



London Borough of Hounslow

Final Report for the Cost Calculator (formerly Multi Agency Demand Management) Project, London Borough of Hounslow – Digital Transformation Project

March 2017

Summary

The Cost Calculator (originally Multi Agency Demand Management or MADM) is an open source, web-based intelligence tool capable of:

- processing data sets that contain financial information about service users from different council departments and, potentially, partner agencies;
- matching individuals between data sets; and
- producing reports about patterns of multi-service use (based on cost) linked to each service user.

The objective of the tool is to facilitate those working in council and other public services to be able to see the pattern of service use of an individual, or group of individuals, beyond one particular service. This will enable the council and our partners to improve our evidence-based commissioning, by using data to tell a more detailed story about service users in Hounslow.

The tool was developed as the output of a project aiming to explore how we could identify the top 20 most costly individuals with complex needs, and those most likely to become homeless, across different services in Hounslow.

2. Introduction

Hounslow Council was commissioned in spring 2016, by the Local Government Association and the Department for Communities and Local Government, to develop a multi-agency digital tool with the capacity to identify the top 20 most costly individuals with chaotic and complex lives in the borough (across a range of services and agencies), with the aim of contacting these individuals and finding out more about whether the services they are accessing meet their needs and whether information from these individuals could be used to help prevent homelessness.

Throughout the duration of the project, as a result of feedback from stakeholders, the focus shifted towards designing a multi-agency cost tool that might be able to be used to support the prevention of homelessness. To achieve this aim, the team added functionality to the tool to enable users to produce reports about the data of people living in temporary accommodation in the borough and, by analysing these reports, to establish whether any patterns of multi-service use could be obtained from this data that might enable us to predict homelessness.

The team encountered challenges along the way, including the initial project sponsor (the Head of the Intelligence Hub) leaving the council; a delayed project kick off; the Data Warehouse project, that was envisaged as being an integral element to the tool, being discontinued; technical issues; and information governance concerns. These are discussed in more detail in the relevant sections of this report.

These challenges had a significant impact on the approach that the project took, as discussed in the next section. However, the core objective of the project - the creation of a working and beneficial intelligence tool - remained constant and was successfully achieved. The project also generated a significant amount of new learning that has been welcomed by, and will be of benefit to, the council. Recommendations based upon that learning are included in this report.

3. Project Approach

The initial project approach was to carry out three phases of work. First, to investigate and identify the top 20 most costly individuals with chaotic and complex lives in the borough by collecting relevant data and building a tool to help analyse it. Second, to carry out research with those top 20 people to better understand issues around homelessness and adapt the tool to support a predictive model for the prevention of homelessness. Then a final phase of refining the tool and disseminating the project's outcomes.

This approach was adapted because of the challenges mentioned above - in

particular the decision not to proceed with the accompanying data warehouse project. Whilst the approach evolved throughout the project, the two most significant changes are as follows.

Firstly, the software development aspect of the project - the actual building of the intelligence tool - became much more prominent. The project team needed to deal with many practical considerations of gathering and processing data, which was originally going to be the role of the Data Warehouse project.

Secondly, it became clear that the space within the project for investigating two issues (the top 20 most costly individuals with chaotic and complex lives and predictors of homelessness) was not available. So a decision was taken to focus more on the issue of homelessness prediction. The project was still, however, constrained by the initial cost-based approach to analysis. So this was built upon by using cost as an indicator of service use; and service use by those experiencing housing difficulties as an indicator of contributing factors towards homelessness. The development of the tool and the data gathered was therefore refocused to meet this aim.

4. Project Progress

In April 2016 the initial project team was established, the project initiation plan completed, the Terms of Reference for the multi-agency panel drafted, and the project presented to the Local Strategic Partnership.

Whilst some progress was made in the initial months of the project, significant delays were experienced due to issues with securing funding and the availability of team resource. The project coordinator, Lucy Watt, joined the team in mid-June. The original lead for the project, Jo Sage, left the council in July, and the Intelligence Hub within the council was restructured, and responsibility for the project was transferred to the Director Joint Commissioning, Martin Waddington.

In May and June, further progress on the project initiation phase was made and initial research undertaken. In July and August the project plan was developed, specifications for the technological solution written, potential approaches for the procurement of a technological solution researched and considered, potential providers of data identified and some initial sample data sets collected.

During August and September the process of procuring a contract developer to help with the building of the analysis tool was undertaken. The developer began work at the end of September.

Work began in August to define, and through a process of internal procurement, commission the setup of the technical infrastructure necessary for the

development and eventual operation of the analysis tool. This process turned out to be far more challenging than anticipated - in part because the requirements of an open source development project were outside the usual experience of the organisation - which created delays in the development process.

From September onwards, a process of contacting potential providers of data, working with them to define data schemas, and collecting and refining actual data sets took place.

In September, the specifications for the technological solution were reviewed, decisions about the scope and approach taken, and a development plan agreed.

Work on the development of the analysis tool began in October and has continued throughout the project. Originally we had intended to first create a prototype or mock-up of the tool before starting development proper, but because of delays this approach was revised and instead we began immediately on creating an initial operational version of the tool. This first version was ready at the beginning of January 2017. Since then the tool has continued to be developed, undergoing several iterations as we have reviewed and tested it, demonstrated it to stakeholders, and added functionality. A more detailed explanation of the development process can be found in the "Tool Design and Build" section below.

Following feedback from senior Hounslow officers, and because of the demands of the DCLG grant funding element of the programme, in October it was agreed to pivot the project to focus more on homeless prediction. Between October and December, various research and stakeholder activities took place to support the new direction.

Throughout the project, consideration of information governance and data security issues has been undertaken. Problems with the restrictive nature of the information governance context have had to be overcome and have caused delays to the project timetable. It was decided that lessons learned from this should be used to encourage changes to the way that the council operates and that this should form part of the project. Team members have been working with other council officers to examine how this might be done.

An on-going process of stakeholder engagement took place throughout the project. This included interviews with data providers, service managers and caseworkers; workshops and demonstrations; regular meetings with the lead director; and panel meetings.

Supporting documentation for the analysis tool and this final report were prepared in February and March. The open source code for the analysis tool was publicly released in March 2017.

5. Stakeholder Engagement and Research

Delivering this project required engaging with a wide variety of stakeholders at senior, middle and frontline level, both within and outside of Hounslow Council.

5.1 Project Board

The project board was established based on the membership of a multi-service, multi-agency group that was established in 2014 to explore the most challenging homelessness and housing cases in Hounslow. The project panel agreed the project terms of reference and met periodically during the project.

Although no members of the project board from external agencies were able to attend the panel meetings, beyond the initial meeting, we continued to solicit their involvement as the project proceeded.

The project board and consequent contacts were instrumental in shaping the project and addressing many of the challenges we faced.

See Appendix 1 - Cost Calculator Project Board Membership

5.2 Fact-finding missions

Making contact with other local authorities working on related projects was extremely useful. Here are details of what we learned from our engagement with these teams.

5.2.1 Hertfordshire Council's You Can Project

The Cost Calculator project is partly inspired by Hertfordshire Council's You Can project. We went to visit the team in August 2016 to learn more about their experience of project scope and length, multi-agency working, data modelling and collection and service redesign. We were very impressed with what they had achieved, in particular the service redesign aspects of their project, but we discovered that their approach to data was quite different to that at the centre of our project.

The Cost Calculator is essentially a big data project, looking to bring together wide a variety of data to analyse spend and service use of individuals, originally with the aim of identifying the most complex cases. However, in phase 1, the You Can project started by collecting specific multi-agency data on pre-identified individuals with complex needs (including housing and drugs) in order to make the financial case for phase 2 of the project: the creation of a new service to support such individuals. The on-going collection of this data in You Can is now

used to monitor the success of the project.

We recognised the difference of emphasis between the two projects and couldn't fail to realise that the Cost Calculator project had more ambitious objectives and less time to deliver them than the You Can project. Completing phase 1 of the You Can project, i.e. getting to the point where multi-agency cost data was collected for a minimum of c. 30 people, took Hertfordshire approximately two years. Whereas, due to a late start, the Cost Calculator had less than a year to gather data and develop a tool to analyse multi-agency data across the whole council area.

