Update on the progress of the 'Local Data Ecosystem' open data project, led by Tameside Council, as at 15th October 2014

Paul Davidson, CIO Sedgemoor District Council, for the Local eGovernment Standards Body (LeGSB), 15th October.

The project has set out to understand how data can be re-used across multiple agencies serving communities, for cross-theme issues. We have chosen to illustrate the approach using a theme of 'troubled families', but we could equally have chosen to look at 'dementia care', or 'flooding' and so on.

The approach is to define and prototype a 'local data ecosystem' where data can be proffered and consumed, by many agencies, based on common standards.

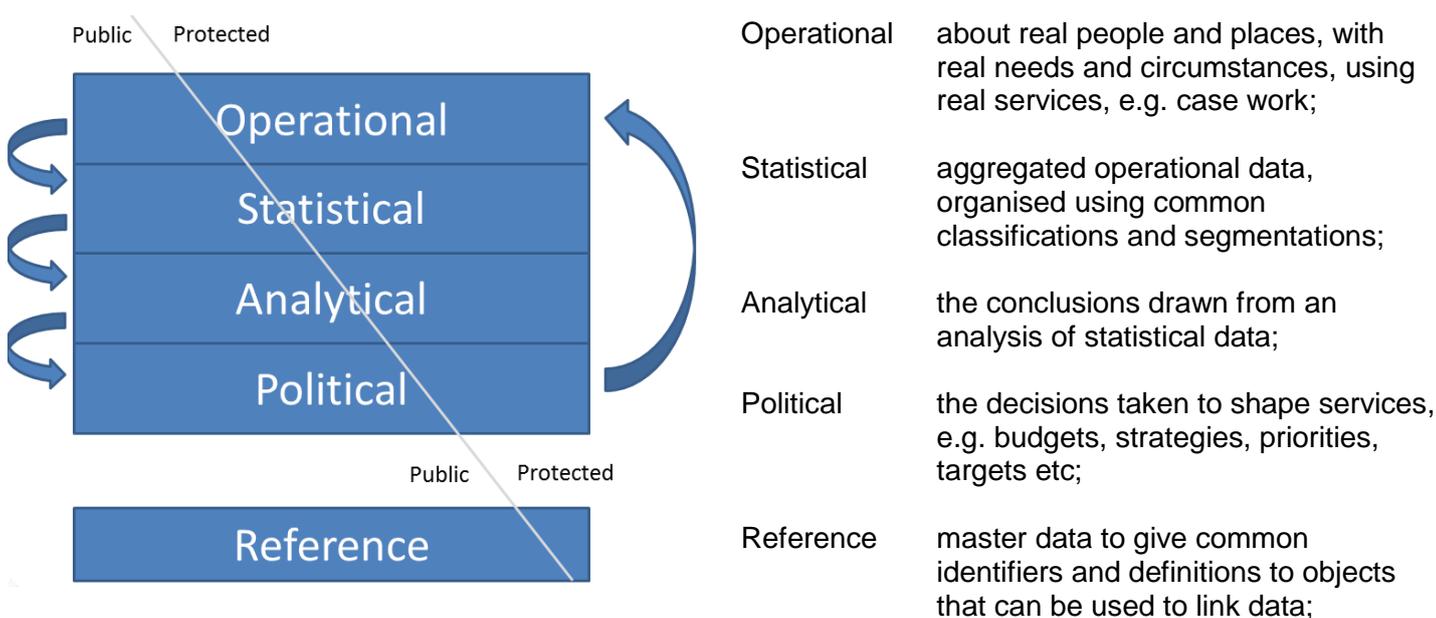|	
Operational	about real people and places, with real needs and circumstances, using real services, e.g. case work;
Statistical	aggregated operational data, organised using common classifications and segmentations;
Analytical	the conclusions drawn from an analysis of statistical data;
Political	the decisions taken to shape services, e.g. budgets, strategies, priorities, targets etc;
Reference	master data to give common identifiers and definitions to objects that can be used to link data;

Fig 1

The scope of the data, and potential for re-use, is as depicted in figure 1, where a 'golden thread' can be followed from operational data, through to evidencing decisions, and observing the impact of those decisions back in operational data.

