

Data Management and Automation: A risk-based approach

Callum Giblett

KBR

Principal Specialist, Programme Management, Government Solutions International

callum.giblett@kbr.com

Charlotte Watson

KBR

Senior Manager, Real Estate Services, Government Solutions International

charlotte.watson@kbr.com

Abstract

The challenges associated with data management across the public sector can often be traced back to the procurement and development of disparate systems. Whilst there is always a 'high tech' solution, for some clients this can be out of budget, appetite, or both. KBR recently worked with a public sector client that wanted to bring their data to life from across a range of tools, hosted on different IT systems, different locations, and in different formats. Some initial work had shown that it was possible to analyse the data, but the manual trawl had taken six months. This was entirely unsustainable and could not be rolled out as a credible business process, but the jump to a fully automated solution was not a realistic next step.

This paper showcases our approach to developing an automation pathway and discusses how we overlaid the pathway with a set of key barriers and potential mitigations, in order to support a more manageable journey towards automation. To ensure our mitigations were at the cutting edge of developments, we assessed novel technologies such as natural language processing (NLP) techniques, including machine learning-based information extraction. Our approach aligns with standard risk-based techniques, but has been applied to data automation to help low tech clients, such as those hampered by system constraints, achieve better data management outcomes.

Keywords:

Automation Pathway, Data Management, Mitigation, Remote Process Automation

Introduction

Background and context

In 2019, the UK government published the National Data Strategy, which sets out 'how best to unlock the power of data for the UK'. The strategy centres on the potential for using data more effectively across decision-making and policy-setting to boost productivity and increase access to a fairer society for all. However, public sector customers often have challenges relating to data management that can be traced back to the procurement and development of disparate management systems, resulting in data sets that are hard to understand in isolation, never mind as an aggregated set.

HarKBRmonic was asked to help a customer show the benefits of its services, as delivered by over 40 local delivery partners, each using different sets of IT systems, software tooling and tracking different effectiveness metrics. The KBR team joined a study group formed of customer representatives, academics from two universities, and local delivery partner resources. Some initial work had shown that it was possible to measure benefits across the tools and systems, but the manual trawl of data had taken six months. While this was a necessary and essential first step, it was entirely unsustainable and could not be rolled out as a credible business process.

Our work focused on **establishing which parts of the process could be automated**, how this could be delivered, and what blockers might be seen or enablers required. KBR's analysts were responsible for leading the discovery and also took on a coordination role for the project.

Our approach

Our approach to the challenge was similar to that used in risk management, focusing on the identification and mitigation of blockers and barriers. We initially set out to **understand the manual processes that had been used to gather and review data**. Previous work had delivered a set of metrics and impact data required across a variety of systems, although identifying, collating and reviewing the data had taken six months. We assessed the metrics for suitability and reviewed the systems and data sets in preparation for automation.

Once the manual process had been reviewed, each part was **assessed for barriers that prevented automation**, in a similar way to a standard risk management approach. We **produced a set of mitigation activities** against these barriers to automation, using best practice for how our customer could overcome each. To ensure our mitigations were at the cutting edge of developments, we assessed novel technologies such as natural language processing (NLP) techniques, including machine learning-based information extraction.

Finally, to enable the assessment of delivery options, we **produced an automation pathway** that applied the mitigation strategies to give a step-by-step process for automating the data capture and analysis. This pathway was key to the outcomes of this phase of the study, as it formed a logical progression for taking data and extracting insights from it in an automatable methodology. We reported on our

overarching findings around wide-ranging topics, such as human behaviours, software suppliers, technology applicability and data management.

Our approach followed the structure below:



Figure 1: Analysis approach

Our findings

We found that, when **moving to a more automated way of gathering and analysing data from multiple disparate systems**, there are a number of potential blockers that are not specific to the public sector or to the systems we looked at for this project. We termed these ‘key barriers’ (KBs) and they apply to all data automation initiatives.

To develop mitigations for each of the key barriers, we established a process map detailing the steps required to move away from a manual data trawl across multiple systems – an **automation pathway**. We found that all the key barriers featured at least once across the full pathway. Our approach, which follows a similar process to risk-based strategies, allows mitigations to be proposed that are specific and relevant to each stage of the pathway.

