Department for Education

Analytical Review

Data Systems (By Roger Plant)

April 2013

Contents

1. Introduction

Data is the fuel for the intellectually curious education and children's services community and is essential to an evidence based system. Data is, however, sometimes felt to be a burden and not a driver for improvement. The terms of reference of this review asked that we find a system where only data essential to the Department and to a devolved self-improving system is collected. Such a system needs to be efficient, trusted and must maintain appropriate standards of quality. To meet this ambition this report recommends transforming the way data is collected, used and held so that:

- Data can be automatically moved from one organisation to another with no manual intervention.
- Data is updated on a real-time basis as part of day-to-day business processes.
- Data is available to all who need it for decision making when they need it.

The Department needs data to inform delivery, allocate funding to schools, safeguard standards, maintain public accountability and develop policy that is grounded in evidence and improves outcomes for children. The recommendations in this chapter, which build on infrastructure already planned with only minimal additional technology, will enable the Department to conduct this business more efficiently, generating savings for the Department and for delivery organisations.

Reforms set out in this chapter will also support an autonomous education and children's services system to drive improvements. Delivery organisations need faster access to more data to support local decision making. It will be important to work in a phased way, initially with schools and local authorities that are already committed to the approach, in order to build evidence of the benefits and identify the challenges and how these should be overcome.

2. The current situation

The technology in the Department is not in line with current industry standards and is almost ten years old. The Department operates reasonably efficiently given the constraints of the systems. However, the systems drive considerable inefficiencies in data collection and use, which not only raise the costs incurred by the Department, schools, colleges and children's services, but also hold back the development of a self improving system. Current

systems will soon become over-expensive to maintain, if not obsolete. They are also not fit for purpose in an increasingly autonomous school system. The quality of data will be compromised and costs will rise as the Department struggles to manage set-piece collections from 24,000 schools, rather than 150 local authorities. Sticking with the current approach is not an option.

The key factors driving costs are:

- The number of different systems used to collect and deliver the data with each needing specific new programming, documentation and user training whenever data is changed.

- The same data is being changed and checked in parallel by different people at different processing stages without modifications feeding back to the source (the delivery organisation which generated the data).

- That the Department is dealing with ever increasing numbers of delivery organisations during this checking process. As more schools convert to Academies the Department must check data with thousands of schools individually rather than through 150 local authorities.

- The fact that data held by schools and other frontline providers on their systems is not the same as that held by local authorities or the Department. This means that data has to be reconciled every time there is a data collection.

Beyond immediate costs the current system also prevents the use of data to drive improvements because:

- School, college, local authority and Department resources are spent checking and reconciling data rather than using it.

- The process of collecting and checking the data is slow and causes delays in identifying risks and issues.

- New requirements on collection and use of data necessitate a development cycle of 18 months from concept, through to development and use that limits opportunities for timely decision making, and can lead to short term ‘solutions’ that may not facilitate wider data usage.

There is a clear need for a real-time system that will manage data faster, make it available more widely and cost less to maintain.

3. Vision for reform

We need to move to model where:

- Data can be automatically moved from one organisation to another with no manual intervention.
- Data is updated and shared on a real-time basis as part of day-to-day business processes.
- Data is available to all of the people who need it for decision making when they need it.

This would mean an end to the manual steps in transferring data from management information systems into the Department's collections - instead data would be sent automatically. It would mean comparative data could be compiled for Ministers, policy makers and third parties when needed with no additional burdens on delivery organisations. The proposals outlined below are not about changing school systems but putting the mechanisms in place to let different systems talk to each other. Delivery organisations would also be able to get data in real-time at the point they need it to drive improvement.

It also opens up new opportunities to share data between delivery organisations. For example, it would be possible for schools to automatically share daily pupil attendance data with the local authority so that welfare staff can directly monitor children who are looked after. This already happens on a daily basis in Warwickshire and twice daily in Hampshire where the model has already been implemented. This is just one example of a vast range of data that could be managed in this way. Academy sponsors could likewise benefit from faster access to data across their schools - "too much time is spent working with data about children that have already left – we need systems that can tell us about what is happening now" (Sophy Blakeway, ARK).

To achieve this vision the Department should put in place the standards and technology that will enable the automatic flow of data. This can be done using off-the shelf technology that is widely used commercially, that will:

- Collect only the data that is needed at a given moment rather than larger set-piece collections. This will support the gathering of timely and less burdensome data exchanges.
- Automate data checking and validation, with responsibility for accuracy pushed back to source.

