

Data Standards Report for the Data Delivery Group – 23rd July 2020

1. Summary

This report provides an update on Data Standards, the activities that are taking place and recommended next steps. It recommends that a Data Transformation Framework should be developed to enable organisations to improve data management and reuse, support business resilience and drive data innovation. The Data Delivery Group are asked to:

1. Note the progress on Data Standard activities and expected outputs in August
2. Provide feedback on the proposal for a Data Transformation Framework by 7th August
3. Agreement to pilot development of the Framework using the Data Standards, Skills and Data Management categories, initially expanding detail to validate the proposal
4. Support the development of a Data Maturity Assessment model and dashboard for use by the Public Sector.

2. Context

Data sits at the heart of digital transformation, and brings to life our vision that data is systematically used to improve decision making: saving time, money and lives. To meet this ambition requires the adoption of a National Data Infrastructure (NDI) with key elements around Governance & Vision, Technology and People & Enablers. The creation of this NDI, with targeted interventions that are aligned to the forthcoming Digital strategy, will help build economic, social and environmental value. It will also support our existing commitments set out under the Programme for Government:

- **Ending Scotland's Contribution to Climate Change** – innovating for (incl. Skills) net zero, climate change adaptation
- **A Successful, Fair and Green Economy** - Supporting innovation and adoption of technology, Digital and Data, Digital Infrastructure, Digital Public Services, Data Driven Innovation, digital and data ethics
- **Improving Outcomes Through Our Public Services** - Integrated health and social care service, working across the public sector

The data ecosystem in Scotland is complex and diverse and data is held and managed, at varying maturity levels, by many different organisations, but the value of data lies in its use and reuse as appropriate for the public good and to realise the potential of data for the people of Scotland. To unlock this value the DDG, as part of their Data Action Plan, have set out the priority areas for focus such as the development of an AI Strategy, reform of Information Governance, work around privacy principles and Research Data Scotland.

Whilst a range of work is currently underway to enable better use of data, data standards is the cornerstone to delivering this outcome successfully and sustainably. This paper sets out on an approach to Data Standards¹, that could be used for other core data management components, that aims to move organisations through a data transformation. This approach will help the development of the NDI and will engage with work underway on the AI Strategy and other priority areas identified by the DDG. For example, the work underway on the AI strategy is already highlighting the need for consistent Data Standards² to realise the potential of AI, and the growth of data reuse. There are a

¹ The definition for Data Standards here is how data is described and recorded at a technical level - for exchange, use and reuse. Standards in scope include elements such as unique identifiers, schema and formats.

² The AI Of The Possible : Initial Consultation Report several respondents noted 'the importance of developing effective data and digital infrastructure, and ensuring quality and standards of data, to help ensure the successful delivery of the strategy'

number of other projects and programmes of work that are also currently looking at data standards, these include:

- Scottish Cites Alliance – Data Cluster Steering Group Data Standards activities
- Scottish Government Planning – PathFinder Projects (led by IS) looking specifically at Data Standards for planning data
- Digital Office: Scottish Local Government – Covid-19 cross sector task force
- GDS/ONS – Open Standards Board & UK Data Standards Authority
- Within SG: New Horizons – Digital First, Platform Authority, RDS, Open Government Partnership
- Thematic areas: Police Scotland, Health & Social Care, Social Security, HES

3. Impact of Covid

COVID has shown us why good data is essential, and fostered a greater understanding of the value of high- quality data and data analytics. The crisis precipitated data innovation at scale to provide information, products and services to support both government and the citizens of Scotland. For Public Sector organisations, data became essential for emergency planning, health interventions, resilience, and service delivery. COVID 19 data analysis had to be carried out at pace to inform the FM and Ministers, this raised the level of innovation in Government, accelerated Research Data Scotland timings and led to the establishment of a COVID 19 Data and Intelligence Network whose aim is to ensure that the public sector is provided with real-time intelligence to make rapid decisions.

It has also demonstrated once again that not all organisations have the data maturity³ to optimise their services and that drivers are needed to move organisations through data transformation just as is being done with digital transformation. To do this requires organisations to understand their current data status and identify priority areas to address.

4. Data Standards update

Online workshops⁴ are currently underway with specific data communities (Scottish Government, local government, health, the environment and society) and these workshops are highlighting the complex situation around data standards in Scotland. There are many examples of good practice but appreciation of the importance of data standards is limited to pockets of expertise and to meet our data vision, we need to have more people engaged and active. This is a consequence of limited recognition of the value of data across large parts of the public sector. If the value of data is not recognised, the value associated with the development or adoption of data standards will not be recognised. The workshops have also highlighted the need for a clear set of data standard principles such as metadata, schema and formats - to support consistent data management and enable data sharing in public sector organisations.