The You Can experience of collecting relevant data and developing a cost tariff, where costs were investigated and agreed across multiple services, was a time-consuming and, for the most, part manual process.

We also discovered that multi-agency collaboration was built into the You Can project at an earlier stage. You Can is funded by a pooled budget to which all the various agencies involved contributed. This aids collaboration and makes data sharing more likely.

5.2.2 Lambeth Council's Data Warehouse and Interactive Dashboard Project

The Lambeth project aimed to create a data warehouse and develop a system to match individuals across data from different internal systems, and then create the user interface to interrogate this data. We made contact with them to find out about their data models, matching capabilities and the extent of customisation employed.

We visited the Lambeth team in November 2016 and their experience helped shape our approach to overall tool development:

- Data sharing issues: Lambeth addressed data sharing issues by adding a statement of intent to all forms that residents fill in to submit their personal data, covering consent for its wider use. For example: "We will use your data to improve services across the council". This forms part of this document's information governance recommendations.
- Matching logic technology: Lambeth experimented with "off the shelf" matching solutions, but the matching rate produced was too low, largely because their data was very specific (in type and accuracy). As a result they decided to write their own custom logic. This information helped us to opt for developing our own custom matching logic.
- Matching logic, timescales: it took Lambeth about three months to develop

their matching logic. This information helped us with adjusting project timescales and managing expectations. We knew that developing a perfect matching logic would be lengthy and that we would have to limit development to the time available in the project.

- Choices of technology: Lambeth used the Microsoft Suite for their project. Although we could not change our choice of technology at that stage of the project (even if we had wanted to), it was still useful to know what technology other projects adopted.

5.3. User engagement

Throughout the project we engaged with users to gather their input in order to shape the product.

5.3.1 User profiles

We created two user profiles most relevant to the tool:

- Group 1: Service managers and senior managers
- Group 2: Frontline case workers

See Table below: User Profiles

At the beginning of the project, we identified members of staff working with people with complex needs and chaotic lives who fitted these two profiles and interviewed them about their work and how the tool could support and extend what they do.

We received positive feedback, gained good insight and were able to test some of our early assumptions.

For example, we confirmed that the majority of frontline workers would most likely already be aware, anecdotally, of who their top 20 most costly individuals were within their own services. However, there was recognition that they might find the tool useful to provide broader information about their clients from across different services, which could support internal decisions about addressing clients' needs.

Table: User Profiles

Group 1 - Service Managers and Senior Managers	Group 2 - Frontline caseworkers
Manage services that are used by or come into contact with people with complex needs	Work directly with people with complex needs and chaotic lives – at the front line.
Have responsibility for, or make decisions about the allocation of, resources on which people with complex needs place demands	Work in public facing roles providing some kind of crisis service
Take a more macro view of how services are used and delivered, and can influence service redesign	Work in public facing roles providing services likely to be accessed by people with complex needs
Who would benefit from knowing who the top 20 are and what they cost	Who may already know who their difficult cases are
Whose job may focus (in part) on service redesign – they may collaborate with other service managers in different services and agencies when redesigning services	Who would benefit from a top 20 list backed up by data

5.3.2 User interviews and feedback

During early interviews with users, key themes emerged around the potential impact of the tool and how it could support their work. The key themes were:

- Planning for the future – the tool might have the potential to support better understanding of future service needs and requirements and help shape future services. Having such a tool would mean using data to be proactive, rather than reactive and would provide data on real people as robust evidence for the much-needed funding of services.
- Prevention and early intervention – the tool could support identifying candidates for intervention and measuring the impact of preventative work.
- Enable improved joined-up working - within council departments and across other agencies. By collecting and having access to shared data, teams might become more aware of the big picture and break out of working silos which stymie effective work.
- Greater knowledge and insight – the tool would create a fuller picture of an individual and their story. Discovering who the top 20 most costly

individuals are, what traits they share and whether they have fallen through the gaps of any existing services, would be beneficial.

5.4 Desktop research and workshop

As the project became more focused on homelessness, we worked to deepen our knowledge of the issue and draw on other specialists within the council for their input.

We carried out extensive desk research into homelessness, its various definitions and causes. We reviewed a variety of source material, including recent government papers on tackling the issue (e.g. Parliamentary Select Committee Report into homelessness); case studies on homelessness prevention and detection; and voluntary sector reports from Crisis and Shelter.

Based on this research we delivered a successful workshop about the prevention of homelessness with eight housing/homelessness specialists at Hounslow Council. Our aim was to identify and obtain further data sets that might help predict those at risk of homelessness and to understand more about a potential predictive model.

Through user journey exercises, we discovered further data sets that could support the project with its renewed focus on homelessness prevention. We also learnt that housing and homeless specialists were keen to be able to find out information about households from the tool's reports, as well as individuals, because people often present as homeless in family groups, not just as individuals.

The workshop brought to life the complexity around the issue of homelessness. Many who become homeless are struggling with a range of issues and circumstances which reinforce each other and result in homelessness, such as poverty, debt, family breakdown, troubled childhoods, mental health issues and addiction.

To develop a tool that can encompass data relating to this broad swathe of experience would require the resolution of quite deep-seated data protection and information governance issues, for example permission to use and share a client's personal details. Nevertheless, we decided to prioritise and pursue a realistically attainable group of data sets.

5.5 User feedback

Once we developed the tool prototype and populated it with real data, we held a series of demo-based feedback sessions where we presented the working tool

and asked for feedback. Having discovered that group 1 users (service managers and senior managers) would get best use of the tool, we focused on getting feedback from this user type. Group 1 users told us they would benefit from using the tool by having access to management information on service and client spend that they could interrogate and analyse to help make decisions about services.

User views were considered and reflected upon to design on-going changes to the product and to provide recommendations for future development.

The key themes echoed findings we encountered in the earlier user interviews, highlighted new insights and reflected some of the challenges we faced through the project. See below.

5.5.1 Functionality and data

The project objectives hinged upon pulling together a wide variety of data from different sources for analysis. It was not an easy task to find a sufficient variety of source data: in existence, in a useable format, from both from Hounslow Council internally and from external agencies. Internal systems, varying levels of engagement, data protection, information governance, and challenging timescales created barriers to the smooth collection of data.

As a result we decided to build the tool with limited, but robust, data sets. We felt this was the best way to create a proof of concept and show what the tool could do.

Most users agreed that the tool has great potential. Bringing together different data sets and being able to track spend by service, individual and group over time opened up exciting possibilities for people working in various council departments.

Even with limited data, the feedback received from users was that the results were interesting and merited further analysis and investigation, but to be really useful, the tool needs more robust data sets: well-structured data which reflects more aspects of an individual's life e.g. data from health services, DWP, police, probation services, community safety and voluntary services. With further data and some development the tool could represent an exciting beginning for Hounslow's real use of big data.

5.5.2 Homelessness prediction

The tool is at the early stages of being able to support homelessness prediction. The solution that was settled upon to enable the tool to produce some insights into the likelihood of an individual becoming homeless was to compare patterns

of service use of those in temporary accommodation (i.e. a homelessness indicator that can be measured through council spend) with those not in temporary accommodation. Populating the tool with far more data (for example data from health services, DWP, police, ambulance, probation services, community safety and voluntary services) is required to test this hypothesis. Do those in temporary accommodation have specific patterns of service use compared to the general population? And if they do, can this pattern be used to identify people who have an increased likelihood of becoming homeless? This analysis could be used to support identifying an "at risk" group to work with to help prevent homelessness.

5.5.3 Additional uses

Through demonstrating the tool to council officers representing different internal teams we realised that the product has the potential to support work in other parts of the council. For example:

- Troubles Families – to identify the next cohort of families who could benefit from joining the Troubled Families programme.
- Preventative Services – to measure the financial impact of preventative services on specific client groups.
- Commissioning – to scope and commission new contracts (services) and make best use of existing contracts where similar services are provided by different suppliers, with different costs.

