today – a solution blueprint to the increasing problems being presented. Within this framework, risk aggregation, and product risk management in particular, present specific challenges. It is in the process of computation, the organization of computer systems and in the selection of quantitative techniques where the key future challenges lie. In this paper we outline the problems facing your business in understanding and reporting risk in a meaningful and useful manner, and suggest approaches to address these complex issues.

2 Executive Summary

Basel III and Solvency II reporting requirements are, in spirit, oriented towards behavioural explanations of demand for funding capital. By definition, behavioural information within a financial institution must be sourced from operational systems. Given this, Basel III and Solvency II reporting are fundamentally a Business Intelligence and Information Management challenge. In the longer term it seems the regulatory outlook is for more and continuing 'historicisation' of operating numbers to support Predictive Business intelligence. So the data collection and specification process is intrinsically valuable to the success of any Regulatory or Supervisory Compliance project.



Precise estimation of the institutional exposure to capital buffering is fundamental. Thus every quantitative tool you have at your disposal will need to be deployed to achieve this objective. The key source of risk is in the Product Hierarchy, in what can be referred to as "The Tariff Engine"[1] (or as close to that concept as one can get in financial services). That is where the buck stops, since it is in the tariff engine that Risk Based Pricing can be expedited.

Product risk management in particular presents its own challenges stemming from not only the overwhelming amount of information, but also the requirement for appropriate models to be built that accurately represent the risk involved in a sufficiently disaggregated and fully delineated Product "hierarchy". (In reality, the process is more complex than that and we expand upon this below). The model is the basis of the assurance to Executive Management that the financial institution is operating at sufficient margin to be able to contribute to supervisory capital (above all the other calls on operating revenue from normal operations).

There are significant challenges involved in understanding the required computation process and organizing appropriate computer systems. Meaningful calculation of risk together with useful reporting of the results is essential.

We show how tools such as the R language and Business Intelligence suites can help you generate the information you need quickly and effectively. We introduce the process of data modelling, an essential step in processing the vast quantities of data involved in your portfolio. Then we provide a brief summary of the algorithmic elements of risk calculation which sit at the heart of the information management framework, outlined here as a solution blueprint to the increasing challenges of supervision in financial services.

3 Context

Banks (indeed all Financial Institutions) are facing not only a long struggle to rebuild balance sheets amid a worsening global recession, but also tough questions about their future as independent institutions. The credit turmoil in financial markets of 2008 has split out so damagingly into recession that they will be exposed to bad debts on everything from credit cards to business loans. Retail banks have operated with relatively low levels of reserves for bad debts in the past few years, because rising home values and economic growth reduced defaults,[2] although it has now become a "political hot-potato" in

[1] http://www.commercial.hsbc.com.hk/1/2/commercial/customer-service/tariffs/tariffs-index

[2] http://www.ft.com/cms/s/0/5b57ee50-d5e0-11dd-a9cc-000077b07658.html

the courts. Crystallization of losses[3] is the loudest call in the US right now; foreclose as quickly as possible, write off provisions and move on. This will come to Europe as the “Extraordinary Measures” of Quantitative Easing (QE) come to an end, as what the US does, Europe inevitably follows.

The singular focus on sovereign credit issues is perhaps the second shoe to drop following the loss of confidence in financial institutions that characterized 2008 and 2009. Europe and its nagging sovereign woes have become the epicentre of global jitters. In a climate of “unusual uncertainty” the spike in volatility and the volatility of volatility, which has become more of an issue over recent months[4], means that the perceived level of uncertainty in markets is at a level never seen before. Now, on top of all that, we have Basel III (and Solvency II) being implemented stringently.

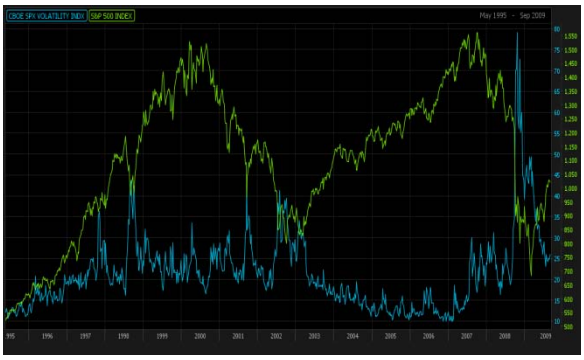
Since Basel II matured and implementation projects for Pillar One came to an end there has been a great deal of focus upon analytic techniques for the quantification of marginal risk or marginal capital (as an aspect of the focus upon ‘economic capital’ initiated by tackling Pillar Two of Basel II [B2P2]). The debate, from 2004-07, centred around banks' adoption of Basel II and their reluctance to invest in internal models. There will always be an economic capital based argument to risk-align capital reserves; Pillar One is arithmetic whilst Pillar Two is econometric. Pillar One is operationally specific; it is derived from the historic asset type exposure matrix of each institution. Pillar Two is holistic; it looks at the overall risk exposure of the entity and at its strategy going forward. Pillar One was in general implemented between 2004 and 2007; Pillar Two was not.

Just look back at the Northern Rock “blow-up” in the UK to see that the FSA did not rigorously deploy the B2P2 supervisory enablement at that time. The UK House of Commons Treasury Select Committee report into Northern Rock highlights failures in Stress Testing and Quantification of Liquidity Risk as the key failures of both Northern Rock risk management and of the UK FSA supervision of Northern Rock. Before the same committee, the Governor of the Bank of the England, Mervyn King, considered aloud the under-specification of Pillar Two before the UK Treasury Select Committee with particular reference to Liquidity Risk; “*Unlike capital regulation, there is no international set of regulatory requirements for liquidity, apart from requirements under Pillar Two of Basel II. At the time when the Basel capital regime was being negotiated the Bank of England did start an initiative to begin a parallel Basel liquidity adequacy regime, and it never got off the ground; other central banks were not so enthusiastic. It is a shame, but maybe we need to get back that.*” [5]

3.1 The New Supervisory Environment

Basel III then requires that the banks implement modern systems which will support financial analytics in an holistic sense; systems which are so holistic that the banks' boards of management cannot use the 'we were not told' or 'we did not know' excuse for risk exposure becoming so excessive that it threatens the stability not only of the financial system but of the real economy.

