The Evolution of Data Collection
The role of data collection in tax compliance

Data collection describes the process by which financial data is sourced and collated for tax compliance purposes and the systems used to perform this function. It comes at the very start of the tax lifecycle, preceding the utilisation, reporting and analysis steps that follow, and is fundamental to the success of the tax reporting process.

The way in which data is collected, the amount of automation versus manual intervention, will directly impact the efficiency and accuracy of the computations and filings. It will also affect how effective the tax function will be in meeting the evolving compliance and reporting requirements as tax authorities move to digital tax reporting.

Additionally, data collection and storage affects the tax function’s preparedness in responding to and defending audits in a timely and efficient manner.

Data collection methods can vary widely. Approaches range from manual processes that require the user to gather and consolidate data before rekeying into the tax software; to partial automation, which sees some elements of the process mechanised e.g. by using data collection templates that can be automatically generated and imported back into the compliance solution to gather data; through to full automation using tools that identify and extract relevant data from multiple systems, locations or entities and automatically populate the tax compliance software.

If data collection is carried out manually or is only partially automated, then the process can constitute up to 75 percent of the effort taken to complete tax returns, materially increasing the overall time required. Tax professionals and advisers can spend days or weeks obtaining and querying the requisite data from different sources and/or data owners.

Our customers tell us it’s the most inefficient and costly part of their tax compliance process. Data collection is often not only inefficient, causing delay and diverting precious resource away from value added activities such as tax planning, but the way in which it is approached can threaten the integrity of the data and increase risk for the tax payer. This makes the automation of data collection a priority.

Where businesses currently stand on their road to automating their tax data collection processes will depend on numerous factors. These range from the size of the company to the quantity of its tax obligations and its propensity for change. However, regulatory change means that automation is firmly back on the business agenda.

The problems associated with non-automated data collection

- Inefficient and costly data collection: Manually sourcing, aggregating and collating data is time intensive, taking days or even weeks, with queries going backwards and forwards between the data provider and the tax compliance team.

- Challenges of an evolving digital tax regime: Tax authorities are committed to digitalising tax, mandating a digital journey from source to submission. Poor planning could see the organisation incur higher costs if they make reactionary changes as requirements evolve.

- Lack of visibility: Continuing to rely on data sourced from multiple systems, applications and spreadsheets denies businesses the power over their data to run analytics and provide quality information to support business decisions and underpin compliance.

- Burden of dealing with compliance queries: Manual data collection from disparate systems means that dealing with queries and audits can be a highly onerous task with compliance teams having to laboriously revert to source data found in different systems.

- Increased risk of error: Every human touch on a piece of data introduces the risk of error.
2. Data collection today

Tax professionals, whether operating within an in-house tax team or as an adviser providing an outsourced service, are spending an inordinate amount of time on manual data collection. Typically, the process involves the tax team engaging with multiple departments and accounting systems and accessing data in numerous formats before reviewing this for quality and integrity and then reworking and consolidating it into formats required by the tax compliance engines.

There are five options or system approaches used for data collection today. In order of what we believe to be the most common (and longest established) first, they are as follows:

i. Copy and paste or manual keyed data entry

Is still the most common method used to create computations today and involves copying and pasting from the accounting system or manually produced schedules (including in-house built ‘Tax Packs’ ) into the tax compliance engine. A recent survey1 found that up to 60 percent of those surveyed carry out notice, information/data requests and tax audit management using internally developed software templates such as Excel spreadsheets. It is also the most time intensive which is why this role is often outsourced to a Shared Service Centre (SSC), bringing another team and further risk into the process.

ii. Tax engine generated data packs

Are Excel spreadsheet templates that some commercial compliance solutions can generate directly from the draft tax computation. These spreadsheets are sent to data owners to complete and return, mainly via email. The key difference from the option above is that these packs can be automatically generated from and imported back into the tax compliance engines. However, relying on email for transferring such data is now viewed as less than ideal due to security concerns.