6. Tool Design and Build

6.1 Development process

It was clear from the beginning that a part of this project was to develop a technological solution that would help to process and analyse data. As originally conceived this project would have run alongside the development of a data warehouse. The technology used to provide the data warehouse would therefore have formed a significant part of this solution. The specific technology developed by this project would then have been something that complemented or provided a layer on top of the data warehouse. This would likely have been something that supported data modelling and produced the interactive dashboard element.

So inevitably, when the decision was made not to proceed with the data warehouse project, this had significant consequences for the nature of the technological solution we would need to provide. Simply put, we needed to find a

mechanism to replace those elements of the technological solution that the data warehouse would have provided that were no longer available.

We did spend some time researching whether there were pre-existing technological solutions, in whole or in part, that would do this, but we couldn't find something that adequately met our needs – in particular the additional requirement in the project brief for the technology to be open source. So it became clear that, without the data warehouse, we would need to create our own tool or set of tools. Therefore, a significant component of the project became a development process in which a custom technological solution was designed and built. This has resulted in both a working analysis tool and the release of the open source code for that tool.

Originally this development process was to include an initial specification stage, a prototyping stage, and a stage of iterative development to build the solution. To enable this, it was decided to augment the project team with a developer familiar with the common open source technologies (essentially PHP and SQL databases). In practice, because of the impact of delays on the timetable of the project, the prototyping stage was effectively merged into the creation of an initial operational version of the analysis tool.

As would be expected, the design of the analysis tool has changed and evolved throughout the project.

Our first step in the development of the tool was to establish the requirements. To develop an initial understanding of the essential features that our proposed technological solution would need to have. To do this we began with what the project brief was asking us to deliver and then thought about how it might be achieved. The results were written up into the 'Technology Solution Requirements' document.

This requirements document also sets out a proposed approach to the technology solution.

Firstly, we emphasised the role of pseudonymisation as a way of overcoming the significant challenge created by the inevitable restrictions around the use of personal information.

Secondly, we identified three distinct user groups around which the solution would be structured. These were Data Providers (those who supplied us with data about services), Service Managers (those who would want to analyse the data and receive reports), and Caseworkers (those who potentially would be working with the individuals and groups identified by the analysis).

Finally, we proposed to create a suite of two tools; pseudonymisation tool or

service that would process the data supplied to us, and the intelligence tool, which would manage the data and run the necessary analysis.

The following diagram demonstrates the various elements of this proposed approach and how it was intended that they should fit together.

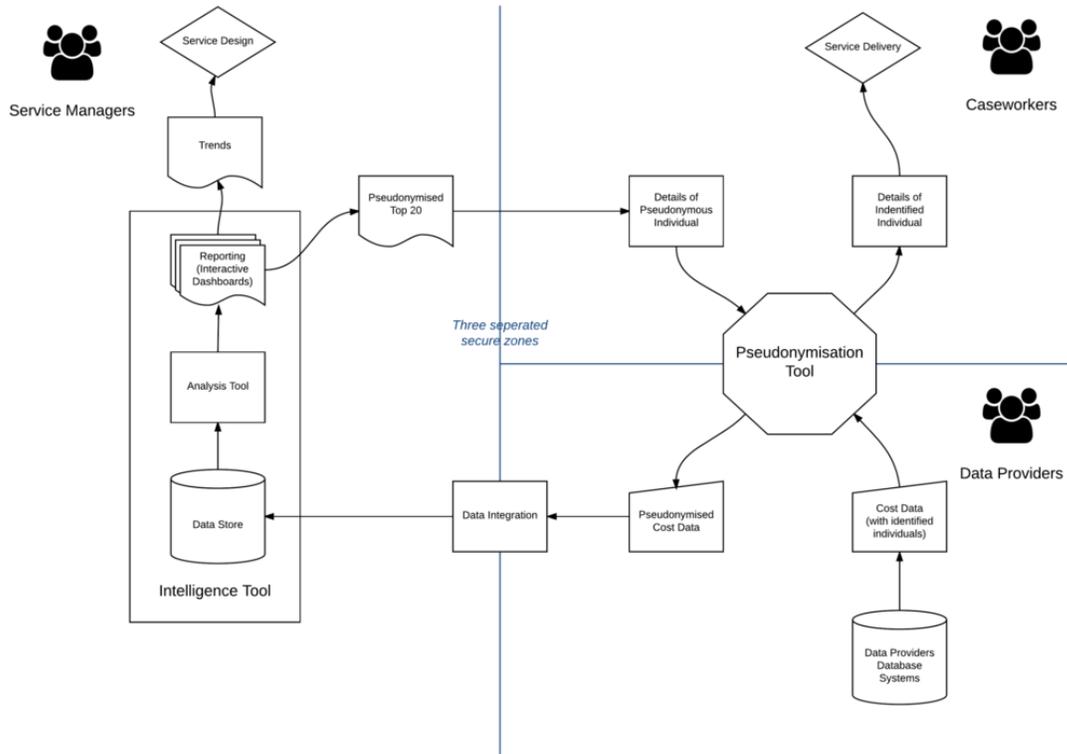


Figure 1 – Design of proposed approach

Taking the approach outlined in the requirements document, we next began to think about the nature and structure of the data we needed to work with. We set out the different assumptions we were making about the data and outlined the different calculations that would be required. Finally, we made a judgement about the different types of data sets that would be required. This work resulted in what we called our initial data model.

We then fed the conclusions of the initial data model into the work we had begun of defining schemas and gathering data sets. In particular, we used it to identify a series of questions we could ask potential data providers.

Expanding on the Requirements Document and underpinned by the Initial Data Model, the next step was to developing a Technical Specification. This was intended to outline in more detail the design of the technology solution and form the starting point for the actual development process.

Once the developer joined the project team, at the end of September 2016, this technical specification was reviewed in order to take on board his advice and experience. This resulted in some adjustments to the development approach.

Firstly, it was decided to remove the physical separation between the pseudonymisation tool and the intelligence tool components. Instead we would attempt to build a single combined analysis tool. This was a simplification that made the development of the tool easier. The main reason for the clear separation was to support the pseudonymisation of multi-agency data sets prior to them being shared with the council and so encouraging their inclusion. But by this stage the focus of the project had shifted to concentrating on internal data sets, so ultimately this wasn't a disadvantage.

Secondly, rather than extending the initial data model and designing a set of generic schemas and process based upon it, we decided to work with the actual data sets we had received and design the structures and processes of the tool around these. This approach had the disadvantage of making the initial versions of the tool inflexible, something that we have had to progressively address in the later stages of the project. However, it had the advantage of ensuring that the structure of the tool matched the data available in the real world. For instance, as the tool has been developed we discovered that some of the types of data we had assumed would be needed in our initial data model turned out not to be useful as no data sets in that format were offered to us.

Following this approach an initial version of the tool was developed. This included the basic framework for a web application, the necessary features for user authentication and management, the essential components of the database and a data import process, and a user interface that implemented the four functions of uploading data, displaying reports, client lookup, and configuring settings. This version of the tool included two data schemas providing data about adult social care and temporary accommodation.

This initial version of the tool was both a successful "proof of concept" and a solid framework on which we could build.

The user interface of this version of the tool was then reviewed by the team and a number of adjustments suggested. The tool was also extensively tested to ensure the integrity of the data. From this testing phase a list of issues and improvements was compiled. This list was then worked through and the results again tested. This gave us an operational version of the tool.

We then turned attention to ensuring that the tool would work in the proposed hosting environment provided for its on-going operation, and gathering wider feedback from stakeholders. This led to some further adjustments and added functionality, followed by a further round of testing.

This completed the tool was set up ready for on-going operation by Hounslow and the source code was released on GitHub.

6.2 An introduction to the tool

The Cost Calculator is a web-based intelligence tool designed to support local authorities in analysing the use of the services they provide. It uses the cost of service provision as the primary measure of service use.

The tool is capable of;

- processing data sets that contain financial information about the users of services,
- matching individuals across the different data sets to produce a "client index", which is used to calculate all costs associated with those clients,
- and producing three types of reports about service use, which can be filtered in different ways to aid analysis.