Whilst no capital conservation buffer existed under Basel II, regulatory requirements under Basel III will require banks to retain a capital conservation buffer of 2.5% as a means of "withstanding future periods of stress." As well as bringing the total common equity requirements to 7%, such a move "reinforces the stronger definition of capital agreed by Governors and Heads of Supervision in July and the higher capital requirements for trading, derivative and securitisation activities to be introduced at the end of 2011.” The capital conservation buffer is to "sit on top of Tier One capital." "Any bank whose capital ratio fails to retain the stipulated limit (which is in excess of the buffer), faces the threat of ‘restrictions’ from supervisors on payouts which include dividends, share buy backs and bonuses." The purpose of the counter cyclical buffer is



[3] http://uk.reuters.com/article/idUKN1711873920101017
[4] http://www.asymptotix.eu/content/inside-volatility-zone
[5] http://www.parliament.the-stationery-office.co.uk/pa/cm200708/cmselect/cmtreasy/56/56i.pdf

considered to be the achievement of "the broader macro-prudential goal of protecting the banking sector from periods of excess aggregate credit growth." Further, the counter-cyclical buffer is aimed at compelling banks to commence build ups of such extra buffers. As is the case with the capital conservation buffer, counter cyclical buffers did not exist under Basel II. Basel III imposes a requirement of a counter cyclical buffer within a range of 0% to 2.5% of common equity or "other fully loss absorbing capital" and will be implemented according to national circumstances.[6]

Every cent, (to the nth significant digit after the point) of common equity (note "common equity") is critical to the financial institution, but prior, it's to the shareholders this matters. Therefore precision estimation of the institutional exposure to that buffering is fundamental. The management in charge of such financial institutions is going to have to get serious about estimating the business cycle and its implications in relation to their risk appetite and risk profile. You cannot do this in an Excel spreadsheet.

Now that this challenge has come upon the financial sector, it looks as though the "Dynamic Provisioning" (DP) approach to the computation of Bank Capital is gaining the ascendancy in the UK and Europe as a solution to ensuring banks hold sufficient reserves to ensure that a credit crisis never occurs again. In reinforcing the tripartite arrangements to Banking Supervision and Regulation in the UK, the Chancellor appears to be favouring this approach which was originally developed in Spain and was perceived to be successful (relatively). One challenge that Dynamic Provisioning addresses is the issue of procyclicality of capital buffers which was perceived to be a problem with Basel II. In effect what DP does is to integrate Basel II Pillar One and Pillar Two by making total capital explicit and transparent over the business cycle. DP requires, however, an econometric model of the business cycle to drive the capital estimates predicated on outlooks for default (as a function of the business cycle). This is an interesting approach since finally an accounting standard is being putatively harnessed to an econometric technique."[7] This today is what "Holistic" risk analytics really means; compliance and supervisory approval just got exponentially harder.

3.2 The Requirement

As stated above, Basel III and Solvency II reporting requirements are, in spirit, oriented towards behavioural explanations of demand for funding capital. By definition, behavioural information within a financial institution must be sourced from operational systems. Given this, Basel III and Solvency II reporting is fundamentally a Business Intelligence and Information Management challenge. In the longer term it seems the regulatory outlook is for more and continuing 'historicisation' of operating numbers to support Predictive Business intelligence. So the data collection and specification process is intrinsically valuable to the success of any Regulatory or Supervisory Compliance project'.

From a Supervisory Perspective then, the quality and usefulness of your Risk Model is entirely dependent upon the robustness and precision of the statistical methods you have implemented in a true Quantitative model. Banking regulators will increasingly expect a sound internal capital adequacy analysis from all banks, and not just those that participate in complex securitization transactions and secondary credit markets. Given today's challenging market conditions, it is a good time to invest in enterprise risk solutions that can provide deep and timely insight into all risk-generating activities, firm-wide[8].

Executive management need to be assured that the institution is conducting its lending business and risk position taking in a manner, or over a hurdle rate, that ensures in-astute practice will never occur again. However, it is in the process of computation, the organization of computer systems and in the selection of quantitative techniques where challenges lie going forward.

The Set Theory, or the Taxonomy of Banking Transparency, is widening to include what are essentially statistical estimates and some arithmetic interpolations of statistical estimates. For example, with a good Predictive Analytic toolset, a bank can statistically estimate its Risk-Weighted Assets (RWA) far faster than the arithmetic approximation of the Basel II Pillar One capital rule. That statistical estimate is also available on demand, likely to be more accurate and is certainly more useful for further analytics.

[6] http://mpra.ub.uni-muenchen.de/25291/1/MPRA_paper_25291.pdf
[7] http://www.asymptotix.eu/node/1247
[8] http://www.asymptotix.eu/content/ubs-transparency-financial-reporting-setting-standard

The holistic[9] analytic solution architecture for banking must be inherently predictive, since transference of credit risk (any risk) has to be on the basis of a robust understanding and quantification of risk in the future. Risk is always with us; we either learn the language to describe it (mathematics) and go some way to approximating it or we are not actually in the risk business.[10]

The banking transparency taxonomy includes accounting numbers but the problem is well known in relation to banking. The standards are not agreed; they are very fluid and very political. All one can say is that the clear direction of travel is towards incorporation of data analytics, calculated values and yes, predictive estimates, far more closely aligned with an econometric background than an accounting one. Confusing BI and Predictive Analytics, i.e. Rear View Mirror Arithmetic with Extrapolative Prediction, is a common error, all too common these days.[11]

3.3 The Current Challenge (Complexity)

In a large enterprise, assets are generally managed by a wide range of disparate systems, often in their own "silos" with little or no connection between them. If there were only two or three systems containing a small to medium size set of data it would not be too hard to design a simple system to extract this information into a data warehouse—a repository of the collected information—for further processing. When the number of systems and the data volumes are much larger then the problem explodes in complexity. Thus, there are significant underlying obstacles to implementing platforms to support holistic Risk analytics today; multiple IT systems - poor integration; partly manual processes; time and resource consuming efforts to keep up with regulation. It may be a cliche but Banks do commonly have complex, geographically disparate IT solution landscapes with complex domains of influence if not control. For the puroposes of this requirement however it is the anlaytical platfrom which is at the centre of our universe. Compliance and Risk Management is often presented as a potentially bolt-on module which you might even get as a free by-product of the implementation of a core-banking solution; to coin a phrase: "Governance Risk and Compliance" [GRC]. However in our view it just doesn't work like that. 'One size fits all' GRC modules applicable the same way in "Big Pharma" as they are in Banking or Insurance, just do not exist.[12]

3.3.1 It's Just not Like that!

So for Risk professionals it is simply not like that. Risk analytics are the core of the banking business and everything else should be driven off that. No-one argues against the proposition that finance and risk management, compliance and reporting will be integral parts of a bank's core business. It is a short step then to strategize towards the integration of finance and risk (numbers or computation).

Integrated Finance and Risk is the buzz-term which both IBM and SAP use now to sell their software into this requirement. Actually meeting this type of requirement in the out-turn generally takes alot longer and is a lot more complex than it may first be presented by the mega-vendors. That is why there is a need for an Information Management Framework for Financial Services Supervision today.

3.3.2 Example: IFRS7

In the specific instance of IFRS7, given its Predictive Modelling element, there are a number of concerns familiar in principle to the audit profession. Models are generally not subject to the input, processing, output and program change management controls traditionally built into typical integrated financial

[9]http://www.asymptotix.eu/node/61
[10]http://www.asymptotix.eu/content/credit-economic-capital-financial-predictive-analytics
[11]http://www.asymptotix.eu/content/risk-consultancy
[12]http://normanmarks.wordpress.com/2010/07/06/why-i-hate-the-term-grc-platform/

packages, significantly increasing the risk of errors. The introduction of numerous intermediary models between sub-systems and the general ledger increases the risk of either information not passing, or passing inaccurately to the general ledger. Input to such models may be obtained from existing data warehouses where data may not be as reliable as core systems. The overall control environment may be weakened through the excessive use of spreadsheets. Finally, moving substantial financial processes from the generally tightly controlled information systems arena into less controlled user domains may result in increased security risks, segregation of duties, data integrity and operational issues.