The following process flow shows the automation pathway at the core, with key barriers and mitigations building outwards.

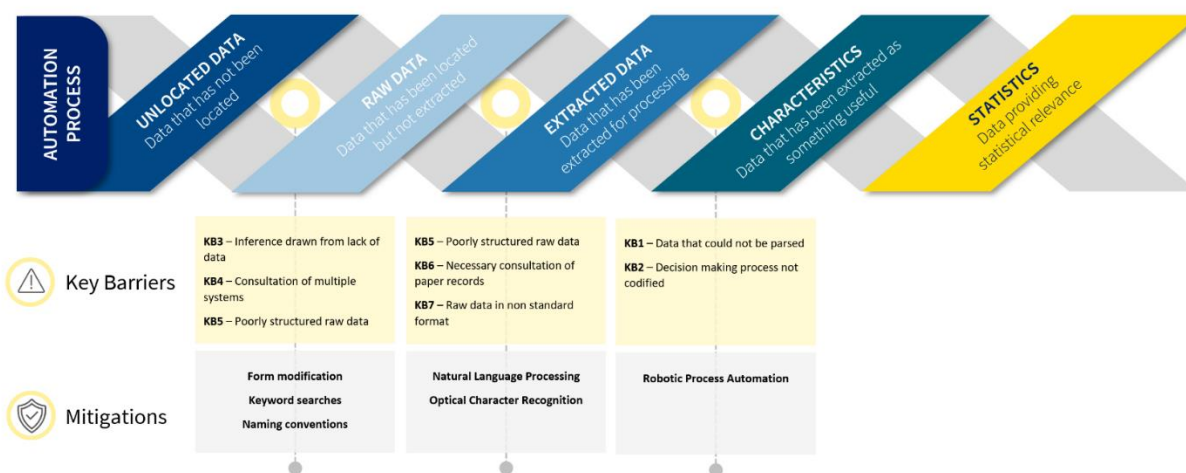


Figure 2: Automation pathway, barriers and mitigations

Developing a staged process for automation allows manageable steps to be taken, focusing on the end goal of better data management. If the proposed mitigations were to be applied across the automation pathway so that the end-to-end design was entirely automated, this pathway would be a form of **hyperautomation**, in which multiple automation tools that enable intelligent automation – including machine learning and robotic process automation – are used to scale automation initiatives.

Conclusions and recommendations

This piece of work covered one of the most critical challenges that we believe our customers will increasingly face in the coming months and years – the need to **automate data gathering and analysis processes from disparate data sets**.

We reviewed the data needed to show value and took an approach similar to that used within risk management to identify the major barriers to automating the process. We mapped these across a roadmap towards automation, which we called an automation pathway. We proposed a set of mitigations for each barrier in order to give a route that could be implemented on a stage-by-stage basis, removing the need for a holistic high-tech system that was unaffordable and unrealistic.

To support our findings, we also proposed a set of recommendations, structured by theme below.

Data and system management

- Use of proxies – we recommend that taking a holistic view and finding proxies that can decrease the reliance on multiple systems is advantageous.
- Investigation of systematic changes – several of the systematic changes covered during the establishment of the automation pathway (for example, better use of structured data, digital tools and forms) are simple, low-cost, local changes that could have benefits in terms of adopting a standard methodology. It should be noted that these local changes may be low cost and simple when applied to a single location but are likely to be costlier and more complex to standardise across multiple locations and systems.
- Use of keywords, including the development of a centrally-owned data dictionary that can be flowed down to local delivery teams, is critical.

Software suppliers

- Collaboration across suppliers – most large public sector software providers have collaborative working groups or other mechanisms by which user needs can be understood. Real progress can be made when such groups work together to allow future versions to better interact, and therefore for data to better flow across major systems.

Technology advances

- Robotic process automation (RPA) and other orchestration tools allow the use of elements of machine learning in order to remove the requirement for human interpretation regarding the assessment of value.

References:

¹ <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>