The three types of included report are:

- Total Spend by Client
- Breakdown of Spend by Client
- Total Spend by Service

The tool operates as a web application and includes user management and authentication features, the ability to assign different user permissions, and file upload capabilities.

The objective of the tool is to facilitate those working in local authorities and other public services to be able to analyse financial information and to gain an understanding of the pattern of service use of an individual, or group of individuals, beyond one particular service.

The source code and full technical documentation for the tool is available on GitHub. Go to: <https://github.com/LBHounslow/cost-calc>

The tool was primarily built in PHP using the Laravel framework and works with SQL Server and MySQL databases.

The tool is made available as open source software under the MIT license.

Full details of the various elements of the tool and how it works can be found in the "Guide to using the Cost Calculator".

See Appendix 2 – Guide to using the Cost Calculator

6.3 Issues

6.3.1 Matching algorithms, data quality and standards

We initially assumed that the matching process would require a sophisticated algorithm and the use of fuzzy matching techniques. In the end we took a much simpler approach, directly matching against three personal identifiers; last name, postcode, and date of birth.

Using this approach we discovered that, provided the data had those three identifiers, the quality of the matching was surprisingly good and therefore we believe that the quality of the data in the tool is good. We have added some logic to the tool to deal with duplicate data, and there is more that could be done to improve the matching process, but our conclusion is that you can achieve good results with a relatively simple process.

However, this approach has resulted in limiting the amount of data we have been able to source for the tool, because we needed to exclude data sets that did not include all three of the matching personal identifiers. This has meant that the data we have included has been of a good quality, but that we have also had to exclude information that would have been valuable in our analysis.

In future it would be possible to increase the sophistication of the matching process and this is worth exploring. Although ensuring that the data in the local authority includes an established number of personal identifiers that can be used for matching could potentially be more useful. This is an approach that requires establishing common standards for data rather complicated technical development - a more achievable goal within a local authority context.

We recommend that the development of data standards, particularly around common personal identifiers should form an integral part of the local authority's future information strategy. The local authority should also consider extending this approach into the development of a number of authoritative registers.

6.3.2 Commitment to open source

One of the significant challenges the project had to face was the consequences of developing a tool based upon open source technologies within an organisational context unfamiliar with such an approach.

We discovered that the local authority's IT team, upon whom we needed to rely for technical support, were unfamiliar with the requirements of open source software. Understandably, they were also reluctant to commit to the on-going maintenance of a tool based on open source technologies that they did not have the in-house capacity to support.

We dealt with this challenge in two ways. For the development process the project commissioned an external developer to help build the tool. To overcome issues of hosting and on-going maintenance we did additional work to make sure that the tool would be able to operate within a Windows OS environment that the IT team were more comfortable supporting.

However, this lack of capacity to support open source technologies remains an issue that any future similar projects will face. So we recommend that the local authority needs to consider whether open source approaches are something the council wants to pursue in future. If this is something that they do choose to do then some degree of in-house capacity in this area would be required. This need not be extensive, but the on-going ability to support open source applications would be essential.

6.3.3 Information governance

One of the major challenges that the project faced was dealing with the many information governance issues that constrained it. Obviously these had an impact on the development process. The lesson here is a simple one. While good design and technical features can be used to accommodate and mitigate information governance issues, technology on its own cannot solve problems. The environment in which the technology operates must be suitable.

Our recommendation, as is dealt with elsewhere in this report, is to propose changes to the local authority's information governance policies and procedures. These are designed to create a permissive environment that encourages innovation, but works in accordance with ethical and statutory obligations around data protection and privacy.

7. Data

Developing a multi-service multi-agency tool required gathering data from a variety of sources. From the start, the project board expressed concerns about the time it might take to gather data from every relevant agency, and encouraged us to start developing the tool with data that is easy to obtain and then add harder to obtain data as it became available. We followed this course.

As the project progressed, further data-sharing and information governance issues arose which resulted in us having to limit the data-gathering exclusively to data sets that were available and useable within Hounslow Council itself.

Although this was frustrating, with agreement from the project board, we resolved to proceed with this more limited range of data sets in order to create and test proof of concept.

7.1 Data provider engagement

We identified and made contact with a broad range of data providers both within and outside Hounslow Council. Following this initial exploration, we targeted key departments (Housing and Adult Social Care) and sourced sample data from them to inform the development of the tool.

Although getting data from external agencies proved unachievable during the lifecycle of this project, we nonetheless made good contacts and gained some useful insight into what external data we could use in future iterations, once data-sharing issues are resolved.

We had frequent meetings with internal data providers and familiarised them with the project then worked with them to understand which data sets would be available and useful.

As the project progressed, we were very grateful to receive a lot of goodwill and support with regard to satisfying our data requirements. However, in some instances some teams we approached chose not to engage, because of competing demands on their time.

7.2 Data gathering

Through liaising with data providers we identified and explored the possibility of including a number of collections of data sets (or schemas) in the tool.

The process we followed was as follows:

NOT PROTECTIVELY MARKED

- Initial discussions with the data providers to identify potential data sets
- The assessment of sample data from the potential data sets sent to us by the data providers
- The development of a draft schema definition to specify the format of the data (a process of matching the data available with the requirements and capabilities of the tool)
- A request to the data provider for data sets that match the draft schema
- An assessment of the data provided and if necessary a revision of the schema definition
- A request for the final (up-to-date) data for use in the tool

This was a process of progressively refining the data available to make it suitable for use in the tool. This involved attempting to solve problems and in some cases modifying the design of the tool as we went along. Obviously, not all the potential schema made it to the end of the process.

The potential schemas we identified are listed below.

Data internal to Hounslow

- Temporary accommodation – PSL
- Temporary accommodation – PLA
- Temporary accommodation – B&B
- Temporary accommodation – Other
- Adult social care services provided
- Adult social care contributions
- Housing benefit entitlement
- Housing benefit switch (from payment of benefit to individual, to payment directly to landlord)
- Council tax
- Troubled Families programme

- Council tax arrears
- Council tax single person's discount
- Rent arrears
- Discretionary housing payments

Data from partner organisations

- Mental health inpatients (West London Mental Health Trust)
- Community mental health patients (West London Mental Health Trust)
- Ambulance Service Frequent Callers
- Probation Service

7.3 Data gathering challenges

We encountered a number of challenges that we had to resolve:

7.3.1 Access to data

The lack of a council-wide data warehouse meant this project was very much starting from scratch in terms of bringing different data sources together. This meant that we had to identify and approach teams individually. In doing this we discovered a variety of approaches and circumstances in place when it came to recording data.

Some teams were able to extract data from their internal automated systems for us to use, but others provided us with data from manually maintained spread sheets, designed to support a specific task. This meant that in some cases data was incomplete and not well structured, at least for our purposes.

At the other end of the spectrum, some data we required is managed by contractors and there came a point in the project when Hounslow Council had to pay to extract and receive its own data.

In general, it took a while for us to receive data. This was in large part due, understandably, to team priorities and workload, but may also be due to the difficulty of extracting data from existing systems in the required format.

7.3.2 Data quality

Data quality was generally good, but the manual nature of much of the council's data collection that we encountered means there is a wide margin for human error.

7.3.3 Data incomplete

Manual data collection, designed for a specific task, meant that providing us with (seemingly straightforward) additional data could be time-consuming, resource-intensive and therefore not possible. For example discretionary housing payments data did not include date of birth and postcode. Some teams were only able to provide us with "snapshot" rather than "historical" data (e.g. Rent arrears).

7.3.4 Cost data

The tool was designed to analyse service use using information about the cost of service provision. This was based on the assumption that data about the cost of services would be readily available. Whilst we did find a number of data sets that included cost data, we also found others that contained useful information that would have supported analysis, but that did not record costs. This limited the amount of data that we had available for use in the tool.

We did overcome this problem to some extent by changing the design of the tool to include an additional schema type. This allowed us to include data sets that would identify the use of a service by an individual and so flag that individual as having a particular "status". We then made use of the data provided by this schema type in the filter functionality of the tool.

