Central Digital & Data Office

11:00 AM

1st October 2024

10:00 AM -

Virtual Teams meeting

DSA/DQ Hub Peer Review Group Meeting Summary

Attendees

Jenny Brooker, Chair (DSIT) Firoze Salim (DSIT) John Olatunji (DSIT) Ade Dayo (DSIT) Abdullah Hakim Alasdair Gray (MHCLG) Allan Jamieson (OS) Andrew Newman (ODI) Matthew Andrews Samera Allam (Office of National Guardian) Aaron Beck (Planning Inspectorate) Ben Salisbury (IPA) Martin Champion (DEFRA) Deborah Wilson (DVSA) Gavin Ajomale-Evans (DVLA) Amanda Greenwood Hannah Mckenzie (UKHSA) Jesus Alvarez-Pinera (FSA) Laura Hill (DESNZ) Rob Lee (HMRC) Tej Manik (DfE)	Michael Whittaker (DVLA) Michal Poreba (DBT) Neil Cholerton (HO) Paul Davidson (Tameside) Laura Payne (ONS) Priyanthi Perera-Nathan (Companies House) Rachel Davies (DVLA) Rich Clarke (HO) Ria Sanderson (ONS) Shona Nicol (SG) Stephen C Phil Swan (GMCA) Tim Waters (HMT) Matthew Webber (RPA) Yogesh Babbar (CO) Suzanne Fry (DSIT) Jody Parker (DSIT) Phil McGinniss (FCDO) Mark Lewis Graham McKenna (DWP) Megan Clokey (DCMS) Michael Comer (Companies House)

Record of discussions

1	**Welcome, introductions and agenda** Jenny Brooker
	Jenny opened the meeting by welcoming all participants and provided an overview of the agenda and objectives for the session. She highlighted the importance of data quality and how the Data Quality (DQ) Hub Peer Review Group aims to facilitate collaboration across different government departments to address common data challenges.
2	**Data Quality Project Update** *Suzanne Fry*

Suzanne Fry provided an update on the progress and current status of the Data Quality Project. She outlined the project's background, principles, scope of work, and key milestones achieved so far.

Key Points Discussed:

1. **Background and Principles:**
    - The Data Quality Project is a joint effort between the Central Digital and Data Office (CDDO) and the Government Data Quality Hub, which is part of the Office of National Statistics. The project is focused on fulfilling data quality commitments set out in the **Transforming for a Digital Future Roadmap** under **Mission 3: Better Data to Power Decision Making**.
    - The overarching goal is to instil a culture of data quality management across government departments, ensuring it becomes a routine and essential activity.
    - Suzanne emphasised that by 2025, all departments are expected to have resolved 50% of their 'high priority' data quality issues as per the commitment within the Roadmap.
2. **Scope of Work:**
    - The project's primary focus is on Essential Shared Data Assets (ESDAs), which are critical datasets shared across government departments. Improving the quality of these assets will enhance data-driven decision-making.
    - The team is leveraging the **Government Data Quality Framework** to create a sustainable approach for continuous improvement of data quality management practices.
3. **Data Quality Commitment Subgroup:**
    - Suzanne highlighted the establishment of the Data Quality Commitment Subgroup, consisting of 29 members from 16 departments. This group has been instrumental in providing feedback and expert advice on developing materials.
    - The subgroup has conducted four meetings so far, contributing to landscape reviews and development of implementation guidance and a prioritisation framework for data quality management.
4. **Landscape Review:**
    - A comprehensive review of existing materials was completed to support departments in applying the Data Quality Framework and developing Data Quality Action Plans (DQAPs).
    - Suzanne mentioned that short-term recommendations from the review are currently being implemented, while medium and long-term recommendations will be considered in future phases.
5. **Guidance Material Deliverables:**
    - Key guidance materials, including the DQAP Implementation Guide, Data Quality Issues Framework, and Issue Priority Calculator, have been developed and are now ready for the pilot phase.
    - Additional supporting documents, such as guidelines on data structure, data lifecycle, and root cause analysis, have also been prepared for testing.
6. **Pilot and Early Adopters Initiative:**

	○ Suzanne provided an update on the pilot with Natural England which was due to start w/c 7th October, where the new guidance materials will be tested. The objective of the pilot is to assess the effectiveness and usability of the guidance. ○ The Early Adopters initiative was introduced, where departments are invited to test the guidance materials with their own ESDAs and provide feedback on their applicability and comprehensiveness. ○ Suzanne encouraged departments to volunteer as early adopters, highlighting the value of understanding how the guidance fits into both mature and less-established data quality environments. The results and feedback from these early adopters will shape the final version of the guidance materials before they are presented to the DSA Steering Board for endorsement. 7. **Next Steps:** ○ Following the pilot phase and feedback from Early Adopters, Suzanne indicated that the guidance materials would be refined and prepared for endorsement by the Data Standards Authority (DSA). ○ Once endorsed, the guidance will be published on GOV.UK to facilitate wider adoption across departments. ○ Suzanne stressed the importance of embedding these data quality practices within departmental processes to achieve long-term impact. 8. **Questions and Participation:** ○ Jenny Brooker asked Suzanne to clarify how departments can get involved as Early Adopters and whether there were specific requirements for participation. ○ Suzanne shared her email address in the chat—**suzanne.fry@digital.cabinet-office.gov.uk**—and invited interested participants to reach out directly. ○ Suzanne concluded by emphasising the need for continued support from departments and active participation to ensure the success of the Data Quality Project. **Action Items:** ● **Suzanne Fry**: Continue coordinating the pilot and Early Adopters initiative and refine the guidance materials based on feedback. ● **All Participants**: Review the shared materials and consider volunteering as Early Adopters to support the development of the guidance. Jenny expressed her appreciation to Suzanne for the excellent work being done on the Data Quality Project and its progress. She then introduced the next agenda item, inviting Paul to present on the Vulnerabilities Concept Model.
3	**Vulnerabilities Concept Model** *Paul Davidson, IStand UK*