iii. Trial Balance import

Is the method whereby the output from statutory accounts software, i.e. the trial balance or similar, can be directly integrated through providing the capability to map trial balance values to related values in the tax compliance engine.

iv. Online data packs

Provide a replacement for the completion of spreadsheet based data packs. Such products allow data owners to provide requested data by logging in to a secure web portal housed in the Cloud, eradicating the need for emails or attachments. Once the pack is completed, it is automatically imported into the tax compliance engine.

v. Accounting software/ERP data extraction

Uses software to connect to accounting systems and either deliver data directly from source into the tax compliance engine or deliver the data into a centralised data pool, for controlled manipulation, before it is then mapped into the tax compliance engine. This data flow will then provide the capability to drill down to transactional level information making it much easier to review and query original data.

Each option is listed independently above, however, these do not necessarily need to be used in isolation and could complement each other. For example, a trial balance could be imported into the tax compliance engine with the subsequent pre-populated data pack providing structure to any queries which need to be asked around the data.

We explore how these approaches may be combined and the future of data collection later in this paper.

**Use case scenarios**

<table>
<thead>
<tr>
<th>Large Corporate</th>
<th>Adviser</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background</strong></td>
<td>The firm provides services for over 1,000 clients and uses common tax compliance engine(s) to deliver services. The sources of data vary and the firm is under consistent pressure to find efficiencies in their process to provide more competitive fee quotes.</td>
</tr>
<tr>
<td><strong>Automation status</strong></td>
<td>The firm has historically used the format and method their client was familiar with last year. This is commonly copying and pasting from accounting system and ERP downloads or Excel based Tax Packs. Both processes rely on sending Excel files over emails. In addition to security concerns the firm also has no visibility of the data being provided by their clients until the packs are returned. The firm then often asks questions about the data.</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>The back and forth between parties requesting more details and answering questions means that the process can take months end-to-end.</td>
</tr>
<tr>
<td></td>
<td>Preparing questions, responses and the data exchanged at each stage, followed by the need to then manually update the compliance system, means that this consumes more time for both parties than planned.</td>
</tr>
<tr>
<td><strong>Consequences</strong></td>
<td>The firm is incurring more costs than they budgeted for and finds it difficult to ask for a higher fee as their client also feels like they are putting in a lot of work.</td>
</tr>
<tr>
<td></td>
<td>Both parties would rather be focused on value add activities which deliver greater benefit for everyone concerned.</td>
</tr>
</tbody>
</table>
The impact of non-automated data collection

Manual data collection processes directly impact operational efficiency within many companies. Research suggests that 69 percent of CFOs still use spreadsheets to analyse financial data, with 50 percent expressing concerns over whether documents and disclosures have been updated with the latest changes to accounts. These results reveal that current data collection practices simply aren’t capable of meeting the demands required by real-time data access.

The same survey found that almost half of CFOs were unsure if their data was always trustworthy and accurate, reflecting issues surrounding data integrity. Furthermore, 36 percent admitted they didn’t have on-demand visibility of reporting status, with manual checking required every time a change was made. The report concludes that ‘CFOs believe they have a reporting problem, when in fact it is a data problem’.

It’s an issue regulators acknowledge with respect to the tax process: it’s not just the end submission that needs to be automated but the entire chain due to the fact humans increase the risk of errors. According to a report by Bloomberg “Top Tax & Accounting Mistakes That Cost Companies Millions” (2015), 27.5% of businesses blamed accounting errors due to manually entering incorrect data into an enterprise system. Other issues included deleting customised Excel formulas (17%), and overwriting system data with numbers calculated elsewhere (13%), revealing how human error can impact the process.

Impetus for change

Whilst there is a real business need to reduce the workload and streamline the efficiency and accuracy of data collection and management practices, the main driver for change is the digitisation of tax compliance globally.

Tax authorities around the world are under pressure to increase tax revenues, close the tax gap and improve compliance by reducing the risk of error and mistakes. To achieve this, they are digitalising the tax administration, with jurisdictions such as Brazil at the forefront of developments.