Operational data is often protected, particularly when it refers to identifiable people, whereas the insight that could be derived as statistics and analytics are more shareable, both with other agencies serving the same communities, and with citizens.

The 'Troubled Families' scenario has all of the dynamics of this desire to re-use data.

Operationally
- councils need to identify families, where the members collectively have certain circumstances.
- councils need to be alerted when a member of family in their programme, has had a role in a service, often provided by another agency.
- councils need to commission intervention services from other agencies

Statistically
- the case load needs to be monitored and reported over categories such as risk factors, circumstances, interventions applied, neighbourhoods affected
- the throughput and cost, of 'reactive services', across many agencies, needs to be monitored to evidence the impact of intervening with families in the program.

Analytically
- data can be analysed, 'in lab conditions' to draw out correlations and patterns, to reach

conclusions about what type of intervention works well for families of certain risk types.
- statistics can be analysed to consider the value-for-money and success rates of programmes.

Politically
- the budgets and targets for a programme can be reported, linked to objectives and assumptions

To make the 'local data ecosystem work, the project has worked with the Local eGovernment Standards Body (LeGSB) to consider seven types of data standard.

- Semantics – the meaning of information
- Syntax – the format of information
- Data Quality – the confidence to use information
- Rights – permission to use information
- Trust – who is accessing information
- Transport – how to move information
- Information Governance – the behaviour and culture to protect and exploit information

The ecosystem will ultimately require an agreed standards framework for each of these, but initially, we have focused on 'Semantics' and 'Rights'.

Semantics

The blueprint for the prototype includes an 'extensible data dictionary' in which the data structures required for a theme can be built, and then built open further by other themes.

Via workshops with councils in the north west, the project has defined the 'shapes' of data that can populate the 'operational' layer of the model, and repurposed data to fit that model. One of those shapes, a Case, is shown in fig2, together with how a dashboard view of data can appear in the ecosystem, that matches that shape.

As each pattern in the model is established, the approach is to consider if that pattern also exists generally in other service delivery scenarios. For example, the idea of a 'Case' could also be applied to a planning application or a pothole, in that they are raised within a SERVICE managed by an ORGANISATION. By extending the idea of (in this example) CASE, the ecosystem can discover and report on, all CASEs managed by an ORGANISATION of whatever type. In this way, the dictionary is founded in a series of core concepts and truisms.

