Book Industry Communication is shortly to launch its Metadata Capabilities Directory (MCD) after testing and trialling with a small group of 'Pioneer' members. The new facility, which is part of BIC's 'Metadata Map' project, has been under development for the last couple of years, but it is now hoped to be rolled-out to the entire BIC membership within the next few months.

So what is the MCD and why has it been created?

The project originally started as a result of repeated questions raised at BIC's Metadata Committee. Whilst it was recognised that there were plenty of excellent 'Best Practice' guidelines available for the creation and supply of product data, it was much less clear that we knew what happened to metadata after that, once it had entered the metadata supply chain.

Do we really know what is actually sent by individual suppliers and how? Are we able to be certain what data recipients do with the data when they receive it and apply it to their systems? Or if they send it back out again, what changes are made to format or content? Perhaps there are useful data attributes that could be sent that are not? Or attributes that *are* sent, but are ignored by receivers?

The conclusion was that if we could establish a way to allow organisations to share what they actually do with the metadata they create, receive or pass-on we might learn some practical answers to these questions and improve the data supply chain for everyone.

The original objective was to "*Build a sustainable Metadata Map of the UK metadata supply chain and to share the aggregated results with all participants to provoke positive change in data behaviours by users.*" (LBF launch in 2018)
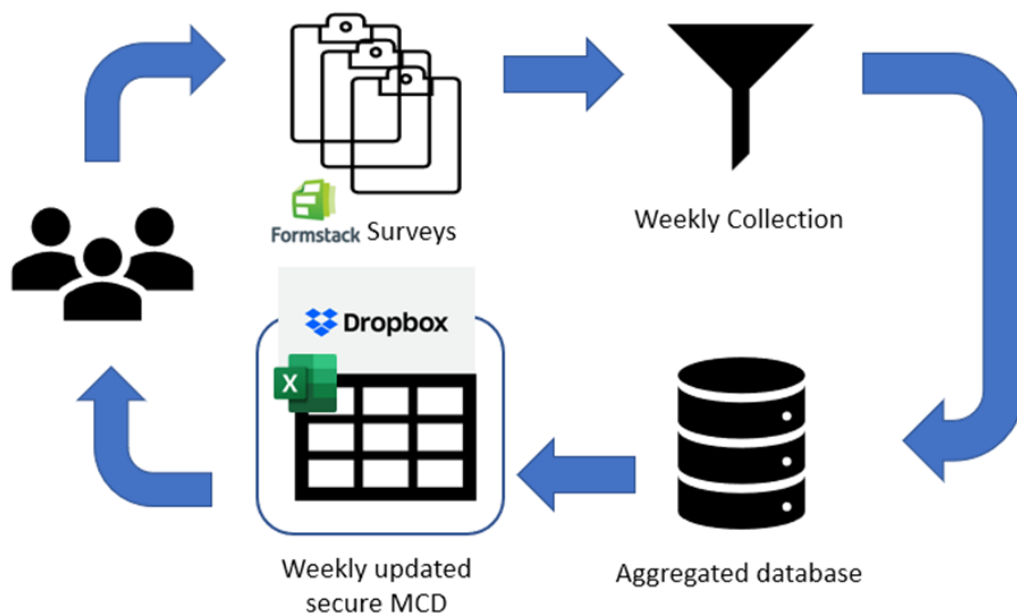
However, we were unsure how much information organisations would be willing to provide about their data activities and how detailed we could legitimately and practically make our questions. How could we best present the resulting information to make it useful to the whole supply chain by providing top-level overview analysis of the data supply chain at the same time as the drill-down detail of organisation-to-organisation and specific data attributes? Perhaps most importantly, did this project have sufficient widespread support across the supply chain to generate useful data? How could we unify the very divergent views on what this project should deliver?

Our ambition therefore needed to be tempered with a dose of realism. We have created a detailed survey for participants to answer which we believe is sufficiently detailed, but practical for all organisation types active in the data supply chain to be able to answer. It is important to note that we are not looking to capture the actual metadata being shared within the supply chain, but rather we are seeking to understand the policies and system limitations that define and shape metadata behaviours within our industry.