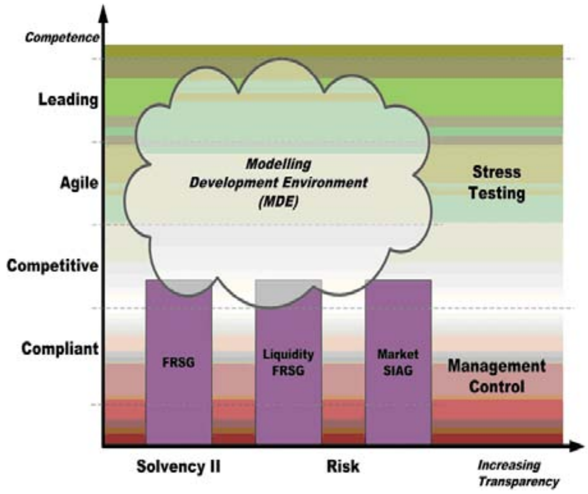
3.3.3 Risk Silos

Financial institutions often use conventional generic BI technology to calculate the risk for each separate area, subsequently combining the risk factors using BI dashboards to gain a view of the organisation's overall risk. This simplistic approach to risk calculation is often spawned from the fact that the institutions operate risk silos – systems that are used to calculate a particular type of risk in isolation. Many of these silo systems have evolved over time because of the way that financial institutions are divided into departments (often with no direct communication channels between them), each department requiring its own systems and only being interested in its own area of risk. The risk calculation will often be performed by a department's core business systems that have been modified to provide risk reporting (overloading the system's capabilities and distracting from its core functionality). The calculation of isolated risk in this manner has several drawbacks:

- Ignores the complex interaction between different types of risk;
- Requires several different systems;
- Lacks consistency;
- Is hard to control;
- Involves duplication of data in different systems;
- Lacks a true overall view of the institution's risk.

3.3.4 Solution Landscape

In a way the mega-vendors (ironically) undersell their magnificent technology platforms[13]. Take SAP Bank Analyzer[14] for example. It is possibly the best designed and engineered layered platform to support Financial Predictive Analytics out there. The Bank Analyzer product is designed to optimally meet the requirements of Legal and international compliance; Basel II, IAS, Accounting for Financial Instruments (AFI) and thus the IFRS7 standard[15] [16].



Alternatively let's take the example of IBM. The tight integration between the IBM Industry Data Models[17], Rational Data Architect and IBM InfoSphere Information Server allows organizations to exploit industry-specific business and technical metadata to accelerate data integration projects such as master data management initiatives or data warehouse development. That two layer model is then fused with the COGNOS BI[18] layered architecture[19] to support an

[13]… because they don't position them to integrate with other specialist tool-sets
[14]http://www.asymptotix.eu/content/sap-bank-analyzer-just-too-complex
[15]IFRS7 is the new Accounting Standard which subsumes IAS32 and IAS39; the IASB is beginning to spell out the statistical methodologies necessary for an institution to comply with IFRS7, effectively it is the accounting standard which reports risk quantities as notes to the accounts, it is the IASB attempting to take turf back from BIS. IFRS7 is a cultural sea change for the accounting profession & is recognition by the accountants that statistical techniques are going to be fundamental to financial reporting in the future.
[16]http://www.asymptotix.eu/content/market-risk-solving-basel-ii-and-ifrs7-siag
[17]http://www.asymptotix.eu/content/building-soa-solutions-industry-models-and-ibm-rational-sdp
[18]http://www.asymptotix.eu/content/ibm-cognos-reporting-risk-cognos-credit-risk-performance
[19]Essentially both SAP and IBM are using Layered Architecture Patterns to present their respective architectures of SAP General Ledger and Bank Analyzer and IBM IFW and Info Sphere. Their presentations look proprietary but they are not, they are simply specific presentations of the Layered Architecture Patterns (LAP) generic approach, which is remarkably consistent across the different organizations using it. Microsoft uses the same technique to present .NET compliant architectures, Ericsson uses LAP for development of GSM software. Each layer is built on top of another more general layer. A layer can loosely be defined as a set of

overall data management platform or environment which will support Financial Predictive Analytics. This is a world beating, if complex, proposition. It can get a financial institution off to a flying start in having a quick win because the industry data models arrive out of the box with 25 years of data analysis and deep thinking in the data model; assets farmed from almost every single one of the great banking names of the past twenty five years.[20]

There is no need to re-invent the wheel. IBM and SAP have been thinking about it for years! It is, however, our experience that no single software solution can solve all the risk management and transparency requirements which are being heaped upon financial services today. If you believe that there is an easy single vendor fix (panacea) then you have insufficient 'scar-tissue'. Inevitably a "one size fits all" approach is in our view inappropriate. Risk Management is just not as simple as that. This is where the IPL and Asymptotix partnership can help you.[21]

3.4 Towards a Solution Design

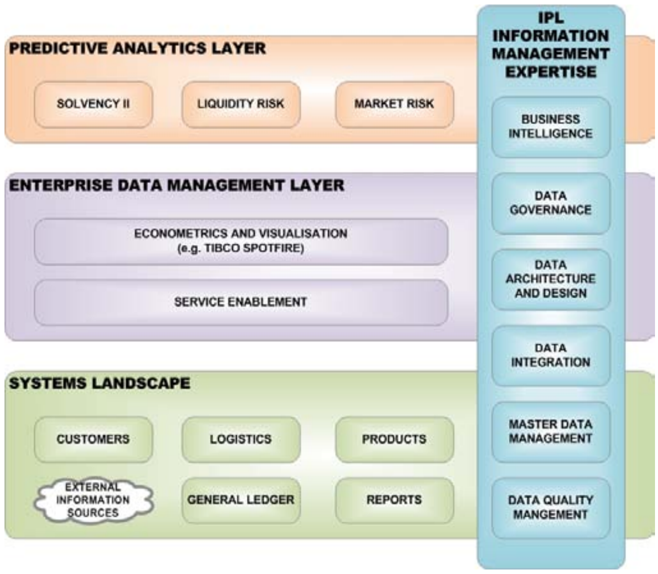
The salient point is that the real constraint on future transparency in banking, the limit of the validity of the best concepts of transparency, is technology. If we are to enable full-blown (Basel III or Solvency II) compliant transparency capabilities in large financial institutions, then Model Development Environments (MDEs) do not fit closely with that universe, or do they? It's the appliance model which fits that on-demand, publish/subscribe-type world. If not archetypal appliances then highly engineered application package point solutions are the route to that kind of transparency. Standard numbers, if not intra-day, then end-of-day (EoD). These appliance-produced core numbers and analytics are then aggregated and blended with other macro data sets to support MDE validation and further outlook i.e. Stress Testing.