The expected final outcomes from workshops, due to be delivered in August, will be:

³ Digital Office COVID response workshop 2 July 2020 – data lacks unique referencing, problems with interoperability, data re-use

⁴ Workshops are being facilitated by Anderson Solutions commissioned to: create a contact list of those in Scotland's public sector with an interest in the development of data standards; facilitate workshops that develop the ambition for data standards and build a roadmap of key steps to be taken; and develop case studies that can demonstrate need for and/or benefit of data standards in Scotland's public sector.

- greater knowledge of the landscape for data standards in Scotland’s public sector, including greater understanding of barriers and opportunities, lessons learned and complementary initiatives;
- greater clarity around direction of travel and priorities for data standards;
- a community of interest;
- an understanding of why there might be a lack of engagement and what needs to be done to address this issue.

However, the workshops will not help with questions of how to define a clear set of principles on data standards, or how to raise their visibility. In tandem there is the challenge of engaging more people and helping organisations understand how to transform data to make it more accessible, interoperable and ultimately more re-useable. The Data Transformation Framework discussed in the next section will help with these challenges by providing a viable approach, including a method of measurement for organisations that can be easily understood, motivation to collaborate and support in identifying priorities for improvement against their business outcomes.

5. Data Transformation Framework⁵

Why it’s needed –

- Inconsistent data quality and management is a known and common issue, highlighted in responses to COVID 19
- Disparate working practices are barriers to and reducing opportunities for data interoperability and reuse
- Lack of a coherent data infrastructure, data hub and models means organisations are creating their own without shared standards, architecture or protocols
- The lack of data skills at all levels is an increasing barrier to realising the value of data⁶
- New data types, uses and the impact of new data growth through developing use of AI in the public sector, with growing use of shared data, has highlighted a need for an ethics framework that is citizen centred and adapts to change
- There is a growing demand from public sector organisations for clear governance, policies and principles to guide data management and decision making on use of data
- changing technologies and use of data requires wider consideration of user needs and mitigation of risks as well as awareness of ethical and social responsibilities for data use.

The aim of the Framework will be to improve and enable data reuse in the Scottish Public Sector. It will outline realistic and measurable maturity stages – setting out what ‘good data’ looks like and the process by which organisations can improve. It will highlight drivers and associated data activities using data maturity pathways that organisations and agencies can align to their priorities. A core driver for the development of this Framework is to highlight the need for Data Standards, but other categories such as Ethics and Social Responsibility, Skills and Data Management are also central components to its development (see Appendix 1). Integrating and standardising across these categories will ensure organisations and agencies can develop and embed data in their business activities and improve data reuse.

⁵ See Appendix 1 for the Framework’s aims and objectives

⁶ There is recognition of this need Example initiative: the Data Lab is developing the Data Skills for Work (DSW) programme which sits within the wider Data Driven Innovation Skills Gateway (part of the Edinburgh and South East Scotland Region Deal). (<https://dataskillsforwork.com/>).

The foundation of the Framework will be a high level Data Maturity Assessment model and dashboard where organisations can measure their current status and review progress, and choose to benchmark with other bodies. Data pathways will be designed to help organisations identify priorities to address, developed around data transformation categories and the Framework maturity stages⁷. The categories define areas of data transformation which organisations need to consider. To support the data maturity pathways, the Framework will provide a gateway to a range of resources both existing and developed. It will be a living resource, a Knowledge repository that organisations will add and update over time.

The Framework is not developing from scratch but will be building on existing projects and resources, the work of expert groups, and building too on a collaborative data community which already exists in Scotland. Organisations will co-design the Framework, seeking participation from potential users as well as developing it using Open Government principles. As organisations use the Framework, priorities for further development will be identified, reacting to user needs, changing technologies and priorities for the sector. The Framework's core goal is to provide a shared vision for data transformation in Scotland to enable responsible and beneficial data reuse and innovation. The Framework will be developed using co-design principles and open government principles.

Benefits (post COVID renewal)

- National benefits
 - Supporting the vision for Recognising Scotland's Data Potential
 - Supporting the delivery of the Data Delivery Group plan
 - Facilitating the development of the National Data Infrastructure
 - Align with/enable/support Digital First Standards (D3 standards)
- Strategic benefits
 - Developing an open knowledge bank and data community open to all (organisations realise the potential of their data)
 - Building wider citizen collaboration and participation to sustain greater trust in data management, publication and use
 - Identifying partner opportunities for data improvement and innovation
 - Building cross sectoral collaboration to scale up existing practice and data innovation
 - avoid fragmentation of local initiatives that involve the management and publication of data
- Organisational benefits
 - Support organisations to develop & realise value for the public from their data through sharing & reuse

⁷ The Framework design will help organisations to choose pathways according to their business priorities.. The Framework can also be used to support individual pieces of work, providing resources to address particular requirements.

