

Analytics, Big Data and the art of the possible

VAT/GST and
data management



Most organisations' tax teams struggle with the accessibility, accuracy and integrity of their data.

Data management is the structured approach to managing these inherent issues, and it requires a sophisticated design and leading edge technology to be truly effective. When considering this it is important to be pragmatic and realistic in what can be achieved, and to weigh up the cost/benefit of each decision, moving away from conceptual discussions (of which there have been many) into the practical.

This publication will

✓
Provide context to the discussion around data management in VAT and GST, why it is important, what has been done and how it is evolving.

✓
Discuss some of the key data management challenges and illustrate some cutting edge techniques being applied to those problems.

✓
Clarify how 'Big Data' concepts and techniques fit into the debate and what these mean in practice.

The historic position



A key dimension of data management is data analytics. Before considering contemporary forms of analytics, it is useful to consider the evolution of traditional analytics in indirect tax.

The first indirect tax analytics techniques were developed by the tax authorities to identify accounting errors. The UK and Canada, followed by the French, Dutch and German tax authorities, were the main exponents of indirect tax analytics. The UK took the concept furthest in the 1990s by dedicating specialist officers to investigate VAT registered businesses in this way and to use the data available to them to identify enterprises evading VAT, e.g. from the suppression of retail sales.

Aside from detecting evasion, the core purpose of traditional indirect tax analytics is to identify errors in the VAT treatment of transactions. All organisations get the posting and reporting of VAT/GST on individual transactions wrong to some degree. The errors generally arise through one or more of:

- Poor master data (counterparty, product, supply chain, rates);
- Human error;
- Failure to properly configure accounting systems for indirect tax;
- Imperfect automated tax decision logic in other systems;
- Problematic interfacing of data between systems; and,
- Fraud.

The traditional analytics methodology looks at transactional records (sales, purchases, inter-company), interfaces between finance information systems and journal postings (in reconciliations). It then seeks to identify erroneous postings using a series of exception reports based on either rules or comparisons of trends and values in summary reports. A few simple examples are as follows:

- A business has a transaction with data showing the following fields: Ship from Belgium – Customer Address Germany – Ship to Germany. Belgian VAT has been posted. In this scenario the 'expected' result is an intra EU supply carrying no VAT – the transaction therefore merits investigation.

- A vendor posts ten invoices for a regular transaction with the same counterparty, nine of which carry VAT and one which does not.
The expected trend indicates that the one exception should carry VAT.
- The VAT return has an output VAT total value which represents 23% of the total net amount value in the VAT return. The domestic rate of VAT in that VAT return jurisdiction is 21%.
This points to an error, since the VAT rate recorded is higher than rationally expected for that organisation.

There are hundreds of these simple tests that have been designed and used over the years. The tests have been very successful for the tax authorities to identify underpayments of VAT and GST – this technique has had one of the highest success rates of any of the tax authority's investigative powers.

The same analytic techniques have also led to organisations offering to help taxpayers identify and recover overpayments of VAT and GST. These businesses often take a contingent fee in relation to the amounts identified for recovery by their clients.

“This technique has had one of the highest success rates of any of the tax authority's investigative powers.”

This approach to analytics and testing has been continually evolving over the years and has been spread throughout the world, with the majority of tax authorities now using these techniques. The explosion in interest followed an OECD publication in April 2010 that provided guidance to tax authorities on the most common testing procedures used elsewhere.

There are, however, two factors that have limited the successful use of indirect tax analytics by the tax authorities.

Firstly, there has been a limited number of skilled staff available to conduct analytics-based audits and investigations.

Secondly, there are significant challenges in the extraction of meaningful and detailed VAT/GST data from many accounting information systems. These challenges include the need to combine the logistics, finance and master data from different modules and tables of a system effectively, in order to provide comprehensive and understandable data for analysis.

In addition to analytics, data extraction and transformation, the other major aspect of indirect tax data management has been reporting. Reporting means the mapping of transactional

data and indirect tax postings in accounting systems through to values in VAT or GST returns. Historically most organisations have relied heavily on standard limited reports from their core accounting systems, coupled with the use of excel workbooks.

There has been a gradual upward trend in the number of organisations improving their data extraction capabilities and using software to manage risk. Similarly there has been an increasing level of sophistication shown by different tax authorities in their approach to tackling fraud and managing risk.