The only way to meet the core BAU aspect of that requirement is the appliance coupled with the MDE in the assurance role. The role of the MDE then is in validation and in exposition with the Supervisor - not only does the institution "know" what it needs to know but perhaps even more crucially these days, it can *prove* that it knows. In this context, there is a role for both a classically closed source, appliance model, alongside classically Open Source MDE[22].

The validation process involves running a quantitative model against the same data set as used by systems in production. Building a model that produces similar results to those seen in production has a number of benefits including:

- Validation of the veracity of those numbers;
- A demonstration to supervisory bodies that a bank understands its operations;
- The opportunity to tailor underlying algorithms used in vendor packages;
- Provision of data sets on which stress testing and data mining can be performed.

The following diagram illustrates how the MDE and best of breed solutions combine to meet regulatory requirements and provide a framework for developing valuable competencies.



(sub)systems with the same degree of generality. Upper layers are more application specific and lower are more general. The layered architecture is favored for the design of multi-tiered systems where one layer is built on the services of another.

[20] http://www.asymptotix.eu/important-references
[21] http://www.asymptotix.eu/content/data-governance-risk-management
[22] http://www.asymptotix.eu/content/risk-management-differential-diagnosis

3.5 Institutional Pressure Point

IPL understands the primary requirement of organisations to meet obligations in regards to Solvency reporting. Indeed, the primary driver in meeting these obligations is an holistic approach to solvency risk, with stress testing and contingency funding plans being key tools in demonstrating that a firm fully understands the nature of this risk. IPL believes its vision enables Risk Capital Units to meet not only immediate risk reporting needs, but also to address 'Transparency'. IPL can help Risk Capital Units develop a highly secure Risk Transparency Portal, capable of forward projecting solvency risk, whilst also accounting for liquidity and market risks within an holistic Governance framework.

FSA guidance indicates that firms must take an holistic approach to stress testing to provide increasing levels of transparency. In order to move towards this goal, the vision must consider integration into an overall Stress Testing, Reporting and Governance framework. Taking solvency stress testing requirements in isolation is likely to lead to this framework becoming expensive, unwieldy and potentially non-compliant with future regulations. Therefore this paper proposes an approach which builds on our partners' core strengths to position organisations to deal with the wider implications of taking an holistic approach to stress testing.

A significant barrier preventing the development of holistic models has been the finance industry's purchasing, implementation and integration of packaged solutions to address requirements associated with:

- Risk Management
- Transparency
- Governance
- Compliance and Regulatory reporting
- Supervisory Analytics

As the alternative approach requires banks to hand over all of their information processing to a single vendor, integrating class leading applications ('best of breed') has been used to meet legislative requirements. However this approach is at odds with the development of holistic models, as proprietry, vendor specific approaches inevitably prove incompatible. More recently, a solution architecture gaining prominence in tackling this issue is the use of a Model Development Environments (MDEs). The presence of an MDE at the front and centre of a bank's transparency and governance functions has already been appreciated by the UK and Swiss authorities. Its value is to allow banks to demonstrate that they can validate models used by 'best of breed' applications to produce mandated transparency and supervisory numbers.

3.6 Technology as a Constraint and an Opportunity

The key issue is that the real constraint on future transparency is technology. Transparency is a genuinely very large scale complex data set management problem, since the universe of financial instruments predicates an instrument level data set which satisfies transparency. So Financial Institutions must manage that data set independently of Credit Rating Agencies (CRA), in order to understand how to report transparently that subset of data, at instrument level, which returns confidence in those instruments.

Financial Institutions need to have sufficient systems flexibility to accommodate those changes which are inevitably coming. There are strong regulatory incentives to remove the over-reliance on CRAs and let investors undertake their own due diligence. Risk Management in Financial Institutions, the client applications (of the complex data set) are for the most part either appliance model or Model Development Environments (MDE).

A focused vendor application may meet primary requirements with regards to solvency reporting. However, if additional obligations requiring an holistic approach to stress testing are required, then some form of MDE is needed. By developing this Supervisory Analytics Capability for liquidity, an organisation will be strongly placed to deal with similar obligations baked into future legislative requirements for transparency.

As data is best sourced from operational systems, fundamentally this becomes an Information Management (IM) challenge. IPL has an extremely strong IM capability with a number of industry recognised experts and published authors, and is therefore ideally placed to guide and support organisations in this undertaking.

3.7 The Key Data Disciplines

Addressing the key data disciplines of good information management is an essential component of any attempt to model risk. Over the years IPL has developed a number of tools and best practices to help enterprises validate their capabilities and accelerate development of their competencies. For the purposes of the exposition here in specific relation to Data Management in the specific context of the challenges of Risk Management analytics, we confine the discussion to those elements of Data Discipline which are of relevance. These are:

- Data Governance
- Master Data Management & Data Integration
- Data Modelling, Architecture and Design
- Business Intelligence

4 Data Governance

How do you get a handle on that information resource you are building? (In the Risk Data Management environment). How do you govern it? In the case of Risk management, how do you know that you know that your line management has a handle on Risk data structures? You have to automate the Data Governance process itself; there is no other way. The answer to this governance challenge or, logically prior, the answer to the understanding problem is not in human beings; it is in advanced software.

Data governance is an emerging discipline with an evolving definition. The discipline embodies a convergence of data quality, data management, business process management, and risk management surrounding the handling of data in an organization. Through data governance, organizations are looking to exercise positive control over the processes and methods to handle data. This evolving definition clearly emphasizes the combination of people, processes and technology.[23]

Data Governance is the key to exercising positive control over the management of information. It ensures that the right stakeholders are involved at the right time when changes are made to data resources. Often these Data Stewards are excluded from the system development lifecycle; data governance bakes their domain expertise back into an enterprise's change processes. IPL also has a wealth of experience implementing Data Modelling, Architecture and Design capabilities based upon a number of different modelling repositories. The benefits of this approach include improved impact analysis, re-use of existing assets, support for Data Governance activities and executive level engagement.

Data governance for risk management starts at the level of the supervisor. From there on, the standards are the interface to the individual financial institutions, who can implement their own data governance programs to ensure that everything is run the way it should run.

[23] http://www.asymptotix.eu/content/data-governance-risk-management

Only in this way can banks provide the numbers required to match them with their peers and sufficiently assess risk, and only in this way can the supervisors make sure that the provided numbers mean exactly what they should mean[24].

It is not easy, but the problem is not going to go away by not doing it. People need to be able to trust the figures again, all over the line: managers from different departments inside the bank, as well as anyone external to the bank.

5 Master Data Management & Data Integration

The standard operating model is to capture the data from all the feeder systems and store it in a data warehouse so that risk analysis can be performed on it later. There are several problems with this approach:

- The same data is often stored in different formats in the data warehouse;
- Different risk analysis is performed on the data by different systems;
- The complex interaction between the different types of risk is ignored.