Indeed, we have imported the planning applications published by the Hampshire County Council led project into the prototype ecosystem, and query all CASEs which then includes Planning Applications and Troubled Families. Of course, that view of a mixture of data over many themes would require that the rights element had been passed. If the data is viewed statistically, it might list the caseload of all services offered from many organisations.
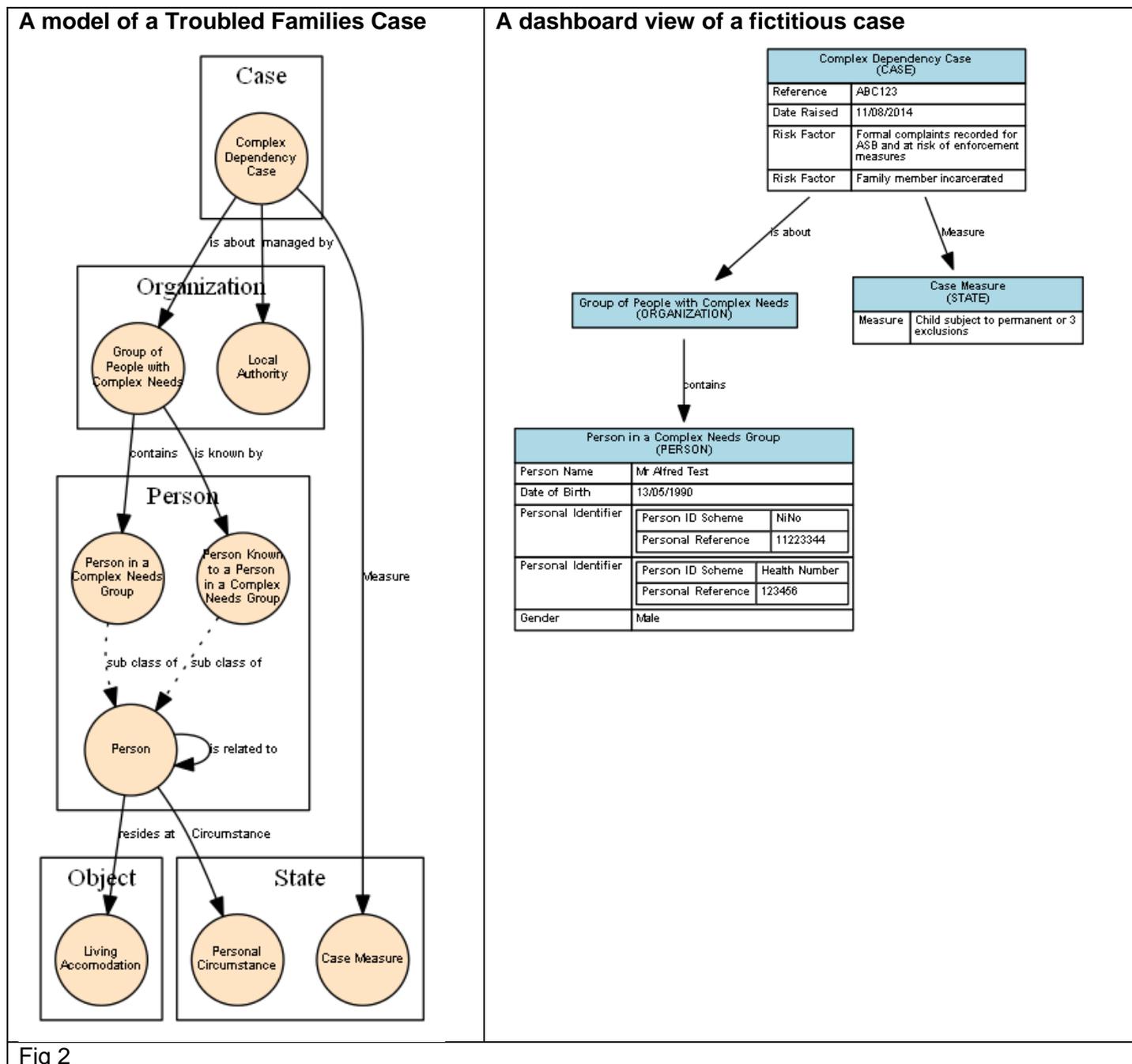
A model of a Troubled Families Case	A dashboard view of a fictitious case

Fig 2

Rights

Representatives from the Information Commissioner's Office have joined our workshops, together with council information governance professionals, in the north-west, to generalise an extendable approach to ensuring that data is re-used used legally. The prototype for the Local Data Ecosystem is designed to have rights and trust built in, rather than just as a series of document that describe data sharing protocols.

The workshops established a 'handshake' for sharing data in which

> A **person** is **empowered** by their **organisation** to act in a **role** that has a **right** to a set of **data items** for a **purpose,** and agrees to the **terms** by which the data is to be used and handled.

The project is therefore building into the prototype, the ability to register a series of 'purposes', 'data protection schedules' and 'legal gateways', 'accreditations', and 'licences' to be associated with datasets so that the data known to the ecosystem can be properly protected, and discovered.

Next Steps

The project is planning a further series of sessions in the north west, to take on the phase 2 requirements of the Trouble Families programme, and further develop the prototype ecosystem to illustrate the potential for using these types of standards.

We also intend to run an open day for developers and suppliers of solutions to local authorities to present the prototype multi-agency ecosystem and debate how a market can form for components that implement the approach.