In the UK, in continuing their transition toward digital taxation, HMRC is introducing Making Tax Digital (MTD). Under MTD, mandatory digital record keeping and mandatory digital links between the records and the submitted return are required for VAT, compliance from April 2019, with Income Tax and Corporation Tax expected to follow no earlier than April 2020.

Mandated organisations will be required to use MTD compatible software to create VAT returns from ‘source to submission’ and to keep digital records for up to six years. MTD is also expected to lead to more frequent reporting for corporation tax, increasing the need for more efficient data collection.


MTD will, therefore, not only see the entire compliance process digitalised, it will also impact how tax teams handle queries, with the mandatory digital link giving teams the ability to drill down to underlying detail, aiding efficient and effective response.

MTD is proving more costly for business than originally predicted. However, it will deliver real benefits for the CFO, improving efficiency and visibility, reducing risk and increasing accuracy. This should enable them to demonstrate productivity and cost gains by dramatically reducing the compliance workload, while, for tax professionals and firms, digitalisation promises to provide the opportunity to focus on more demanding tasks that require human expertise to deliver tax sensitive strategies.

3. Journey to automation

The need to automate the data collection and tax return generation process is indisputable. How firms and businesses respond to this need may vary. Contributory business factors include an organisations appetite for change, attitudes and IT skill sets, and building the business case to justify investment in tax technology.

Some could choose to take a phased approach in their journey to automation, responding to regulatory change and finding solutions that deliver compliance with minimum change to systems and processes. As MTD expands to other taxes, these businesses will find themselves evolving their data collection capabilities in line with digital mandate.

Other businesses will look on MTD as a major disruptor and a catalyst for change, seeking budget and resource to transform systems and processes beyond what is required by HMRC. They will use this as an opportunity to make process improvements that support data analytics to deliver competitive advantage.

Some may already be fully automated with integrated systems supporting automated data collection. These organisations will have harnessed the power of the available data collection solutions, mentioned in section 1, to collect and exchange data between e.g. ERP and return preparation software. This could be using one or a combination of the following: data collection packs (online or on premise), an accounts integration solution to import P&L and balance sheet data straight from the statutory accounts, or a sophisticated data collection solution that provides connectors to automatically harvest relevant data from the ERP system into the compliance engine.
Review of the automated options available today

For the large corporate or firm looking to embrace data collection automation there are a number of stages to the data gathering process which could be improved upon today. These include extraction of data at source, processing the data into standard formats, importing that standard format into tax compliance engines and generating queries for further information.

We describe the solutions for these stages, as follows:

- **Automated Data Extraction**
  Extracts the data at source. It simplifies the collection of data from multiple ERPs and accounting systems, including structured data from the trial balance and unstructured data and information from non-ERP sources into a data warehouse for cleansing, manipulation and onward transfer.

- **Accounts Integration**
  Automates the import of the statutory accounts from a standard format into the tax compliance engines, with some customisable capability to ensure data templates and reports can be manipulated to accommodate deviations in compliance requirements.

- **Data Entry**
  Automates the process of populating data into the compliance engine and supports raising and responding to queries about the data received by presenting it back in a consistent format and automating the import of any responses.

The diagram below shows how data gets from the ERP system into the tax compliance engines using the different solutions described above:
i) Automated Data Extraction

Automated Data Extraction draws data directly from raw data sources such as ERP and accounting software systems using software connectors installed on customer systems. Data is securely extracted and entered into tax compliance engines to populate the majority, if not all, detail required to form the basis of the computation and return. Automated Data Entry could map data directly into the necessary fields for tax computations. Alternatively, it is possible to map data into a standard template. While this mapping does involve human intervention to confirm where each piece of data should flow, e.g. which tax compliance engine field should a nominal code fall under; Artificial Intelligence is beginning to emerge within the solution, which takes away some, if not all of the human element of mapping nominal codes.