Reporting systems need to marshal large data sets from numerous systems across the IT estate. ETL (extract transform and load) tools have been designed with this task specifically in mind but a good grasp of the marketplace is required. Data virtualisation is another component that can help extract key data from legacy systems obtained through acquisition.

Master Data Management systems help to maintain consistent views of the key entities of the business. Typically these systems are used to manage and consolidate information about customers, products, employees and physical assets.

When it comes to ETL processing, particularly in a real-time or "trickle-feed" environment, Complex Event Processing (CEP) may actually provide a better approach to traditional ETL. CEP provides complex data manipulation directly against the individual record. The architecture is inherently storage-efficient: if a second, third, or fourth application needs access to a particular data element, it doesn't get its own copy. This prevents the unnecessary or reckless copying of source application content.[25]

5.1 Towards Holistic Risk-Data Management

A more sophisticated approach to complete risk analysis is to capture the required data from the feeder systems and bring it into the risk analysis system that uses Complex Event Processing (CEP) to calculate the required risk types. By allowing a dedicated system to carry out the analysis, a complete view of the institution's risk can be produced that takes into account the relation between the different types of risk. The results from the risk analysis can then be distributed to the institution's senior decision makers and the various departments.

Having a single risk analysis system may seem like a risky strategy to adopt. How do you know whether your single system is producing the correct results? This is a problem that should not be ignored but is mitigated by the fact that many departments within the institution will have visibility of the results produced by the risk analysis system. These departments will have their own Line Of Business (LOB) systems that will normally provide some analysis that can highlight anomalies in the risk analysis.

6 Data Modelling – The Product Domain

The data modelling of the product table and its related entities was always perceived as an absolute minefield; you can't do Bank-BI without an integrated Product Table. Product Table Data Design is different from Customer domain modelling in one important respect: it is possible with product to have a bright analyst come into the office one morning and re-design everything you have been thinking for the last three weeks, which is less typical of Customer. Product is where the mathematics is, where the relationships are. In a way the General Ledger is an abstract of Product, filtered by Business Unit.[26]

Your financial institution almost certainly has many different products and instruments, each with a different asset class. In order to understand the risk elements of a financial institution we need a way to

[24] http://www.asymptotix.eu/content/risk-management-differential-diagnosis

[25] http://www.evanjlevy.com/2010/07/complex-event-processing-challenging-realtime-etl.html

[26] http://www.asymptotix.eu/content/new-banking-transparency-inevitably-diy-%E2%80%93-position-banking-line-management

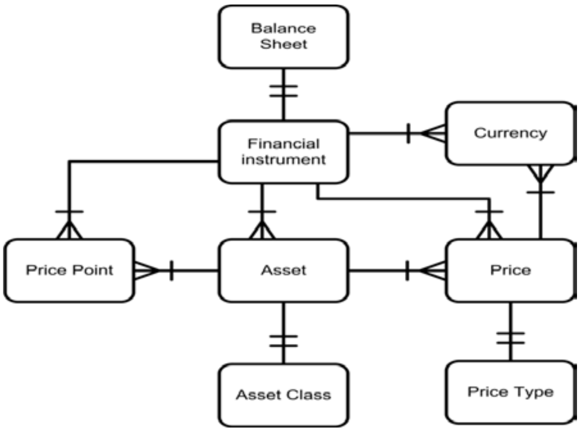
fully capture and model the products. Data models allow us to model the products, assets and instruments to enable us to understand the various risks associated with them. Until the correct models of the structure of asset hierarchies and products have been produced it is very difficult and expensive to identify the underlying data and risk elements. For a suitably large portfolio, it may in fact be impossible for all practical purposes to perform this identification without the data model.

One of the positive side-effects of modelling is that the model assists the visualisation of the interdependencies of risk. Products change, and new, more diverse and exotic products are constantly being added to the product list. However, the basic concepts that apply to all products can and should be modelled.

Products can be modelled at different levels appropriate to the audience. There are three levels at which products are typically modelled:

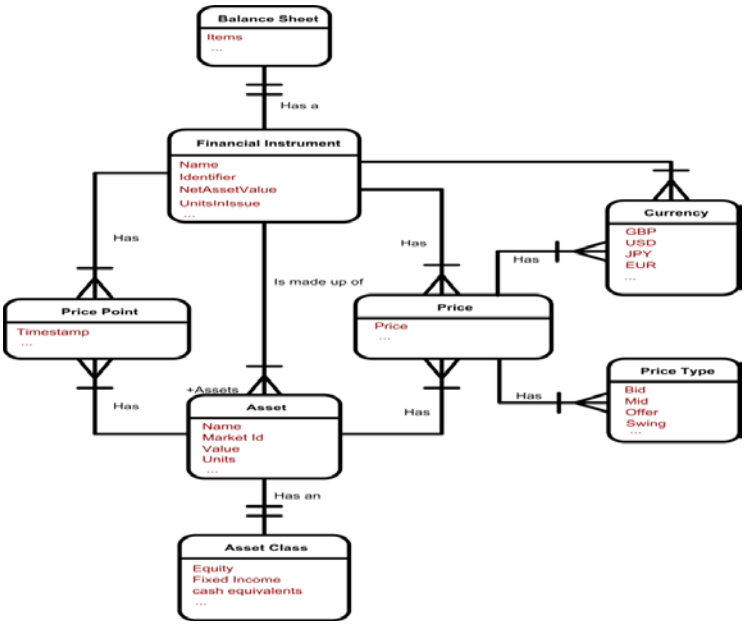
6.1 The Business Level

The business model describes business concepts and entities of which the products are comprised, and the interaction between them. This model will be understood by the business users and will be the way in which the products are viewed within the institution.