- Provide real-time feedback to delivery organisations and third parties in a format that can be used for comparative analysis. The more data is used and is useful, the greater the incentive to increase accuracy and so the system becomes self-cleaning.

- Enhance the efficiency and effectiveness of the data warehouse and the School Performance Data Programme as data will be more readily available and up-to-date.

- Draw on existing industry-wide standards so that data sets can be efficiently matched together, whilst also allowing new categories of data to be added for use in schools, authorities and the Department at minimum cost.

- Automate data collection wherever possible for all education and children's services. For example, through agreement with schools, attendance data or free school meals data could be regularly accessed.

- Support teaching and learning directly. The system will be able to cater for broadening data demands particularly in relation to performance and pedagogical data held in systems such as learning platforms.

- Reduce the costs related to the management and use of data both within the Department and throughout the education and children's services community.

4. Moving from collections to automatic exchange of data

Three things are needed to automatically move data between different management information systems:

- A data warehouse to hold data – this is already being put in place.

- A data model to define common standards for the data - this is already being developed by the Department's Information Standards Board.

- A transport system to move the data – this is the key piece to add that will let the different systems talk to each other. It would consist of servers and software which direct the flow of data between existing information systems.

Minimal technology would be needed to move data around the education system in this way. The technology available today allows this approach to run on existing infrastructure. There are a range of different technology providers already in the marketplace delivering off-the-shelf solutions. Many schools already have the software in place that could direct the flow of data. Expertise in this technology already exists in the Department through ongoing collaboration on Information Standards. Such data exchange systems are often referred to as interoperability systems.

The Department should be responsible for overall governance of the system, from defining requirements through the design and development of the system to its delivery and operation. It is likely that the development and delivery stages can be best managed through outsourcing.

5. Benefits, risks and feasibility

There are three key benefits of implementing this approach. First, the Department could access and use data more efficiently. This approach could reduce the core Departmental staff resource on data storage and collection by over half (currently ~60 FTE). It could also reduce the staff resource on statistics and data distribution (currently ~70 FTE) by around a third. Staff savings of between £1.4 and £2.6m per year could be achieved once the system is fully up and running. The Department would also avoid £300,000 per year programme costs of upgrading the collection system. On top of this, officials and Ministers would have faster, easier access to data when they need it.

Second, it would improve efficiency in delivery organisations. Local authorities would not need to spend so much time quality assuring data. Schools, colleges and local authorities would also save time by not having to manually manipulate or cut and paste data between different systems. Schools could also save money by not having to manually contact other organisations to chase data, for example on admissions or free school meal eligibility. Based on examples from local authorities that have already partially implemented this model, tentative estimates suggest these savings could total around £20-£40m per year across the country in reduced administrative burden. These savings are likely to be higher with full implementation covering a wider range of data items.

Third, there would be additional benefits to delivery organisations from sharing data at a national level. This might include, for example, sharing

timetabling, workforce information and pupil assessments. This would help planning and improvement activities at a local level.

This model has already transformed efficiency and delivery in the areas in which it has been used:

- In Warwickshire, core student and teacher information, student attendance data and free school meals data is being exchanged in real-time between all schools and the local authority. Schools report a significant reduction in administrative burden and a huge benefit from real-time comparative data within the authority.

- Interoperability has been introduced into a number of states in Australia to enable the real-time sharing of data across schools and states. This enables pupil learning and attainment records to be shared across a multiplicity of different providers, improving learning experiences.

- In Northern Ireland, an interoperability system has resulted in administrative savings in schools by sharing data between management information systems and learning platforms without requirement for re-entry. The intention is to take this approach further so that data related to timetable, classes and parent information is automatically fed into the central Northern Ireland education system data warehouse, with scope for greater comparative analysis.

The biggest challenge lies not in getting the technology in place but in changing attitudes to data sharing and use. Some providers may be uncomfortable with the Department or other delivery partners having access a wide range of data and may want to 'sign off' any data before it is used centrally. If delivery organisations are to be accountable for their data more formally this represents a significant cultural shift in the current approach to data input. This model will also depend on providing the right incentives to keep source data accurate without rounds of re-checking. Phased implementation and data tools that meet delivery organisations' needs will be important to tackle these challenges. It will be necessary to make sure that all changes to the use of data comply with relevant legislation, including the Data Protection Act. Financial data in particular will still need to be audited, with checks and balances in place to guard against fraud.