7.3.5 Unique identifiers

In order for the tool to assign the cost of service use, derived from different data sets, to the same individual, the tool needs to perform a matching process against the records of individual clients. This means that data imported into the tool needed to contain a range of unique personal identifiers.

Whilst most of the data we looked at contained personal identifiers, we discovered that these were not consistent (or consistently used) across different data sources. As discussed above, we chose to match against three personal identifiers; last name, postcode, and date of birth. These were the most common personal identifiers, but there were not included in all data sets we looked at. Again, this limited the amount of data that we had available for use in the tool.

7.3.6 Team resource

Our team included a data analyst from the start of the project until Dec 2016, when he left Hounslow Council to take up a new job.

7.4 Data selection

Due to the issues outlined above we had to carefully assess and select data to populate the current iteration of the tool.

7.4.1 Data used in tool

- Temporary accommodation
- Adult social care services provided
- Housing benefit entitlement
- Housing benefit switch
- Troubled families programme

7.4.2 Data obtained, but not used in tool

- Discretionary housing payments
- Rent arrears
- Adult social care contributions
- Council tax
- Council tax arrears
- Council tax single person's discount

7.4.3 Schema definitions

The final schema definitions for the data sets included in the tool are as follows:

- ASC01: Adult social care services

- H01A: Temporary accommodation
- RB01: Housing benefit entitlement
- RB03: Housing benefit switch
- TF01: Troubled families
- GS00: General Services (a generic schema)

8. Information Governance

Throughout the process of developing the Cost Calculator tool, concerns about information governance were raised by the Information Governance Manager and needed to be resolved.

We learned that for such data projects to take place, the council needs to change how it obtains consent from customers about data usage in order to comply with information governance regulations and best practice. The scope of this project was therefore extended to work with the Information Governance team to change the internal data collection processes and to work with our partners to support them to do the same.

The Cost Calculator team has created a report containing recommendations for the changes that are needed to be made to Hounslow's Information Governance Policy to enable the council to use individuals' data with proper consent in all future big data projects.

It was decided that the research exemption of the Data Protection Act applied to the development of the Cost Calculator, and the project plan was amended to remove the element where the data would be used to identify and contact a "top 20" most costly service users with the aim of obtaining information that would support future service re-design or improvement, as permission had not been sought from users for their data to be used in such a way.

See Appendix 3 - A data exchange framework for the London Borough of Hounslow

9. Conclusions

This section highlights the key outcomes, benefits, lessons learned and recommendations coming out of this project.

9.1 Outcomes

The tool has delivered the following outcomes:

- 1 A live and working web-based intelligence tool capable of processing data sets that conform to a number of defined schema, matching individuals and extracting cost data, and producing reports about patterns of service use based on cost for individuals.
- 2 The open source code for the above tool, with accompanying documentation, publicly released so that others can use and extend the tool.
- 3 Details of the algorithms/data logic for matching individuals and for calculating the costs, publicly released so that others can use and extend them.
- 4 The live tool populated with a range of real data sourced from within Hounslow and producing reports about patterns of service use by individuals in temporary accommodation.
- 5 Details of the schemas of the real data used in the above, publicly released so others can replicate the data sets within their organisation.
- 6 A set of recommendations about how Hounslow Council should change their information governance arrangements to allow for a more intelligent use of their data.
- 7 This project report with details of lessons learnt and additional recommendations about suggested next steps.
- 8 A plan for the dissemination and communication of the above products.
- 9 Functionality in the intelligence tool for adding new schemas.
- 10 Recommendations and a road map for further developing the intelligence tool.
- 11 The live tool producing reports comparing the patterns of service use by individuals in temporary accommodation with the pattern of service use by individuals in the 'general' population.

- 12 Recommendations about how to use the intelligence tool to predict homelessness.

9.2 Benefits of tool

The tool has enabled and established Hounslow Council's first real experience of using big data. It brings together a variety of data sources and enables its analysis to aid a whole raft of evidence-based decision-making. Users can now bring together data from a variety of sources, match it and analyse it to gain real insight to council spend by service, individual and by group. In its small way, the tool has pioneered many of the key elements required to make any future big data project work. The project has highlighted important issues and launched key changes required to make Hounslow digital transformation aspirations possible. The tool represents a successful proof of concept.

The tool has been identified by a variety of internal teams, beyond the Housing team, as being able to support and extend current and new initiatives. The teams so far include: Troubled Families, Commissioning and Preventative Services. We are in the very early stages of understanding the tool's impact, but most agree it has great potential. Realising the benefits hinges on embedding the tool into Hounslow Council and using and refining it as much as is possible in the coming months.

9.3 Lessons learned and recommendations

This section outlines lessons learned and makes recommendations for the future of the tool and future projects of a similar nature.

9.3.1 Technical infrastructure

The corporate decision not to proceed with the Data Warehouse project had a significant impact on this project. Many of the issues we encountered that would otherwise have been addressed by the team developing the Data Warehouse, had to be taken on by the team delivering the Cost Calculator. This meant that more of the team's resource was spent on the basics of data gathering and processing and that the technology solution required us to develop a "from scratch" tool rather than something to complement the data warehouse technology. As a result there was less space in the project for pure analysis, data modelling, and refining the reports.

Recommendation 1: If the aim of future projects is to produce sophisticated data analysis such as the prediction of homelessness, the basic systems and tools to support this need to be in place.

The lack of capacity within the London Borough of Hounslow to support open source technologies remains an issue that any future projects with a requirement to use open source software will face.

Recommendation 2: The council needs to determine the level of commitment to open source technologies that it wants to make; and then develop the appropriate level of internal capacity to support that commitment.

9.3.2 Information Governance

If information governance issues are not resolved prior to a project starting, it can cause significant delays to the project. We were also not able to proceed with the multi-agency working side of the project, including data sharing agreements, because we had not resolved our own information governance issues.

Recommendation 3: Achieve information governance approval from all agencies involved before designing and commencing a big data project.

The council's information governance procedures, particularly the permission sought when collecting data from individuals, do not readily support big data projects, where the sharing, analysing and use of data to support policy development and service delivery improvements are integral to the success of the work.

Recommendation 4: We have produced a series of recommendations which if adopted will support the future development of the tool and future big data projects within the council.

See Appendix 3 - A data exchange framework for the London Borough of Hounslow

9.3.3 Data

The experience of this project is that for some council staff the primary mechanism they use to store, manage, and exchange data relied upon using spread sheets. This suggests the possibility that people may not be using existing information systems to their full advantage or that they lack the skills to make use of them. It also limits the potential development of more sophisticated data integration and reporting activities.

Recommendation 5: That the council reviews the internal use of low level data storage systems, the take up of existing information systems, and relevant staff training and support needs to identify ways to get more benefit from the data it holds.

Contracts for one of the council's third party contractors did not include the charge-free supply of council data about service users. More budget was needed to pay for some of the data sets to be supplied. We also faced difficulties in securely transferring the data between the council and a contractor, needing to resort to a manual transfer using external hard drives.

Recommendation 6: Third party contracts should include free provision of data to the council. There should be systems established to enable secure data transfers between contractors and the council.

We were not able to use some relevant data sets we encountered because basic personal identifiers were not present.

Recommendation 7: The development of data standards, particularly around common personal identifiers should form an integral part of the local authority's future information strategy. The local authority should also consider extending this approach into the development of a number of authoritative registers.

9.3.4 Project buy-in and support

We did not manage to cultivate the level of engagement from other agencies throughout the project. This was partly because of the delay in us resolving our information governance issues, but also because representatives of other agencies did not attend our board meetings, beyond our initial meeting.

Recommendation 8: High-level commitments to such a multi-agency project needs to be sought at a very early stage of such a project (e.g. tender stage) for it to succeed. A financial commitment from partner agencies as part of the project would ensure closer working.

Now that we have a working tool, we can demonstrate it widely to partner agencies. For example, colleagues in the health service (e.g. West London Mental Health Trust). The aim would be to generate buy-in and support, and pave the way to joint working and data-sharing across the agencies. Hounslow could take a lead in this endeavour and champion the development of big data projects within the local area.

Recommendation 9: Showcase the tool to external partner agencies.