This stage in the evolution of VAT/GST data management has been most relevant to the last 20 years.



The current trends



There have been three real shifts and trends developing in recent times:

- The tax authorities are increasingly understanding the importance of data to them. They have been investing heavily in both regulation (from e-invoicing, e-filing and e-auditing, to regulated transparency, data warehousing and gaining access rights) and their capability (software, people, the use of predictive analytics in the fight against fraud) to get access to better data more frequently and to have more analytical powers.
- While a number of regions and jurisdictions are playing catch up, there is a real trend toward the use of common ‘visualisation’ or ‘business intelligence’ tools. These essentially allow the user to move away from having to look at the transactional detail and focus on key trends and exceptions, displayed in user-friendly graphics and dashboards. They can then drill into the details as required.
- There has been a rapidly growing trend toward centres of excellence for tax reporting (outsourced or in-house), using dedicated tax reporting software and bespoke point solutions to collate, analyse and process the returns and declarations.

The visualisation trend has been driven by the impact of smarter user interfaces. This is a theme of all modern analytics technologies (as the user experience constitutes a very high element of the design decision process) and it is not a trend isolated to tax.

Using vision to think is a well-established principle. Research shows that well-constructed visual representations can dramatically improve a person’s ability to process information and solve problems.

The external visualisation is a substitute for trying to hold detailed information in your memory while working on a problem – see Card, Stuart K., Jock D. Mackinlay, and Ben Shneiderman. *Readings in information visualisation: using vision to think*. Morgan Kaufmann, 1999.

This approach has brought with it some real benefits, and some of the descriptive analytics and management information they provide is a real leap forward.

In contrast, many of the new entries into the market have had a tendency to miss some of the basics that had been working before. For example, the need in the reporting cycle for contained audit trails and resolution strategies, the need to provide more than simple snapshots for analysis, and the often needed separate extraction and transformation layer underneath these tools to deal with underlying data challenges and performance issues.

The key to these tools being a success is to make them work to the practices that need to be embedded and, as with all analytics, to spend the time upfront to access the right data for the right question. This approach is stepping away from solely focusing on the front end visuals to ensure the benefit is gained through integrated and seamless connectivity to the source data, workflow management, strong intelligent analytics being used on meaningful data, with relevant conclusions and the use of good assurance practices. Where the full set of requirements like these are defined from the start, the best software and practices to utilise can then be determined to gain the desired outcomes, not the other way around.

“Spend the time upfront to access the right data for the right question.”

This movement on visualisation and reporting solutions will be relevant for a few more years as the general landscape matures – especially as some big hitters from the software world are releasing new business intelligence type technologies and improvements right now. These improvements will allow for further (more end-to-end) benefits to be gained in this space.

The reality is that these trends are the start of the future journey and the user interface will only get better and better as the technologies improve. However, with the technology advancements and data driven demands to come in the near future, these trends will likely be overtaken quickly by improved and advanced new practices.

Advanced data management



The next step in this evolution is to talk about the future. This is an incredibly interesting topic because it is clear the analysis and reporting needs of the future will continue to get more sophisticated and data from different sources will become available.

To consider this we need to look at both the first movers (mainly multinational corporations) and the tax authorities, in particular those investing heavily in this area.

This starts with the ongoing tax data warehouse discussion. The concept is sound, i.e. that you can create a database of all your tax sensitive data from which you will be able to generate the necessary analytics, reporting and management information required to run a compliant, effective and insightful tax function across the world.

It has taken up a lot of column inches in the last 24 months – in particular with the new country by country reporting requirements (CBCR), standard audit files for tax (SAF-T) and its derivatives, and real time information reporting and data interchange obligations. Additionally, the purported capability of tax software providers and main ERP providers themselves has been promoted heavily. However, there are not many, if any, organisations truly operating a ‘one stop shop’ tax data warehouse. This is because while the technology to achieve this exists and the need for this solution given the scrutiny and obligations being put on taxpayers across the world is high, the functionality, content needs and data challenges are enormous.

What is true to say though is a number of organisations and providers are now on this journey, either in attempting the utopia of covering all taxes in one data store, or more commonly, in seeking to deal with the elements they need to prioritise – often separating the challenge out by the individual tax areas and obligations. This later approach is currently proving to be more cost effective. This is also being echoed to some extent by the tax authorities who have set out certain obligations to capture cross-tax information while still focusing others on discrete areas of tax and duty.