The small sample of local authorities and schools we spoke to (maintained and Academy in both primary and secondary phases) have reacted positively to the idea of real-time data. There will be some cultural and trust issues to

overcome to make it a complete success. The issue of governance and access is paramount, and this was expressed particularly by local authorities with regard to data connected to children in need. Some schools did not like the Department having access to their data whenever it wished.

Concern was also expressed over the length of time likely to be needed to implement a real-time system, as put by Dame Reena Keeble, Headteacher at Cannon Lane First School: “my only regret is the likely time needed to implement such a system”. Sophy Blakeway, Director of Education ARK Schools, whilst welcoming the proposals expressed concern over the timing: she felt that at a time of transformational change to both curriculum and assessment, there was an increasing need for a more immediate and supportive data system. With this in mind the Department should take steps to expedite the work necessary to bring forward the delivery of the data warehouse together with the development of new data standards.

Benefits stack up very favourably against costs. Set-up costs for the system are not prohibitive (<£2m) and it builds on existing IT infrastructure rather than being a large centrally mandated IT project. Further work would be needed to scope accurate costs in detail, but we estimate there would be ongoing programme costs of between £2m and £4m to run the system, depending on how much of the infrastructure is already in place. However, as there are clearly evidenced savings and benefits to schools and local authorities, some or all of this cost could be redirected from existing funding streams rather than being an additional cost to the Department. Taking into account savings to delivery organisations and the Department, we estimate the net present value over ten years would be between £110m and £190m.

6. Recommendations

Recommendation A: The Department should procure an interoperability system for the exchange, use and maintenance of all data.

Recommendation B: An appropriate data model should be developed through collaborative design between users and system suppliers coordinated by the Information Standards Board.

Recommendation C: The Star Chamber should be used to ratify the inclusion of data items in the data model. The remit and membership of Star Chamber should be reviewed as part of the changes to the system.

Recommendation D: Data should be gathered and used in real-time, as part of the day-to-day business processes. Some formal collections may

still be timetabled (for example core pupil data - once a year) whilst others will be running regularly (for example attendance data gathering - weekly).

Recommendation E: The data exchange process should include a data validation system designed to ensure that data is always validated at source.

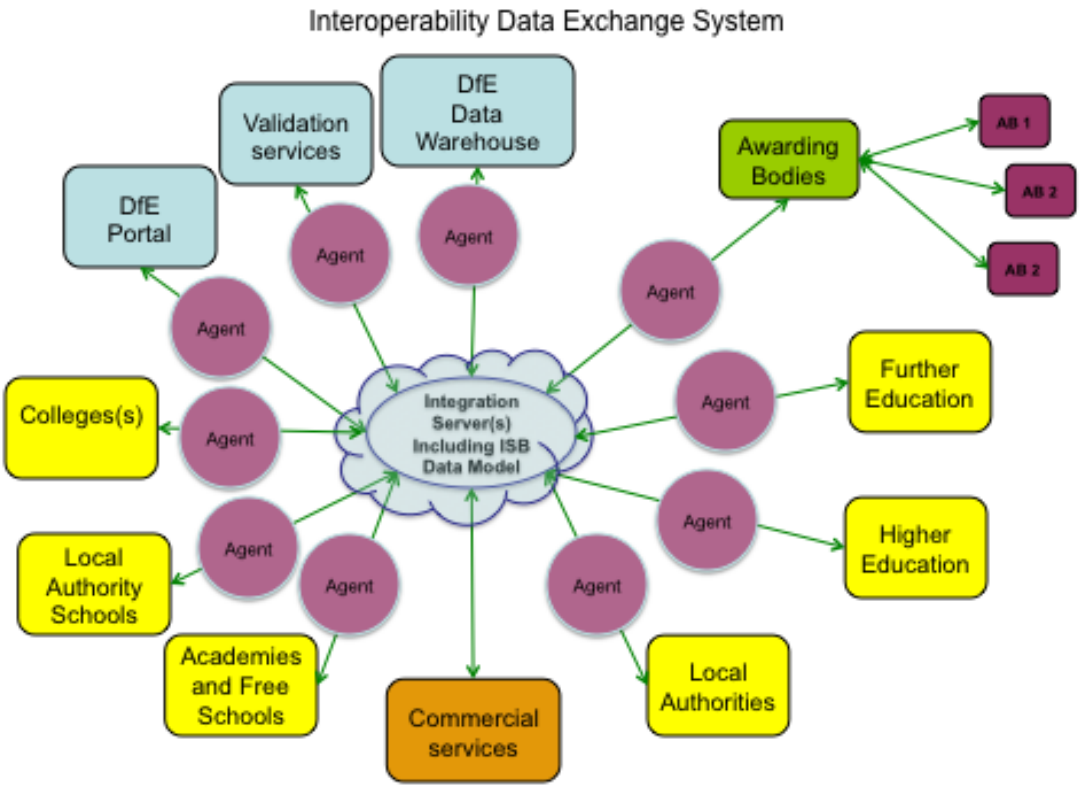
Recommendation F: The data validation service should be able to be used independently of data exchanges to help ensure the accuracy of source data.