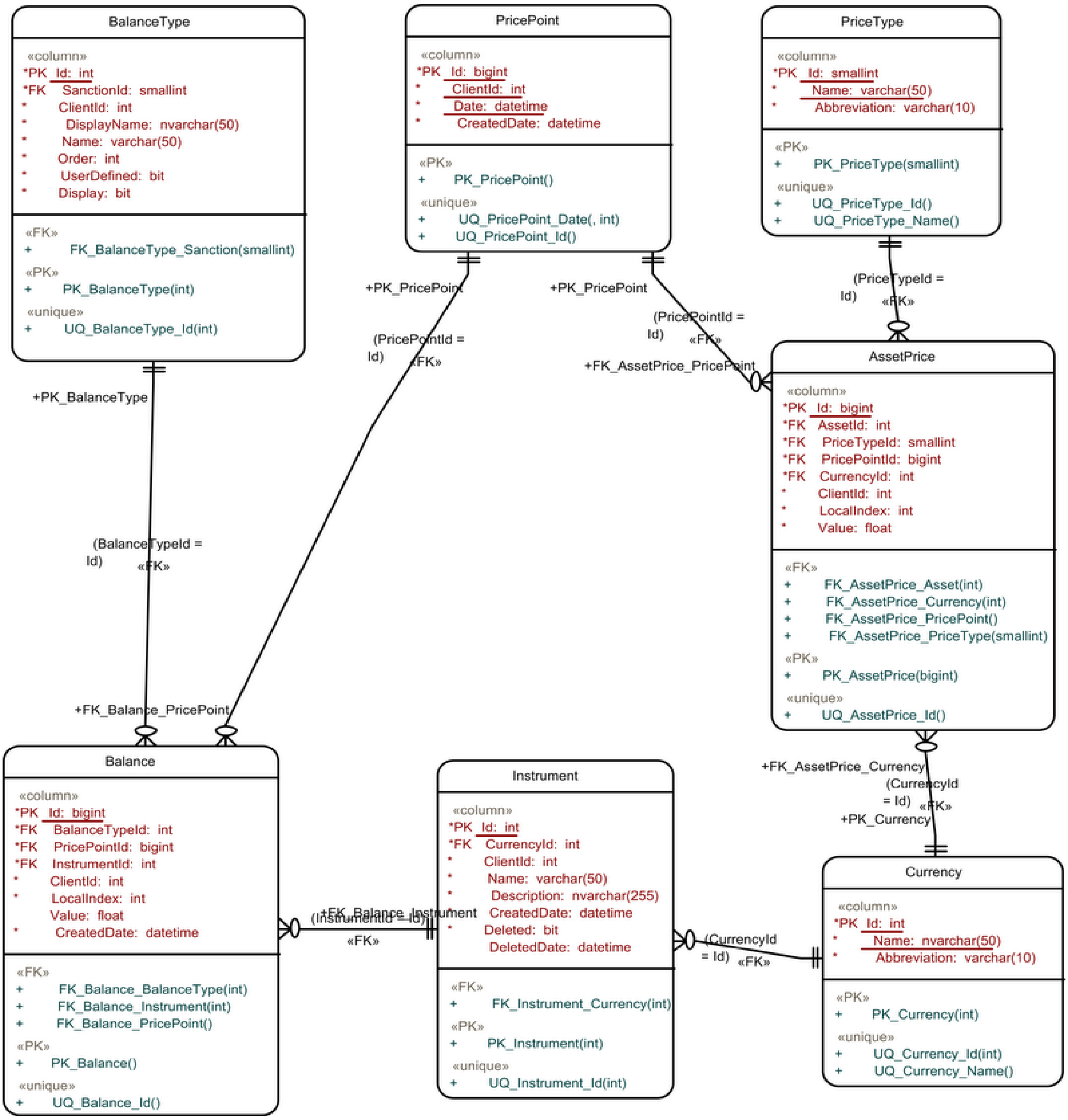
6.2 The Logical Level

The logical model serves as the interface between the business and the physical model (described below). It uses concepts and entities that will be understood by business users but introduces structure that can be used to form the physical model. This level is of particular interest to the process of modelling products.



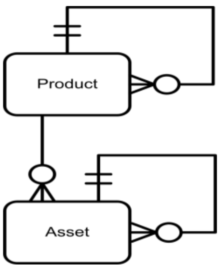
6.3 The Physical Level

The physical model represents the way in which the data will be structured in a database including all interfaces and constraints. The physical model may be used for in-memory databases as well as for representing the data structure of data streams used in Complex Event Processing (CEP).



6.3.1 Challenges in Modelling the Product Domain

It is simple to model hierarchies at the logical level with a recursive link. For example, a product can be made up of assets or other products, and an asset can be made up of other assets. However, when it comes to implementing such a recursive hierarchy at the physical level complications can arise. Although many modern databases will happily support recursion, it is far from ideal and often the database will start complaining before too long. There are several well-known techniques to avoid recursion at the physical level, such as flattening the data structures (typically used in a data warehouse), or forming a fixed hierarchy (i.e. different levels in the hierarchy kept in different tables). Care must be taken when deciding how to model product hierarchies. Each system must be examined on its own. There is no general rule that can be applied across the board.



6.4 An Environment for Data Design

In order to produce accurate and high quality product data models the right environment has to be in place:

6.4.1 Data Modelling Team

It is essential to have a team responsible for the data modelling to ensure the quality and integrity of the models. Team members need to be familiar with the data modelling process as well as understanding the ways in which financial products are structured.

6.4.2 Design Standards and Principles

A set of design standards and principles must be developed before any modelling can take place. These principles will be applied to all the product data models and used as part of the quality assurance process. A set of templates may be used to help ensure the standards are followed.

6.4.3 Version Control

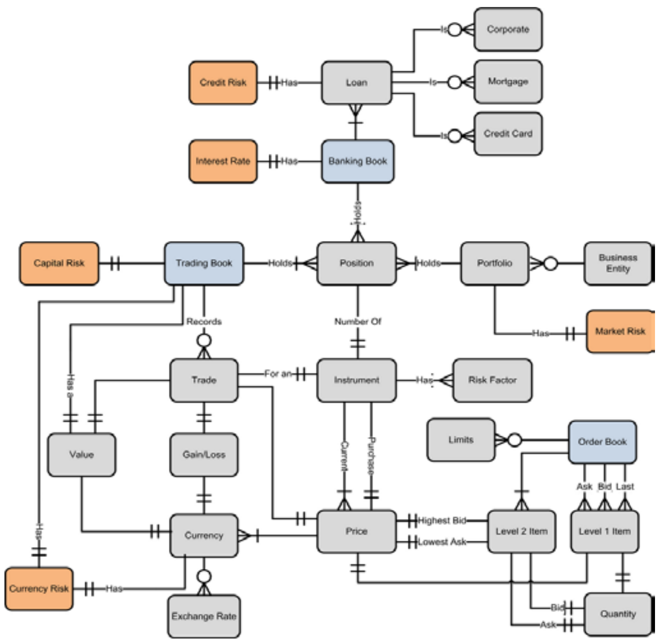
All models must be version controlled to ensure that only the latest models are used and changes can be easily tracked. Preferably, the version control system will integrate with the modelling tool used in the model production process.

6.4.4 Industry Standards

An industry standard modelling tool should be used when producing models. The product model may need to be supported over many years, so the modelling tool will need to be available over the same time span. At the very least, the tool should have the ability to export models in an industry standard format, so that the model can be picked up in another tool.

6.5 Product Domain: Business Level Data Model (Canonical)

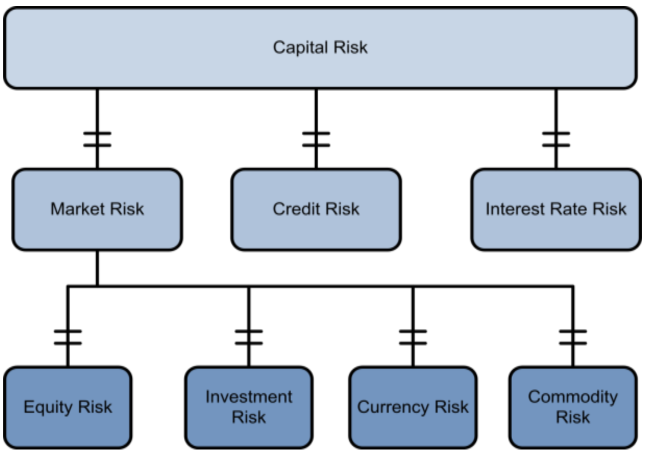
The ability to model financial products is essential for the implementation of a robust risk analysis system. The figure below shows an example business level model of the structures in a modern financial institution.