We also recognised that we could not justify spending too much of BIC's members money on developing a very sophisticated solution until we better understood what was wanted and what was possible. To find this out we needed to create sufficient trust and demonstrate value amongst potential participants to enable them to share information in the first place.
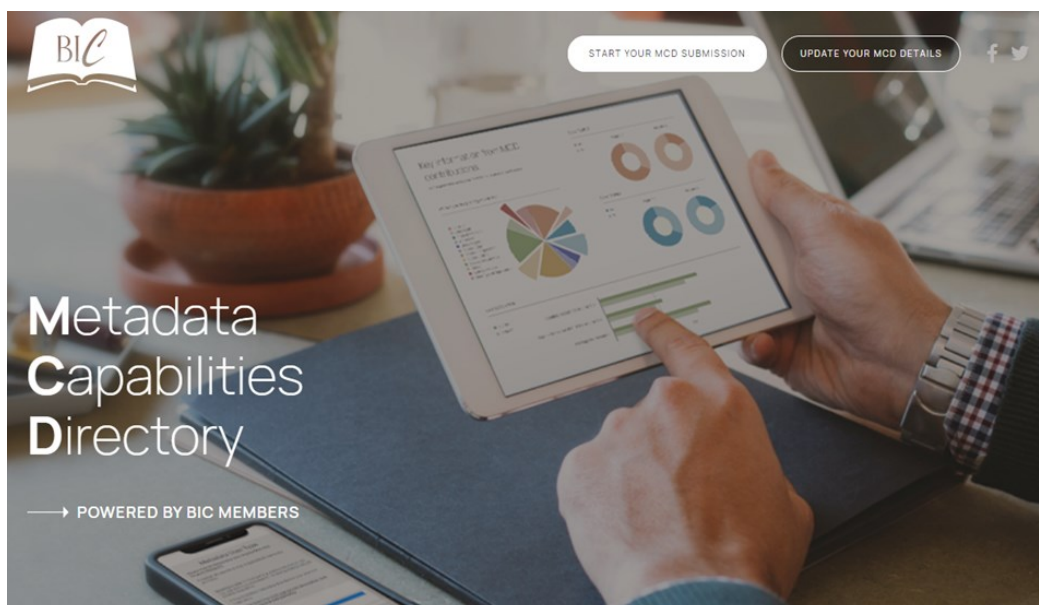
The MCD system

The result is the MCD, a simply-presented summary listing of responses to our survey, detailing the data activities of each of the participants, that is password-protected and only accessible to BIC members who have contributed to the directory.

Participants complete a Formstack survey, covering their full range of metadata activities, whether they are a data supplier, recipient, or both. Newly submitted surveys are downloaded weekly and added into a secure database, from which we extract a simple spreadsheet of all submissions. We make this available securely within Dropbox, so that only contributors can view it.

Organisations needing to update their submission do so by the same process, using an update survey which allows them to change only the information they need to change, which is then collected, aggregated, and re-presented to users in the updated MCD each week.

To attract new participants to the MCD, we will shortly launch a portal webpage within the BIC website which provides an introduction to the MCD, hosts the documentation created for users, provides the links to the Formstack surveys and showcase a few key overview statistics collected from MCD submissions.

How does the MCD work?

The scope and complexity of the product metadata that is used and communicated within the supply chain is significant and it follows that we needed our survey to reflect this to make the MCD useful. Our question set has therefore been designed, and refined by our Pioneer Group testers, to allow a comprehensive overview of the whole range of product data created and moved within the supply chain and to enable true insight into the processes that shape and maintain it. The scope is the same for both Data Suppliers and Receivers, with questions and responses reflecting each perspective. Most questions offer a set of standardised multiple-choice options. This helped to frame the type of responses we were looking for, but also provided a more practical way of aggregating the responses in the MCD. Filtering on standard terms is much easier than trying to make sense of free-text.

Nevertheless, there is still plenty of opportunity for organisations to elaborate or refine their responses with free-text, supplementary boxes available in most sections of the survey. We hope that the MCD becomes a trusted source of comprehensive information for all contributors to use and we believe that this is best achieved by encouraging a spirit of honesty and full disclosure.

We have taken considerable steps to ensure the security of all data provided, which we believe is paramount to the success of the MCD. We have required each contributing organisation to nominate a single authorised individual who takes responsibility for that organisation's submission and is the authorised user to access the MCD file on Dropbox. The spreadsheet file is read-only to emphasise that data supplied into the MCD should only be seen and used within the context of the MCD by all users and to ensure that only contributors to the MCD are able to see its contents. In addition, each participating organisation must create a Unique ID which must be used to ensure any future updates are aggregated with their previous responses. All new submissions are validated by BIC with a completeness threshold required to met before access is granted to the MCD spreadsheet on Dropbox.

Why only a spreadsheet? It helps keep the cost down – we have not needed to pay for the design or development of an interface to deliver the contents – and everyone knows how to use standard spreadsheet functionality like 'Find' and 'Filter' to enable the data to be easily interrogated.

How will MCD users benefit?

We recognise that organisations are unlikely to participate just on the promise of improving the whole supply chain. They will want to know what they can get out of it themselves to benefit their own business and the effectiveness of their own metadata.

Pioneer Group contributors have already identified that simply by taking the survey they have shone a light into areas of metadata use in their own business that warranted attention. We believe however that the main benefit will be by MCD users benchmarking their own activities against other similar organisations, competitors and/or data partners and against the overall trends of all participating organisations, and thereby able to identify beneficial improvements to their own systems and practices. In addition, it may be possible to justify the cost of system development and prioritise the most valuable improvements to that organisation. Nevertheless, in time, we also believe that there may be valuable lessons to be learnt when looking at overall data and trends in the future that could highlight data usage improvements we can all make within the supply chain.

Launch of the MCD

We have already started receiving live, full submissions from our Pioneer Group and more are promised in the very near future. The system is successfully aggregating these and creating the system reports including the MCD itself, and this will shortly be shared with those contributors for the first time. The dedicated MCD webpage on the BIC website is also being updated routinely behind the scenes and will be published when we formally launch the MCD soon. We look forward to inviting BIC members to participate soon, and enabling the expected, shared, benefits of the MCD to become a reality for all our members and the wider supply chain.