The changing corporate priorities of the council meant that it was difficult to ask for some teams to make time and prioritise the provision of data to the project team.

Recommendation 10: Future projects will be better supported if they are tightly aligned with corporate priorities that are recognised across the

organisation.

In demonstrating the tool as widely as we could within the council, we found many teams were interested in using it to support and extend their existing work. When Hounslow council officers are engaged with the tool, they can envisage other uses for it.

Recommendation 11: Share the tool with more teams across the organisation, so that they can develop their own uses for it. Demonstrate the tool to more senior managers in Hounslow, highlighting the flexibility of the tool and the benefits of big data analysis in their work. This would also create an opportunity to request further data sets to add to the tool.

Many of the team who worked on developing the tool are contractors, who are moving onto new projects outside the council.

Recommendation 12: That permanent staff members own the project from now on.

The tool has brought tangible benefits to Hounslow Council and a lot has been learned. It is important to share this knowledge across the LGA network and for other council's to use and learn from what has been achieved.

Recommendation 13: Work with the LGA and Hounslow Communications Team to develop and launch a communications strategy.

The project faced difficulties when the initial project sponsor left the council.

Recommendation 14: Build contingency about project leadership into future project plans.

The funding to support the on-going development and promotion of the tool will not be available after year 1.

Recommendation 15: Seek additional funding to support further development of the tool, and suggest to LGA that funding is made available for similar projects for a period longer than a year.

9.3.5 Future tool development

The tool has reached an exciting phase in its development, but really we are just at the beginning of what is possible. Further development could yield great benefits for the council and its partners.

Recommendation 16: Seek further funding, in collaboration with partner agencies, to move to a second phase of development. The LGA, Nesta

and other agencies may have digital transformation funds which we can pursue.

The tool would benefit from further development. We have identified further functionality which would improve the tool and make it even more useful to the council. For example, moving towards identifying households as well as just individuals would greatly support the Troubled Families programme. Adding new features to the advanced search capability and improving the sophistication of the matching logic would provide greater powers of analysis.

See Appendix 4 - Development Roadmap

Recommendation 17: Commit resource to implement the development roadmap.

The shift of priorities for the project towards homelessness prediction was ambitious. We created a temporary accommodation filter in the tool to enable analysis to be carried out about people who live in temporary accommodation and are one segment of the homeless community.

Recommendation 18: Use a big data analyst to support the development of predictive elements of the Cost Calculator tool.

The financial and other benefits of the tool are not yet known, and the LGA and other funders are keen to hear about the potential impact and benefit of the tool for the council.

Recommendation 19: Implement a means of measuring the impact and recording the benefits of the Cost Calculator tool and report them to the LGA and other funders.

Appendices

Appendix 1 - Cost calculator project board membership

Appendix 2 – Guide to using the Cost Calculator

Appendix 3 - A data exchange framework for the London Borough of Hounslow

Appendix 4 - Development Roadmap

Appendix 1 - Cost Calculator Project Board Membership

Attended

- Peter Matthew - Director Housing – Chair
- Martin Waddington - Director Joint Commissioning (representing CCG) and Project Sponsor
- Lourdes DeBarry - Head of Housing Strategy, Investment & Performance
- Gerry McCarthy – Head of Enforcement, Community Safety and Commercial Regulation
- Vinesh Govind – Acting Head of Intelligence Hub
- Ben Tomlinson – Performance and Data Manager, Housing Strategy, Investment & Performance
- Alex Tambourides, CEO, Hammersmith and Fulham Mind
- Martin Bowdler, Stakeholder Engagement Manager, London Ambulance Service
- Lucy Watt - Project Co-ordinator (London Borough of Hounslow / Freelance)
- Andy Strange – Technical Consultant (London Borough of Hounslow / Freelance)
- Ingrid Karikari - Project Consultant (London Borough of Hounslow / Freelance)
- Jack Segal, Software Developer (London Borough of Hounslow / Freelance)
- Owen Kennedy – Intelligence Analyst, London Borough of Hounslow
- Fatima Ajia – Information Services Officer, London Borough of Hounslow

Invited to meetings

- Metropolitan Police representative
- West London Mental Health Trust representative

NOT PROTECTIVELY MARKED

- Children's Safeguarding and Specialist Services representative
- Education and Early Intervention representative
- Adult Safeguarding, Social Care and Health representative
- Public Health and Leisure representative
- VCS representatives

Appendix 2 – Guide to using the Cost Calculator

Overview

The Cost Calculator is a web-based intelligence tool designed to support local authorities in analysing the use of the services they provide. It uses the cost of service provision as the primary measure of service use.

The tool is capable of:

- processing data sets that contain financial information about the users of services;
- matching individuals across the different data sets to produce a "client index", which is used to calculate all costs associated with those clients;
- producing three types of reports about service use, which can be filtered in different ways to aid analysis.

The three types of included report are:

- Total Spend by Client
- Breakdown of Spend by Client
- Total Spend by Service

The tool operates as a web application and includes user management and authentication features, the ability to assign different user permissions, and file upload capabilities.

The objective of the tool is to facilitate those working in local authorities and other public services to be able to analyse financial information and to gain an understanding of the pattern of service use of an individual, or group of individuals, beyond one particular service.

The source code and full technical documentation for the tool is available on GitHub. Go to:

<https://github.com/LBHounslow/cost-calc>

The tool was primarily built in PHP using the Laravel framework and works with SQL Server and MySQL databases.

The tool is made available as open source software under the MIT license.

Installation

The tool is a web application so needs to be installed as a website accessible to web browsers via an internal network or the internet.

To run the tool it needs to be set up on a web server running PHP 5.6.4 or above with access to a SQL Server or MySQL database. The web server also needs to have Composer and Laravel installed.

Full installation instructions are available on GitHub here:

<https://github.com/LBHounslow/cost-calc/wiki/Installation>

How It Works

The tool provides the user with three main functions:

- The ability to upload cost data
- The ability to query the uploaded data to produce reports
- And a client look up function that allows the user to see details about an individual that can help to identify them

The tool also has a number of settings that a user can use to configure different aspects of the tools operation, and some logs that monitor activity as the tool is being used.

Upload Process and Client Matching

The tool allows the user to upload and import standard Microsoft Excel files containing data. The structure of each file needs to conform to one of the schemas included in the tool (see below).

Each uploaded file is validated against the related schema and processed according to rules set out in an import script associated with the schema. The data is then stored in import tables within the database.

Once the data has been imported further processing takes place to match each record against an index of individual clients.

This 'client index' is queried to see if the details of the client specified in each record of the imported data matches the details of a client we have already recorded. This matching process currently uses the following personal identifiers:

- Last Name
- Postcode
- Date of Birth

If we have a match against an existing record in the client index then we attach the cost data to it. If we don't have a match then we create a new record in the client index and attach the cost data to it.

Attaching the cost to an individual is done by assigning a unique ID number that represents the individual to the cost record. These ID numbers are used throughout the tool to represent individual clients. This means that we do not need to display identifiable personal information about clients in the tool, thus helping to preserve their privacy. The exception to this is obviously the Client Lookup function (see below).

Using this process, as more data is added, the user is able to build up a databases of cost information that is all indexed against specific individuals.

Cost Calculation

The cost data from the individual import tables is combined (using a SQL view) into a structure with common column headers. This enables the tool to run various database queries against the data which will result in cost calculations.

The common column headers are:

- Client_id
- Surname
- Dob
- Postcode
- Service
- Service_type
- Frequency
- Start_date
- End_date
- Unit_cost

NOT PROTECTIVELY MARKED

Making sure that the data fits into the above format is generally done by the import scripts during the upload process. The bits of logic that the queries carrying out the cost calculations do are:

- Working out the date range of service use (Report Date vs Service use Date)
- Working out the cost of the service use for above date range (Frequency * Unit_cost)
- Applying filters to include or exclude different service types

Reports and Filters

Using the cost calculations the tool presents the following three types of report to the user:

- **Total Spend by Client:** this lists the individuals in the client index in the order of the total amount of cost they have generated through service use.
- **Breakdown of Spend by Client:** this shows the details of the cost of service use incurred by a specific individual.
- **Total Spend by Service:** this shows the total amount of cost incurred by the different types of service.