Recommendation G: The Department's data warehouse should become the one-stop-shop for all education and children's services data in England.

Recommendation H: The system should provide analytical tools within a Departmental web portal offering controlled access to reports and data extractions for the community.

7. Features of data interoperability systems

In operational terms an 'interoperability system' for the Department and the whole sector could be represented as follows:


Interoperability Data Exchange System
DfE Data Warehouse
Validation services
DfE Portal
Awarding Bodies
AB 1
AB 2
AB 2
Agent
Integration Server(s) Including ISB Data Model
Colleges(s)
Further Education
Higher Education
Local Authority Schools
Academies and Free Schools
Commercial services
Local Authorities

The key features of this system are:

Feature	Implications and benefits
It uses the Internet as the medium for data transport.	Communication channels are largely in place with transport systems being added to the infrastructure (e.g. data warehouse) that already exists.
Data is gathered as non-aggregated, granular or atomic data in real-time replacing data collections as we know them today.	Smaller packages of specific yet simpler data are assimilated within the system as-and-when needed.
Data is not held in the system but simply 'transported' through it	The security of the data in transport is not threatened as it is not stored.
The 'agent' is a piece of software that enables data packages to move as messages between source and destination in a common format and in a controlled way.	Individual participating systems can remain unique whilst data can be translated through the agent as it is exchanged.
All parties use a common data model (created by the existing Information Standards Board) in order to identify data items that are transported.	Data confusion is minimised and validity guaranteed.
There is no limit to the numbers and types of members of the interoperability network.	The system is available to all members of the education and children's services community and could be extended beyond. e.g. health or commercial services.
The Integration Servers act as controllers in directing the movement of the data packages from their source	Once governance of the system has been established by the community of users the system becomes self-managing and extendable, when additional data is identified

Feature	Implications and benefits
to their destination.	as being needed for exchange and agreed by Star Chamber.
Data can be accessed automatically or as a response to a directed request from one site to another or others.	This requires less human intervention and thus reduces support costs. The system is designed to be timely and could be daily as in the supply of attendance data; or weekly as in the checking of free school meals data; or termly as in the checking of core pupil data, such as addresses.
As data is moved through the transport system it will be checked for validity and only valid data can be moved once it has been corrected at source.	Data is always corrected at source preventing the need for expensive and often repeated validation processes along the transport route.
The validation system can be called at any time by sources to check upon the validity of the data they hold.	Provides a route to ensure the accuracy of data with the source, thus enabling users to check data validity as it is entered.
There is no limit to the number of data items that can be managed and transported although software 'agents' will require controlled modification which will incur a minimal maintenance cost.	The base system can extend as required over time with new data entities being added by participating parties to the data model – this could be governed through Star Chamber.
Data can be moved into systems as well as collected from them.	For example, free school meals data from local authorities can automatically be fed into schools systems, as could exam results.
Through the DfE portal any authorised user can access the latest data from one	Through a set of system reports, data extracts and user-defined reports, analytical and tabulation/graphic tools can be made available

Feature	Implications and benefits
place.	as a free and/or charged service. This could be made available to parents, learners, teachers, governors, LA officers, Department staff, MPs, third parties (e.g. research groups) or anyone with a legitimate right to access the data.

Department for Education

Any enquiries regarding this publication should be sent to us at www.education.gov.uk/contactus.

This document is available for download at www.education.gov.uk/publications.

Reference: 00047-2013PDF-EN-02