Importantly from a regulatory compliance perspective, Automated Data Extraction is MTD compliant. For MTD for VAT it will extract the data and map this into the compliance engine. A compliance engine, for VAT purposes, can include third party software or spreadsheets, allowing the organisation time to evolve its capacity for automation.

Pros and cons associated with this type of solution are as follows:

### Automated Data Extraction – Pros

- ✓ Futureproof investment i.e. the system is MTD compliant and capable of supporting data input from hundreds of structured and unstructured data sources
- ✓ Eliminates iterative requests as data can flow directly into the compliance engine
- ✓ Low impact in terms of cost and integration. Installation process is often straightforward and solutions can cover a wide range of accounting/ERP systems
- ✓ Mapping enables data to be automatically applied to multiple computations
- ✓ Frees the tax team from manual data entry
- ✓ Enables professionals to review and query data direct from the source, down to nominal codes, without needing to request details from another individual
- ✓ Improves data accuracy by reducing the need for human intervention
- ✓ Clear transparent audit trail that details all aspects of data handling from its raw state through to submission
- ✓ Secures data in flight with the latest encryption technologies
- ✓ Used in conjunction with Accounts Integration gives deep dive electronic breakdown of data

### Automated Data Extraction – Cons

- ✗ Advisers will need customer permission to use software connectors on customer sites
- ✗ Relies on the financial systems in place being up to date and reliable to automatically report from
- ✗ The desire to protect the remit of the tax function could see some resistance to change
- ✗ Automated data collection does not bring ‘tax intelligence’. Tax compliance engines and the tax professionals who use them still need to be able to understand the data journey and question the validity of what they are presented with in order to make judgements on the correct tax treatment
ii) Accounts Integration

A user-friendly automated data collection engine designed to support data capture through mapping data to be exported from statutory accounts system, such as the trial balance, into destination statements in the tax compliance engine.

Effectively, it identifies a ‘category’ by code or name and maps that category and its related value into the tax compliance engine. Data mapping is stored in the form of a template which can apply across the organisation and which enables data to be shared more readily with groups of companies or entities with similar reporting and accounting needs. Data is date and time stamped enabling it to be traced back to source and it’s also possible to perform a drill-back from individual account code line items.

Accounts Integration is the process of mapping from the categorised data in the accounting records to the compliance engine so the data is automatically populated without error. Traditionally, the extraction and categorisation of items will either be done in the accounting system or with Excel acting as an intermediary step in between.

There is clearly the potential for great synergy between Automated Data Extraction and Accounts Integration to get the data out of various accounting systems, into a standard format to create a data pool from where they can be mapped into the tax compliance engines as required.

<table>
<thead>
<tr>
<th>Accounts Integration – Pros</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔ Inserts mapped data into the tax compliance engine in a pre-determined manner</td>
</tr>
<tr>
<td>✔ Flexible mapping from multiple data sources which is transferable, allowing data to be used repeatedly</td>
</tr>
<tr>
<td>✔ Audit trail from source to compliance engine automatically produced</td>
</tr>
<tr>
<td>✔ Easy to use with an intuitive interface</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accounts Integration – Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>❌ Provides peremptory detail only, with no drill down capability</td>
</tr>
<tr>
<td>❌ Needs to be used in conjunction with other approaches in order to complete more complex tax processes</td>
</tr>
<tr>
<td>❌ Separate data request may still be required to gather more granular data to support tax analysis</td>
</tr>
</tbody>
</table>
iii) Data Entry

A simple data collection solution that publishes fully customisable Data Entry Packs, created from the statements in the Alphatax computation, to the secure Data Entry portal hosted on Microsoft Azure, removing the reliance on email for the exchange of information e.g. Excel-based data collection packs.

Once the data provider has signed in to the portal and completed the pack it can be uploaded to the compliance engine, removing the need for any manual rekeying by the tax professional.