The above reports can have the following filters applied to them:

- The date range covered by the report
- The type of services to be included in the output

In addition, an advanced filter can be used to change the pool of clients that the reports are looking at. These filters can be used to include or exclude clients based on whether they have ever received the specified service or group of services. Currently the tool can apply this kind of filter for the following type of services:

- Any temporary accommodation
- Membership of the Troubled Families programme
- Client having had the payment of their housing benefit switched from being paid to them, to being paid directly to their landlord

In the latter two cases the "services" are acting as a form of status assigned to the clients (see below for types of schema).

Client Lookup

In order to protect the privacy of the clients whose information is used by the tool we use a unique ID number to represent these individuals. This means we do not need to display personal information that could be used to identify an individual within the tool.

However, there may be circumstances that require an individual to be identified. The client lookup feature is used to convert a unique client ID into personal information (last name, date of birth, and postcode). Future development of the tool could enable further personal details to be displayed.

File Types and Data Schemas

The tool allows the user to upload and import standard Microsoft Excel files containing data. Each data file needs to be associated with one of the file types already set up in the tool. Each of these file types is associated with one of the included data schemas and its accompanying import script.

In order for the import process to work the structure of each data file needs to conform to the relevant data schemas. These schemas define what columns need to be included in the file and the format of the data in each column.

File Types

File Types have 3 properties:

- **Service Name:** The name of the service providing the data. This property will be combined with the "service_type" field, where included, to produce the service type label in the data.
- **Schema:** The schema that the data files for this file type needs to conform to.
- **Import Script:** Which script is used to handle the import of the file.

The tool initially contains the following four File Types:

- Temporary Accommodation
- Adult Social Care
- Housing Benefit Switch
- Troubled Families

Adding New File Types

New file types can be created from within settings section of the tool.

Data Schemas

The tool currently includes the following five data schemas:

- Temporary Accommodation
- Adult Social Care
- Troubled Families
- Housing Benefit Switch
- General Service

The first four schemas have been developed to handle specific sets of data available within the London Borough of Hounslow and match the four File Types included. The import scripts that process these schemas include custom code designed to meet the specific requirements of this data (e.g. there is some custom logic around getting the dates from Adult Social Care).

The General Service schema has been included to allow for additional data sources to be included in the tool without requiring further development. Provided that the files containing this data adhere to the schema they can be imported into the tool without any customisation.

Full definitions for these schemas can also be found on GitHub [<https://github.com/LBHounslow/cost-calc-schema>].

Types of Data Schemas

There are currently two types of data schemas:

- Service Cost (costs incurred for service use, against each client)
- Service Status (use of a service, but no costs associated)

The Service Cost schemas are the standard type of schema that tool has been designed to work with. These provide the service use and cost information that are analysed and reported on in the tool. The Temporary Accommodation, Adult Social Care, and General Service schemas are of this type.

The Service Status schemas allow the tool to work with data about the use of a service where there is no cost information. Including this information means that in effect the use of a service by an individual can flag that individual as having a particular "status". The Housing Benefit Switch and Troubled Families schemas are of this type. The tool uses these two schemas to provide two of the "status" criteria available in the advanced filter that can be used to change the pool of

clients that the reports are looking at.

Adding New Data Schemas

It is possible to add new data schemas to the tool. Unfortunately, adding new schemas currently does require some additional development, although the process itself is relatively straightforward. A developer can do this by defining the schema, writing an accompanying import script, and registering the new schema with the tool. The developer should be able to use the existing schemas as a model when they do this.

Further instructions about adding new schemas can be found in the technical documentation on GitHub here:

<https://github.com/LBHounslow/cost-calc/wiki/File-Types-and-Data-Schemas#adding-new-data-schemas>

User and Provider Administration

Users

Access to the tool is controlled by authentication against user accounts. Users need an email address and password in order to login to the tool. There is also a "forgotten password" process to allow users to reset their password.

User accounts can be managed from within the settings section of the tool.

Access to the different elements of the tool is also controlled by setting permissions for each user account. The following permissions are available:

- Upload Files
- Reports and Analysis
- Client Lookup
- Settings

User accounts can also be temporarily suspended.

Recommended User Permission Scheme

The tool was built in order to meet the needs of three distinct user groups. These were Data Providers (those who supplied the data about services), Service Managers (those who would want to analyse the data and receive reports), and Caseworkers (those who potentially would be working with the individuals and groups identified by the analysis). The different user permissions in the tool enable these different user groups to be given access only to those elements of the tool that they require.

It should be noted that, in order to conform with information governance best practice, access to the client lookup element of the tool, which provides personal information that can be used to identify an individual, should only be given to those who have a legitimate reason for contacting identified individuals.

The recommended user permission scheme is therefore as follows:

Data Providers

- Upload Files

Service Managers

- Upload Files
- Reports and Analysis

Caseworkers

- Client Lookup

Tool Administrators

- Upload Files
- Reports and Analysis
- Settings

Providers

The tool was built with the assumption that data would be provided by different departments within a local authority and by a range of partner organisations. In order to restrict the data files that these different data providers can upload to only those they have responsibility for the tool allows for the configuration of different providers.

Providers can be set up with a unique name and associated with some or all of the file types available in the tool.

Each user account has to be associated with one of the providers available in the tool. The file types that users with the "upload file" permission are allowed upload are then restricted to just those associated with that provider.

Providers can be managed from within the settings section of the tool.

The tool currently includes the following providers:

- Hounslow MADM
- Hounslow Housing
- Hounslow Adult Social Care
- West London Mental Health

Logs

In order to help with administration and monitor security the tool logs various activities for later review.

The Upload File section of the tool includes details of the current status of the data in tool and logs new uploads and other changes.

The Settings section of the tool includes two logs that help monitor the use of the tool. The User Login Log records both successful and unsuccessful attempts and accessing the tool. The User Change Log records changes to user accounts.

Technical Information

The tool was primarily built in PHP using the following Open Source Libraries:

- Laravel 5.3
- Bootstrap 3
- Bootstrap Tables
- Chart JS
- jQuery
- Select 2
- Date Picker

It works with SQL Server and MySQL databases.

Appendix 3 - A Data Exchange Framework for the London Borough of Hounslow

When a resident or member of the public provides personal information to the council, or generates information about themselves by receipt of a service or participation in other activities, that data is available to be exchanged within the council and with external organisations. The personal information provided by the individual has the potential to be viewed and used by parts of the council or other organisations not involved in its original collection.

If large amounts of information were to be made available in this way new "big data" techniques for matching, analysing and managing data could be applied to it. The use of such techniques opens up a range of opportunities for improved strategic decision making, more efficient and effective service delivery, and to develop greater knowledge about the council and the borough.

However, alongside these new opportunities are the legal, ethical and reputational risks associated with exchanging large amounts of personal information. The council has a responsibility to protect the privacy of the residents and ensure that they can exercise their rights over the use of their personal information. So the difficult task is to find an approach that balances that responsibility while still making it possible for the council to make use of "big data" techniques.

If in the future the London Borough of Hounslow wants to make use of these new opportunities an improved data exchange framework is needed to clarify what the legitimate reasons for making personal information available are and to ensure that the correct procedures are in place to enable this to happen in a way that minimise the exposure to risk.

This document seeks to clarify both the legitimate reasons for exchanging personal information within the council and with external organisations; and the reasons that should be considered unacceptable. It also outlines the components of the proposed data exchange framework; namely:

- An improved Privacy Notice and procedure for gaining appropriate consent
- New guidelines for the use of data in research
- New guidelines for the use of data in service redesign
- Revised Data Sharing Agreements with other organisations

- A communications strategy for informing Hounslow residents about the use of their data

Legitimate reasons for exchanging data

The London Borough of Hounslow should regard the following reasons for sharing personal information as legitimate:

A) Exchanging data to support the provision of a service.

This is where information gathered during the provision of a service to an individual is shared with others so that the service itself can be properly provided. The data exchange takes place to support the provision of only the service for which the information has been gathered.