6.5.1 The Presentation Layer

Modern computing techniques such as SQL streaming—which allows SQL queries to be run upon streams of incoming live data—and Complex Event Processing (CEP)—which integrates events occurring across multiple layers of an organisation—are key to ensuring that risk information can be calculated and presented in a timely fashion and in a way that is suitable for the level of business user. Executing these techniques on High Performance Computing (HPC) grids, and processing the results with advanced BI tools aids the efficient distribution of this information in appropriate forms for the recipients. Senior management and decision makers will often require dashboards that give a high level overview of the entire institution's status with the ability to drill down to fine-grained information where necessary. Department heads, on the other hand, will normally only be interested in the report relating to their own specific department.

In order to gain a better understanding of the risk that a financial institution is exposed to for both regulatory and business reasons, a complete view of the entire organisation's risk is required. To achieve this, the institution must have the appropriate risk systems in place to capture the underlying data from the feeder systems in a controlled manner, to perform the analysis and to distribute the results.



7 Business Intelligence

True Business Intelligence (BI) is more than the construction of mathematical models and raw data processing capabilities. The addition of a rich presentation layer sitting above the processed data assists users in the understanding and reporting of complex data.

BI suites combine impressive reporting functionality with the ability to access and format data from disparate sources. Users can view data "dimensionally", grouping data into sets or time periods.

Integrating a BI suite with R adds tremendous capability, by incorporating the sophisticated statistical modelling and advanced data visualizations that R provides. Introducing the data mining functionality of R into the BI platform brings the best out of two already powerful products.

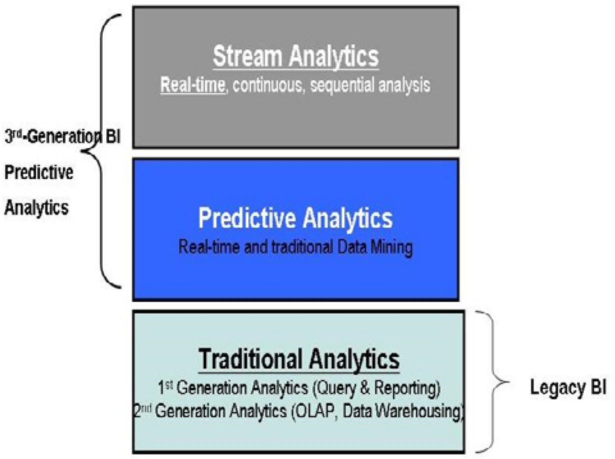
If the credit crunch has taught us anything, it is that you can never tell exactly where the next problem is coming from[27]. While no product can provide 100% protection against the unknown, tools such as R and BI suites that assist with the navigation and exploration of the data enable your business to get both the detailed numbers and the qualitative 'feel' for what the data actually means. This ability to work with the data gives you an edge in spotting nascent problems, so you can take remedial action before they become disasters.

7.1 Financial Predictive Analytics

Treating a quantitatively analytic application like a mashup or a widget is a risky approach. But this is an all too common scenario right now, a consequence of development teams having little understanding of what it takes to make quantitative applications do the job they are required to do; the iterative aspect of quantitative analytics is generally missed. Confusing BI

[27] http://www.asymptotix.eu/content/get-grips-business-intelligence

and Predictive Analytics, i.e. Rear View Mirror Arithmetic with Extrapolative Prediction is a common error made by those with little quantitative background; all too common these days.



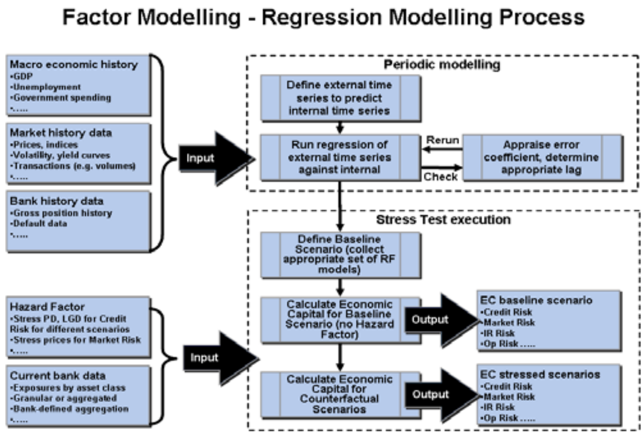
An accounting Number is known as a “first moment”. The associated risk number is known as “the second moment” of that accounting value. BIS is driving accounting disclosure in Banking towards second moments since first moments (it argues) tell us nothing. This nomenclature is associated with Borio and Tsatsaronnis[28], and Phillipe Jorion[29].

In software terms this Supervisory logic has an exact parallel and that is reflected in the exposition of generations of Business Intelligence from the most basic up to the latest available technology. Third-Generation BI (or Predictive Analytics) injects analytical insight into the day to day process of an organization when activity is occurring in real time; it enables broad, real time leverage of insight to achieve business optimization and thus moves beyond “what happened” to “why and what should happen next”. Predictive Analytics requires the marriage of analytical insight with real time business processing. 3rd Gen BI by nature requires a Data Warehouse Platform and MDM system to consume analytical insight, not just source data for BI[30].

Arguably Financial Predictive Analytics is by definition even more complex, requiring even heavier lifting[31] simply by virtue of the requirement to integrate macroeconomic estimates of forward outlook with your internal forecasts of performance. This entails prima facie a large data set and a complex historicisation process.

7.2 The Predictive Modelling Process

Even when you have gathered the raw data together, the calculation of risk is fraught with difficulty. A wide range of algorithms is available for computing and representing product risk; each of them has its own advantages and disadvantages, as would be expected from a process that is turning a large set of data into a small representative summary. Some algorithms are specified for regulatory reporting purposes, but for your internal business processes you may find advantage in using methods that are more closely attuned to the data you have available in your portfolio and the information you need in order to manage your business effectively.



On top of the problem of specifying the appropriate algorithms, working with the data in a productive way is challenging. Econometrics and quantitative analysis are tricky beasts to deal with as they often encompass huge data sets requiring fast, accurate processing. Of course, a computer can be made to deal with these problems using applications developed in almost any programming language. However, some languages are aligned more closely with the problem than others, making the development of the solution much easier.