The next topic worth looking at relates to scenarios covering what can be achieved, over and above the traditional use and visualisation of analytics and software-based reporting mentioned above.

In considering this, the key to grasping ‘the art of the possible’ is to assume you could do anything and not allow the thinking to be constrained by any bias of current challenges, boundaries and limitations.

“Assume you could do anything and not allow the thinking to be constrained by any bias of current challenges, boundaries and limitations.”

The practical possibilities are then best illustrated by examples, as follows:

- **Risk and control** – A tax control framework can be greatly enhanced using analytics. For example: the capture and reporting of risk and control metrics and operations; the segregation of duties over tax processes; and, conducting of exception-based analytics as a control in itself.
- **Change management** – By conducting real time analytics this can enable the tax department to rapidly respond to changes that occur out in the organisation. For example: monitoring the risk of creating new VAT/GST registrations, permanent, fixed establishments or nexus status; identifying where new taxes, duties and levies may become due; or, the identification of master data issues requiring remediation and the tightening of the associated controls.
- **Analytics and reporting** – For example, it’s traditionally been difficult to look at chain supplies end-to-end when they go through multiple companies and plants within the same entity; including where supplies can be deemed to have occurred for VAT purposes but no transfer or transaction is recorded in the system for any other purpose. Similarly, with the increase in needs to report more intangibles and recharges for VAT purposes, the data structure and use of it is key to capturing these scenarios.
- **Intelligent or big analytics** – Talked about a lot in theory and there are practical uses, for example:
 - The identification of outlying supply chains used in practice against tax logic which had not envisaged such scenarios occurring in its initial design.
 - The use of regression analysis to determine future VAT cash flow needs under current patterns and the effect of any operating model, process or tax reporting changes on these.
 - Simulations run over trading data to identify the impact of different Brexit models on an organisations’ VAT and duty contributions, potential margin and pricing challenges, and its future administrative costs.

- **Analysing documentation** – The conversion of hard copy documents required to support the VAT treatment of supplies, into data streams. This process can enable automated matching against the associated financial postings, which has previously been manual and labour intensive at best.
- **Fraud detection** – The use of various data sources – both public and private connections, time series analysis and some simple predictive analytics. These and other techniques are used by both traders in their audit defence and by the tax authorities to identify, prevent and detect missing trader activity.
- **Cross tax analytics** – The analysis of customs and trade data and the tie up with transfer pricing. There is a lot of commonality in the data sources and datasets that can be used for customs analysis, VAT/GST purposes, and for profit and cost analysis and allocations for transfer pricing purposes, so efficiencies can be gained. This now also includes the potential use of datasets the tax authorities will gain. For example, the identification of royalty payments from CBCR and SAF-T data could lead to valuation enquiries being raised in the duty area.
- **Using big data** – For example using ‘found data’, i.e. the digital exhaust, by scouring the internet’s vast volumes of traffic, the tax authorities’ activities and the media and social network data, to pro-actively manage reputational risk and the transparency agenda, detect VAT fraud, and/or identify tax authority changes relevant to an organisation.

There’s nothing listed above that hasn’t already been achieved. Further, the tax authorities are already well on this track too, moving rapidly into a digital world with intelligent uses of the data available.

This shift from the tax authorities is evident in the sharing of information they’re already moving on, the trend for pre-validating records – as we’re seeing in Brazil, Russia, India and China – and the use of SAF-T to increase the tax base in several countries.

They are investing in large scale digital programs, where they are working with some of the world’s largest and most dynamic software companies and their solutions.

The authorities have also identified the benefits they could exploit from key emerging technologies like blockchain or other distributed ledger technologies, where the ledger postings cannot realistically be altered or deleted once they’re created. They’re working with several parties to explore the use of blockchain to prevent the possibility of VAT fraud altogether. While the community challenges to enable this methodology are tremendous, the theory behind the idea is good. Aside from this, decoding this massive source of data can give them a lot of information in the fight against fraud

and as for any auditing function, the potential to monitor the flow of financial data in real time using one source of the truth. The use of blockchain technology will expand beyond the bitcoin environment, as it already has into other trading platforms and more recently into transport and real estate documentation, and this will represent a real shift in the availability of data and the authenticity and reliability of the data, for analysis and potentially reporting too.