This is the standard and most common reason for exchanging data.

B) Exchanging data to fulfil legal and statutory obligations.

This is where the council is required to share information in order to conform to obligations imposed on it. For example: reporting statistics to central government, the investigation of fraud, or ensuring the protection of children or vulnerable adults.

This is again a standard reason for exchanging data.

C) Exchanging data for the purposes of research.

This is where; in order to understand how services are being delivered, the impact of those services, and what is happening in the borough; information is shared to enable analysis to take place.

This is already taking place within the council to some extent. But it is clear that the data the council holds is not being used to its fullest potential.

In order to make greater use of the information the council gathers in research and analysis it appears that achieving greater clarity in the council's approach to the use of data in research would be useful.

D) Exchanging data to offer additional benefit to the resident.

This is where information is shared so that residents can be contacted and offered additional services or interventions that could be of benefit to them.

For Hounslow this would be a new reason for exchanging data. The council does not currently have activities where data from different services is compared so as to influence the services delivered to a specific individual. If this was to be made possible, providing the appropriate procedures are put in place, it would represent a significant opportunity to improve service provision.

But it is important that the data is only used in an appropriate manner. Its use must take place within ethical and legal parameters suitable for a public sector organisation. To do this it is proposed that the use of shared information would be restricted to "making an offer of benefit" only.

The information would be used to make an *offer* of services only, to which the resident has to agree. No changes to service provision or the relationship between the resident and the council should take place as a result of the exchange of data without the residents consent.

The offered services or interventions must be of additional *benefit* to the resident themselves. The reduction of cost, achieving greater efficiency, or the convenience of the council, *on their own*, are not sufficient to make the offer. Whether or not the proposed use of the data exchanged would lead to genuine additional benefit for the resident would be a test that any new proposal for using such data would need to meet. The council's information governance procedures would need to apply this test and proposals that did not lead to genuine additional benefit should be prevented.

Unacceptable reasons for exchanging data

In order to reassure people consenting to having their information exchanged, and residents of the borough generally, that their information will not be used inappropriately the London Borough of Hounslow should make it clear how information will **not** be used. It should be clearly stated that the council will not share information for the following reasons:

- As stated above, exchanging data for the reduction of cost, achieving greater efficiency, or the convenience of the council, *on their own*.
- Exchanging data for the purposes of marketing products and services;

either the council's own or those of a third party.

- Exchanging data for the purposes of commercial and financial gain either by the council's own or a third party (the information will not be sold on).

This is of course in addition to the inappropriate use of data under the Data Protection Act and other legislation.

Information and consent

There are a wide variety of situations where information can be shared by the council provided they do so in line with legislation. In some situations the council can only share personal information where it has received informed consent to do so from the individual who supplied that information.

In the case of reasons A and B consent is effectively implied from when the individual signs up to receive the particular service. Although the individual should be informed of how their information will, and will not, be used.

Due to the legal exemption for the sharing of information for research purposes, consent is not explicitly required for reason C, provided the council undertakes the research in a way that conforms to the published guidelines (for example by using anonymization techniques). Although, again, the individual should be informed of how their information will, and will not, be used.

Explicit consent *is* required for reason D. Without that consent shared information cannot be used to offer an individual an additional service or suggest an appropriate intervention outside of the service that the information was provided for. This is of course appropriate to preserve an individual's right to privacy and give them control of their information. But the lack of this consent across the board creates a barrier to the innovation of service delivery for the council.

So to enable services to be redesigned to make use of reason D the council needs to acquire explicit consent from residents. The suggestion is that in practice consent is acquired in such a way as to provide a sufficiently permissive environment that facilitates innovation but is restricted to only those activities that the majority of individuals would be comfortable signing up to. Hence the limitation to "making an offer of benefit".

LBH should ask people to consent to the sharing of their information for reason D when they sign up to a service, but allow them to refuse if they do not consent to this. For practical reasons this refusal would apply to the information generated by that specific service, not a refusal that applies to all services.

In summary, when individuals begin to provide information to the council as part of the receipt of a service or participation in an activity they should be informed that their information may be shared for reasons A (to support the provision of the service) and B (to fulfil obligations on the council) as usually happens currently. They should also be told that it may be shared for reason C (for the purposes of research) but also that this only happens in a way that conforms to a set of guidelines. They should also be told that it may be shared for reason D (to offer additional benefits) but they can choose to opt out of this if they do not want that to happen. Finally, they should be told what their information will *not* be used for and what rights they have over their personal information.

Sharing with other organisations

The above reasons for sharing personal information should also be seen as legitimate for exchanging data with other organisations. In their different ways each of the four can equally be used to justify exchanging data with an external organisation as with another part of the council.

So it is suggested that when getting informed consent from an individual for data exchange consent is also acquired for sharing with other organisations. However, it is recognised that the barriers to achieving this are considerable.

One step towards achieving this could be to bring the Information Sharing Frameworks that the council has with partner organisations into line with this new data exchange framework.

Existing agreements could be revised and new agreements made to ensure that they take account of the above reasons for exchanging data. Note that it may not be possible to share information for reason D (to offer additional benefits) under current agreements.

Implementation

Implementing this the proposed data exchange framework requires establishing the following components:

- An improved Privacy Notice
- New guidelines for the use of data in research
- New guidelines for the use of data in service redesign
- Revised Data Sharing Agreements with other organisations

- A communications strategy for informing Hounslow residents about the use of their data

An improved Privacy Notice

This is the key statement provided to clients when they sign up to a new service and provide personal information. It is through this that we achieve the necessary informed consent.

The current Privacy Notice should be reviewed to bring it into line with the reasons for the exchange of information outlined above. The reviewed Privacy Notice should include the following:

- A statement that their information may be shared to support the provision of the service.
- A statement that their information may be shared to fulfil obligations on the council
- A statement that their information may be shared for the purposes of research but also that this only happens in a way that conforms to a set of guidelines.
- A statement that their information may be shared to offer additional benefits but that they can choose to opt out of this if they do not want that to happen.
- A statement that their information may be shared with external organisations but also that this only happens in a way that conforms to Information Sharing Agreements.
- A statement of what their information will specifically not be used for.
- A statement of the rights they have over their personal information.

Finally, steps would need to be taken to ensure that the Privacy Notice is being used in a standardised format across all forms and points of contact, including paper based and electronic.

New guidelines for the use of data in research

A set of guidelines about how to use personal information for the purposes of different types of research and analysis should be created and circulated amongst staff. This could include:

- A clear statement of what is and is not acceptable under current legislation and national guidance
- Procedures and restrictions specific to Hounslow
- Details of tools and sources of information available

New guidelines for the use of data in service redesign

A set of guidelines about how to use personal information for the purposes of service design should be created and circulated amongst staff. Specifically this should include details of procedures and guidance to support the sharing of information to offer additional benefits to residents. It should also set out how to handle any opt outs.

Revised Data Sharing Agreements with other organisations

A review of existing Information Sharing Agreements would be needed to bring them into line with the framework as well as developing revised procedures for making new agreements.

A communications strategy

A strategy needs to be developed to ensure that appropriate publicity and communication activities are carried out to support the framework.

Appendix 4 - Development Roadmap

This is a list of suggested new features or areas for improvement that could form part of the future development of the Cost Calculator tool.

- Add a reporting API to enable integration with other systems
- Improve the client matching algorithm with more sophisticated matching techniques
- Improve the client matching algorithm by including external address look up (possibly making use of UPRNs)
- Extend the tool to work with 'households' as well as with individuals
- Improve the handling of duplicate cost records
- Allow for the "manual" refining of client records (matching, merging and splitting of records) during import
- Improve the management of service type categorisation (such as editing labels or combining categories)
- Improve the filter options and filter user interface
- Add additional types of reports
- Make use of the additional data fields recorded but not currently used
- Improve the security features of the data and tool
- Improve the settings management and customisation making configuration available within the user interface
- Add advanced schema customisation
- Add a more sophisticated client lookup
- Allow for data import using additional file types
- Allow for direct data import via connections with external databases
- Improve the process of installing and setting up the tool (including improved working with different database types)