In the modelling process, focus on key risk drivers rather than trying to model everything "perfectly". If possible, capture the extreme market events that can be deadly if you're on the wrong side; the low probability but very high impact events that historically have occurred when demand has reached

[28] http://www.bis.org/publ/work128.pdf?noframes=1
[29] http://econpapers.repec.org/RAS/pjo72.htm
[30] http://zlin.ba.ttu.edu/
[31] http://www.asymptotix.eu/content/predictive-analytics-finance-heavy-lifting-meat-sandwich

a critical point above supply. Some do this through statistical modelling, while others use stress testing or extreme scenarios. It's valuable to use a combination of such techniques. One may not be able to simulate everything and it may not even make sense to attempt this because of the complex dynamics. However, business intuition and experience can meaningfully guide scenarios or stress tests[32].

Models should be tailored to stakeholders and their requirements. Any modelling effort will fail if stakeholders aren't on-board and convinced that their interests are being addressed. For regulatory or financial reporting alone, the most complex model may not be necessary. For internal use and strategic planning, a more complex model may be useful but only if the business side is convinced that the additional complexity truly provides a benefit in terms of management information.

7.3 Appropriate Tools

R[33] has been described as "a programming language written by statisticians for statisticians." It is an open-source product that can be used as an interactive command-line driven tool for manipulating data and viewing the results live, allowing fast production of useful results, or iteration of concepts when trying to develop a new model. R can also be used to create stored programs that can run regularly or continually on huge quantities of data.

R is a great example of a tool that does not "cut against the grain", as mentioned above. It is designed from the outset to deal with large vectors and matrices as part of its natural syntax, so a developer does not have to make special accommodations to deal with these aggregations of data. Operations which need to be applied to multiple elements of a data set are easily specified in a succinct manner without the overhead of complex syntactical constructions that are required in many other languages. The benefits of a language designed this way cannot be over-emphasised. Developers find such an environment easier to use to create new processes because the language is working with them. It is also easier to spot errors early on because there is far less syntactic support structure required which in other languages can often get in the way of bug finding.

On a single desktop computer in the hands of a good statistician R performs very well, producing fast results. Although the language itself is reasonably simple to learn, the strength of R really lies in the vast array of libraries that have been created over the years. These libraries have been developed and optimised to handle almost any mathematical operation that is needed in business and science, such as multiple regression types and time-series analyses. If your particular scenario needs something that has not yet been defined, then you can either write a new function in R, or for extra performance create a plugin written in another language, such as C or C++.[34]

An extensive community of R users is available to be called on for solving specific problems. Almost always someone somewhere has solved your particular problem before, or knows how to solve it if it is genuinely new.

The graphics library ggplot2[35] is especially useful to the interactive user. This library is easy to use with a sensible set of default options. Graphs from simple x-y plots up to complex three dimensional heat maps are easily produced on the fly. With care, these images can certainly be used in reports and publications. R shows its greatest strength when it is deployed in an enterprise environment for large-scale data analysis. R is designed from the outset to handle data in vectors, allowing operations on every item in a vector as part of its natural syntax. This means that if a task is parallelizable then a program written using the language elements of R will be more readily mapped to run on multiple CPUs, providing an enormous gain in speed, and getting the results you need back to you

[32] http://www.asymptotix.eu/content/credit-economic-capital-financial-predictive-analytics

[33] Sophisticated statistical analysis and charting has long been the domain of expensive proprietary languages from vendors such as SAS, SPSS and Insightful. However, for almost a decade there has been a free OS alternative, R, that has legions of almost fanatical worldwide devotees. The R project grew as a fork from the original S language developed at Bell Labs and commercialized as S-Plus by Insightful. Today, R and its worldwide community of contributors support a variety of pre-packaged statistical models, applications and graphics. Indeed, R's core supports many of the mundane capabilities of data management, manipulation, and presentation pervasive across business intelligence. In addition, R has a flexible API that enables both Java integration for decision management application development and support for leading agile languages such as Python and Perl. The latter can be a huge productivity boost for the data preparation tasks that often plague analytical modelling efforts.

[34] http://www.asymptotix.eu/content/quantitative-libraries-financial-predictive-analytics

[35] http://www.asymptotix.eu/node/353

as fast as possible. The multicore package[36] provides support for parallel processing of R code on machines with multiple cores or CPUs.

7.3.1 The Demand for HPC

As described earlier, It is in the process of computation, the organization of computer systems and in the selection of quantitative techniques where future challenges lie. A crucial aspect of any successful solution is the need to retain the attention and “affection” of users. Within a complex development environment like that necessary for econometric modelling, it is essential that the application is performant, that it functions in reasonably acceptable time scales etc for the user to be able to get his job done without too much downtime if the solution is to become widely accepted.

In some instances the predictive techniques are so computing intensive that they require to run for tens of hours on a standard, even large scale system rack. That is why predictive analytics draws in the requirement for the High Performance Computing (HPC) platform, which means putting clusters of processers together in parallel to run simultaneously at one problem rather than running sequentially at several problems. Good data modelling cannot help time series data; time series just is what it is.

If you can't scale, then you must rewrite what you prototype. It's like inventing the wheel again, requires resources and maintenance and more to the point does not create an efficient loop between an event and a response to it, which causes big business problems.

[36]http://www.asymptotix.eu/content/revolution-analytics-r-revolution-fast-powerful-and-cost-effective-analytics-technologies

8 CONCLUSION

IPL in conjunction with Asymptotix have developed a vision of an Information Management Framework to achieve 'Transparency' and to meet the tough conditions of Supervision in Financial Services today.

The Command and Control Centre, the Situation Room of the future in Financial Services, for the Bank Executive Officer is going to include as many essentially predictive numbers as it does historic, just to keep that executive ahead of what the business is reporting to the market; Market Discipline is taking over as the supervisors put the enforcement pressure on the "big boys". We are moving towards the blueprint painted in very high level but presciently by Claudio Borio and Kostas Tsatsaronis of BIS in their paper of 2006[37]; to design a solution in house by climbing that sheer IP cliff i.e. waiting until your organization looks like MIT before starting to integrate something which might address the reporting requirements of the new supervisory environment is a non-starter. You just don't have the time.

IPL and Asymptotix bring together a range of solid experience in the key data disciplines that support the implementation of large data warehousing systems, business intelligence reporting functionality, and complex product risk management calculations.

There is no need to re-invent the wheel. IBM and SAP have been thinking about it for years! It is, however, our experience that no single software solution can solve all the risk management and transparency requirements which are being heaped upon financial services today. If it was, we would all have had it done by now and that is certainly not the position in the UK or Europe today.

This is where the IPL and Asymptotix partnership can help you. It is a powerful partnership of Information Management expertise, methodological rigour and software engineering capability; coupled with deep domain expertise in Risk Management and Solution Architectures for Financial Predictive Analytics (in fact Asymptotix coined the term). If we can help please do not hesitate to contact us via our web presence.

www.ipl.com

www.asymptotix.eu

9 ACKNOWLEDGEMENTS

9.1 AUTHORS

Andy Cox, IPL

Pat Galea, IPL

Tim Franklin, IPL

intelligent business

J A Morrison, Asymptotix

Brussels / Bath, October 2010

[37] http://www.bis.org/publ/work213.htm