- Enable organisations to meet changing data needs and compliance, as well as develop robust cultures of transparent, ethical and responsible data practice
- Cross agency policy i.e. Covid (health & social care) and child poverty where data from multiple sectors is needed i.e. health & mobility data

Data Standards Framework application

Taking Data Standards as an example category from the Framework, we can apply a clear set of data standard principles, for example, the FAIR principles, which are used by a large number of private and public institutions.

The FAIR Data Principles are a set of guidelines for good practice when managing data, originally designed by stakeholders in academia and the private sector. The principles are based on transparency, reproducibility and reusability and applies to data, metadata and supporting infrastructure (i.e. tools and workflows). The FAIR acronym stands for Findable, Accessible, Interoperable and Reusable and were coined in the scientific article from Wilkinson et al. (2016)⁸, they provide a common framework that different users can take for a unified approach to data management – therefore, different interpretations and implementations of the FAIR principles exist in different disciplines.

Example use cases include VODAN, set up in response to COVID⁹ and the Australia National Data Service (ANDS)¹⁰. ANDS asks data owners to implement the FAIR principles and to help with this they provide a [FAIR data self-assessment tool](#) which assesses a dataset and determines how to enhance its FAIRness. They also use FAIR to structure their matrix of resources (see diagram 1 Appendix 2.), users simply click on the tiles to access the available resources. For example, clicking on ‘Rich Metadata’ will take you to resources showing all the relevant guidance on metadata for different types of data (i.e. geospatial schemas), how to store metadata and vocabularies for metadata.

Using ANDS as an example, data owners would be able to measure their status in terms of data standards through the self-assessment tool (for the Framework as a whole, this would be done using a Data Maturity Assessment). Data owners can then move through the Framework using maturity pathways to improve the status of the data, like this pathway defined by FAIR - <https://www.go-fair.org/fair-principles/fairification-process/>. To support the Framework pathways, ANDS have provided this FAIR gateway to policies, strategies, reports, courses, case studies, templates and other resources.

4. Recommendations

This report has noted evidence of demand for both Data Standards programme of work and a Data Transformation Framework.

Recommendations

1. Note the progress on Data Standard activities and expected outputs due in August

⁸ <https://www.nature.com/articles/sdata201618>

⁹ The Virus Outbreak Data Network (VODAN) Implementation Network (IN) was conceived to kick-start a ‘community of communities’ that could design and rapidly build a truly international and interoperable, distributed data network infrastructure that supports evidence-based responses to the viral outbreak. As a GO FAIR IN, VODAN will restrict itself to projects that are directly associated with FAIR data [1] and services relevant to COVID-19. (<https://www.nature.com/articles/s41431-020-0635-7>)

¹⁰ Now the Australian Research Data Commons

2. Provide feedback on the proposal for a Data Transformation Framework by 7th August
3. Agreement to pilot development of the Framework using the Data Standards, Skills and Data Management categories, initially expanding detail to validate the proposal
4. Support the development of a Data Maturity Assessment model and dashboard for use by the Public Sector

Appendix 1

The Framework aims:

- To improve and enable the reuse of Scottish Public Sector data, setting out what 'good data' looks like and the process by which organisations can change/improve
- Help organisations align their business strategy and realise more value from their data
- Support the development of the Scottish Data Infrastructure
- Develop a collaborative and pro-active community building on existing strategies, programmes and partnerships

The Framework objectives:

- Deliver a Data Maturity Framework providing a shared vision for data maturity and the core elements of data management that support this
- Provide supporting resources and models to enable data maturity pathways
- Provide a Maturity Dashboard that allows organisations to measure progress and collaborate with each other
- Support organisations to develop their Data Strategy and plans
- Promote the expert community: stakeholders, experts and practitioners from across the sectors

The Framework categories

| | |
|---------------------------|--------------------------------|
| Categories | Ethics & Social Responsibility |
| Leadership and Governance | Information Governance |
| Data Management | User-centred Design |
| Data Standards | Data Optimisation & Innovation |
| Culture & Skills | Infrastructure & Technologies |

Appendix 2

Diagram 1: ANDS Matrix of Resources