The process is complemented by communication alerts informing the parties involved of progress, for instance:

- At the point when the Data Entry Pack is created by the tax professional the system sends the respondee an email informing them that they need to sign in and complete the pack
- Upon completion, the tax professional receives an email confirming sign-off

All changes are tracked and time stamped for auditing purposes.

## Data Entry – Pros

- Removes need for spreadsheets as packs are housed in the cloud. This prevents the potential for spreadsheets to be tampered with by the provider to fit their data layout.
- Cloud-hosted solution eliminates need for locally installed software and offers more robust security than traditional desktop applications
- Using Data Entry avoids the need to rekey information, reducing the risk of error, tampering etc
- An unlimited number of packs can be completed and consolidated into a single computation, easing the process of collecting data from multiple stakeholders
- Packs can be customised with the appropriate statements to meet the requirements of different computations
- Client status tracking enables the progress of pack completion to be monitored and errors or incomplete data to be spotted upfront
- Data is continuously saved into the cloud ensuring up-to-date figures are always available
- An audit trail is recorded for all user actions.

## Data Entry – Cons

- Still a manual process in that data owner needs to input information
- Data entry by inexperienced users can still take time, particularly if those users aren’t familiar with digital signing etc
- Iterative requests are required to query results

## Combining solutions

It should be noted that the above three options are not mutually exclusive and can be combined. For example, if they were all used together, Automated Data Extraction would take data out of the source systems and transform it into a standard format in a central repository. Accounts Integration would then map the useful data from that central repository into respective tax compliance engines as required. Data Entry then gives the user of the tax compliance engine the opportunity to present that data back to the data provider and highlight where further detail is needed. The format in which this is returned through Data Entry enables automated import of further detail. Importantly, this also creates a clear audit trail throughout the whole process.
Benefits of Automated Data Collection:

• **Scalability** – the amount of structured and unstructured data that the tax function will be expected to access on a real-time basis will steadily increase to a level that is currently unprecedented. Businesses simply will not be able to meet this demand using human resource alone. Automation will provide the tools tax professionals need to be able to function more productively with systems and processes that can scale in line with the evolving demands from tax authorities.

• **Efficiency** – the ability to pull data from hundreds of ERPs and accounting systems, including structured data from the trial balance, and unstructured data and information from non-ERP sources automatically, reduces the administrative burden on the tax team. This enables tax professionals to utilise their skills and to focus on what they do best such as exception handling and tax planning.

• **Process control** – the tax function owns the process, ensuring that data is gathered, consolidated and utilised in a consistent and timely manner, without the unforeseen delays and complications that can arise when requesting information from data owners.

• **Accuracy** – data is taken straight from source to submission, eradicating the need for human intervention, and hugely reducing the likelihood of errors.

• **Clear audits** – data repositories retain a full version history and associated audit trails enabling information to be pinpointed across the ensure compliance and computations process and accessed on-demand to satisfy tax authority enquiries and reduce the burden of dealing with compliance queries.

• **Contribution to the business** – consolidation of data into central data repositories enables the data to be used by a variety of tools including analytics software providing visibility. Real-time data then becomes a valuable resource in its own right, helping the business measure performance against KPIs, inform company strategy etc by delivering insight.

• **Future capabilities** – the organisation is in a position to embrace more value added data handling practices, such as error checking and data cleansing through machine assisted processes. Data integrity will be assessed with administrative checking of elements such as data order, repeats, description inconsistencies, consistent group information descriptions and spelling mistakes. This will ensure the quality of the data is consistently reliable so that it can be used as a trusted resource to fulfil multiple compliance demands and carry out business analysis and projections.

### 4. Looking forward

Automation in the future promises not just to digitalise data collection but to make the entire lifecycle self-sustaining. End-to-end automation will see data sourced, stored, cleaned, analysed, selected, retrieved and applied to a computation. It will only come under human scrutiny for expert tax review and sign-off before being sent to HMRC, at which point it will be automatically submitted. In this scenario, the tax professional is emancipated from the drudgery of collating and consolidating data and is free to focus on high level activity such as tax planning.