While there will be a lot of change, it is still important to note that there will be challenges that technology alone, even with the advent of robotics in process automation and data interpretation and machine learning in tax posting validation routines, will not in the near future be able to completely solve.

A good example of this is the need to improve the structure and quality of tax sensitive data being relied upon. The greatest solutions in the world can be useless if the data integrity or structure in which it is captured in the core system is restrictive. These inherent data problems won’t simply disappear because there is access to more data than before.

To meet this challenge it is critical that the process of understanding how data can work includes a critical look at the configuration of the underlying systems, interfaces and processes capturing it, the data structure and possible limitations this presents. For VAT/GST purposes, this means identifying and prioritising the key data points, documenting these where possible in a tax data dictionary or other form, and working with the relevant financial, logistics and systems process communities at this granular level.

There are further benefits of this integration with these communities in any organisation, especially in the embedding and monitoring of relevant controls within these businesses’ processes for the input and maintenance of the tax sensitive data by these teams.

The potential to leverage investments already made in other wider internal financial reporting and enterprise data hub technologies, and resources that could be utilised for tax data management purposes, should also be considered.

Finally, it would be remiss not to mention the new world and the need to secure the data. Examples of private data being exploited are common and the recent raft of new regulations on data security is evidence of its importance. These principles all apply equally to tax sensitive data, protecting the organisation from anyone who may want access to its data and what their motivations might be.

This phase in the evolution of data management has already begun, slowly. It is the future and will have an exponential rise when it really takes a grip.

The challenge and the opportunity



Big Data, as explained above, is only one of many parts in the history and future of how indirect tax data management has developed and will develop further. With the advanced analytics and reporting opportunities and challenges to come, all facets of the challenge need to be taken in context to demonstrate what's necessary and what is 'the art of the possible'.

It is now inevitable that there will be no let-up in the volume and complexity of data-led obligations being pushed onto tax departments by the tax authorities. Equally there will be no let-up in the sophistication and desire of the tax authorities to improve their ability to interrogate the data, access publically available data, drive down the costs of collection, secure transaction under new technologies like blockchain and rely on organisations' own procedures to gain their assurance that taxes are being collected and the tax take is being maximised.

“There will be no let-up in the volume and complexity of data-led obligations being pushed onto tax departments by the tax authorities.”