To achieve this level of automation, other components must be in place. Data must be able to be automatically cleansed to improve reliability and accuracy, reducing the room for error, and it must be mapped to ensure that it can be tagged and retrieved easily from a data pool of information. This will shorten query times and enable the same data to be used for multiple computations without the need for retrieval from source.

Achieving this level of automation may not yet be possible for some but the organisation can begin to take initial steps towards this goal.
5. Conclusion

Data collection is a discipline which for too long has escaped transformation. Time and labour intensive manual processes have monopolised tax team resources, reducing the role of the tax professional to one of a data entry operative as a result of which many firms and large corporates have opted to outsource to a Shared Service Centre.

Manual data collection is not only inefficient, it also puts the organisation at risk of non-compliance. Inaccuracies can be introduced through applying the wrong tax code or simply mis-keying information. These errors can then jeopardise the reliability of the data, the computation and the return. Problems with data integrity then generate requests for clarification from the firm or tax authority in a bid to validate the data until the situation becomes so complex that it is extremely difficult to create a clear audit trail.

Recognising the need for change to make the process more efficient and more accurate, tax authorities are now forging ahead with reforms that will dramatically impact data collection by requiring firms and large corporates to adopt digitalisation processes and systems. At the present time, digitalisation is set to impact the reporting and submissions process but within the next two years, organisations will be required to digitalise data collection too.

Organisations must therefore embrace the tide of change and begin to look at how they can begin to transform manual or part-automated data collection into a fully automated process capable of handling the sourcing, consolidation, utilisation and submission of data with limited human intervention. This automation will confer numerous additional benefits on the organisation providing a scalable and accountable system which will see data become easier to access and more productive, with the automatic population of associated computations, trackable audit trails, and the capability to inform future tax strategy.

One criticism levied at the digitalisation of tax is that it could call into question the remit of the tax professional and even endanger the profession. Yet, rather than displacing tax professionals, automation has the power to give them greater autonomy. Free from the bind of data collection, tax professionals can provide the adviser or corporate with the benefit of their expertise by carrying out high level analysis that adds value.
About Tax Systems

Tax Systems is a leading specialist in tax technology and services in the UK and Ireland. Our solutions automate the end-to-end compliance process, reduce operational risk and increase process intelligence so that expert professionals are able to focus on delivering more value to their business or clients.

Our objective is to help organisations meet the challenge of a changing regulatory landscape by providing proven and innovative technology and services delivered by a close-knit team with rare skills and capabilities in both tax and technology.

We can support you on your journey to automation regardless of how far you are along the path with a range of compliant data collection options, from Accounts Integration for statutory accounts to Data Entry with a cloud-based platform that allows data owners to directly input calculations to compliance engines, to Automated Data Extraction for the extraction of data directly from hundreds of ERPs and accounting systems, providing complete end-to-end automation.

Our R&D department are working closely with HMRC to devise solutions for emerging regulations and we can help you become fully MTD compliant for digitalised regulations such as VAT and future MTD targets such as Corporation Tax.

We also collaborate closely with our customers to help them embrace technology to reap the benefits of increased accuracy, auditability, control and efficiency of reporting, as well as saving time associated with manual compliance processes. We aim to change the nature of relationships between advisers and corporates by freeing up skilled resources to focus on adding strategic value to their organisations.

Tax Systems is a public company quoted on the AIM market of the London Stock Exchange with offices in Dublin, Farnham and Staines. Tax Systems is the creator of Alphatax, the market leading corporation tax compliance software built over the last 25 years with a £60 million investment into research and development. Alphatax is used by over 1,000 customers including 43% of the UK FTSE 100 and 23 of the top 25 Accountancy firms.

Contact us today on 01784 777 000 or enquiries@taxsystems.com to find out more about how we can assist you in the automation of your compliance data collection systems and processes.

www.taxsystems.com