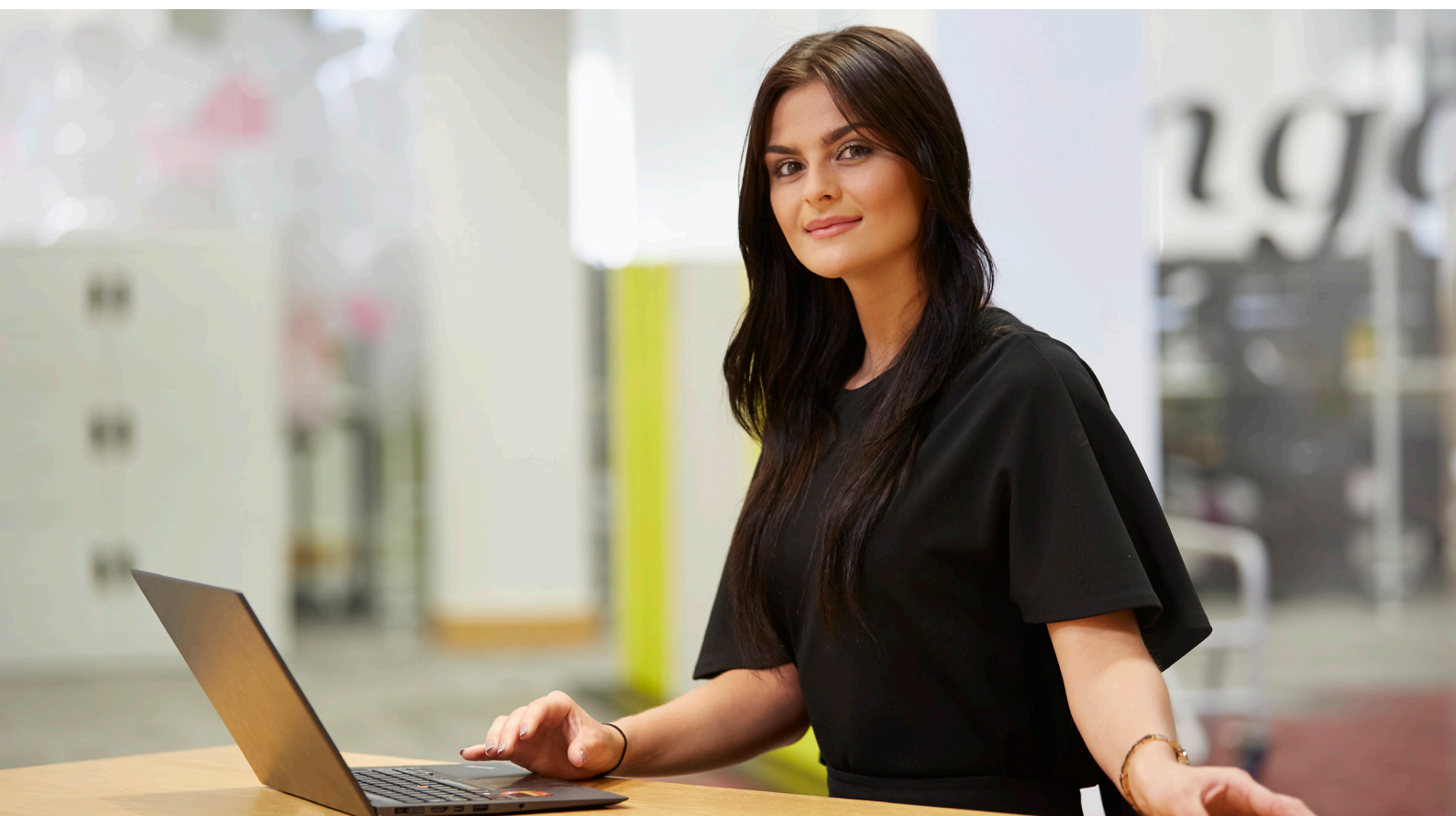
The critical initial question for any organisation is whether it will take a reactive stance and simply do enough to keep up, or whether it wants to do more to future-proof itself and use the data available to gain and provide more insights to add more value to the business.

The yardstick with the tax authorities will be based on how long it will be before they're more tax data literate and thus better than the taxpayer at tax data management, have an organisation's information in real time and consequently know more about the tax-sensitive activities of an organisation than it does.

This raises a separate debate of whether the taxpayer is really entering into a co-operative compliance environment or has simply been entered into a race to keep one step ahead of the tax authorities, while also doing more of their administration and policing for them.

To determine the vision and strategy of the tax function an organisation needs to understand the traditional, the contemporary and the future of tax data management. The strategy should then balance its needs and desires against the reality of what's out there to support it. This will then drive what staff, skills, technologies and sourcing partners will be needed going forward to make it successful.

The challenge, when it comes to implementation and maintenance of the adopted strategy, needs to be understood on several levels. For example: how problematic is the data challenge; how intelligent is the data extraction approach; how much will capacity and performance play a part in the solution(s); how intelligent and agile are the reporting, analytics and continuous improvement techniques being deployed going to be; and, what is the cost/benefit of all this against the vision, strategy and initial objectives of the organisation?



Contact us



For more information, please contact



Tom Birch

M: +44(0)7703 563 631

E: tom.birch@pwc.com

This publication has been prepared for general guidance on matters of interest only, and does not constitute professional advice. You should not act upon the information contained in this publication without obtaining specific professional advice. No representation or warranty (express or implied) is given as to the accuracy or completeness of the information contained in this publication, and, to the extent permitted by law, PricewaterhouseCoopers LLP, its members, employees and agents do not accept or assume any liability, responsibility or duty of care for any consequences of you or anyone else acting, or refraining to act, in reliance on the information contained in this publication or for any decision based on it.

© 2016 PricewaterhouseCoopers LLP. All rights reserved. In this document, "PwC" refers to the UK member firm, and may sometimes refer to the PwC network. Each member firm is a separate legal entity. Please see www.pwc.com/structure for further details.

161027-132445-CR-OS