- Paul Davidson led the main presentation on the **Vulnerabilities Concept Model**. The purpose of this session was to walk participants through the conceptual model, explain its connection to the logical model, and discuss the next steps, particularly around defining entity definitions to support interoperability and data sharing.
- Ade Dayo provided a brief explanation of the diagram on Slide 2, emphasising the relationship between the conceptual, logical, and physical models. He outlined how these models form the foundation for creating data standards that can be consistently applied across different domains and use cases.

Detailed Discussion Points:

Introduction to the Concept Model:

- Paul began the presentation by outlining the high-level structure of the Concept Model. The Concept Model is positioned as a foundational structure that informs the development of more detailed logical and physical models. This progression from conceptual to logical and physical models helps ensure that standards are defined consistently and that the data can be used interoperably across various public bodies and other relevant organisations.
- Paul described the importance of having a well-defined Concept Model to serve as a common understanding of key terms and relationships. The model identifies core concepts such as **Vulnerability**, **Purpose**, **Need**, **Outcome**, and **Risk Factor**. Each concept has a precise definition that is intended to be universally understood and applied in the context of data sharing around vulnerabilities.
- Paul emphasised that having this shared understanding enables better communication between departments and organisations, reduces ambiguity, and fosters a consistent approach to addressing vulnerabilities.

Purpose and Application of the Model:

Paul explained that the Concept Model serves multiple purposes:

- **Promotes interoperability**: By defining common data standards, the model ensures that data can be shared effectively across organisational boundaries.
- **Identifies data integration points**: The model highlights where different datasets can be joined, making it easier to understand the relationships between entities such as persons, households, and services.
- **Provides a basis for the Logical Model**: The Concept Model feeds into the development of the Logical Model, which defines more detailed data structures and attributes.

Key Concepts Defined:

Paul and Ade walked through the key concepts within the model, providing definitions and

examples for each:

- **Vulnerability**: Defined as the increased risk of a poor outcome.
- **Purpose**: The remit an organisation has to carry out an initiative covering one or more vulnerabilities.
- **Need**: An unmet requirement of a person or household that can be addressed.
- **Outcome**: A possible result that can be achieved by addressing the needs of a person or household.
- **Risk Factor**: A single fact about a person or household that, when combined with others, can categorise the risk of a vulnerability and predict needs.
- **Information Type**: A generic description of a dataset that one or many organisations may hold, collected for a single purpose.
- **Service**: The ability of a delivery organisation to carry out one or more predetermined methods designed to deliver an outcome.

The definitions provided for these concepts aim to standardise how data is represented and used, ensuring that all participating organisations have a common understanding of these terms.

Link to the Logical Model and Next Steps:

- Paul highlighted that the Concept Model serves as a stepping stone to developing the Logical Model, which provides more detailed definitions and attributes for each concept. The Logical Model will include specific data elements that are necessary to support data sharing and interoperability.
- The next step is to agree and finalise entity definitions for the Logical Model, a process that will require input and agreement from various stakeholders.
- Paul emphasised that while developing the Concept Model was relatively straightforward, defining entities at a logical level to ensure interoperability and data sharing will be more challenging. This will involve not only defining the entities themselves but also agreeing on how they should be represented in data exchanges.

Encouraging Participation:

- Paul encouraged participants to get involved in the ongoing work of the subgroups, particularly those focused on data modelling and terminology. Paul stressed the importance of active participation to ensure that the models and standards being developed are comprehensive and meet the needs of all stakeholders.
- Paul also mentioned the potential for participants to become involved in the new **Enablers Subgroup**, which will address legal and ethical aspects of data sharing.

Feedback and Additional Insights:

- Rob Lee brought up the Government Data Architecture Community (GDAC)

	and its efforts to align data sharing architectural practices across organisations like HMRC, Companies House, and DWP. He emphasised that effective data sharing begins with a clear definition of terms and a solid conceptual framework. ○ Rob suggested that participants from this meeting could join the upcoming GDAC discussions to explore synergies and ensure that the work being done on the Vulnerabilities Concept Model aligns with broader interoperability initiatives. ○ Paul responded positively, expressing interest in joining up with GDAC and emphasising the importance of aligning the work to ensure that data sharing practices are consistent across different sectors. Firoze advised that plans were already in place to link up with GDAC. **Importance of the Work and Commendations:** ○ Phil Swan added that this work is critical for the future of data sharing and interoperability across the public sector. He commended Paul and the team for the progress made so far and the opportunity for participants to contribute to such impactful work. **Conclusion and Next Steps:** ○ Paul concluded by reiterating the need for collaboration and inviting participants to get involved in the subgroups. He shared his enthusiasm for aligning the work with other initiatives like those being led by GDAC and looked forward to further discussions. ○ Jenny thanked both Paul and Rob for their insights and agreed that it would be beneficial to have a follow-up discussion to explore potential collaboration opportunities. --- **Action Items:** ● **Engage with GDAC** and align initiatives and bring their expertise on the vulnerabilities work. ● **PRG Members** to review the conceptual model and provide any comments to the DSA Secretariat by **16 October**. ● **Paul Davidson and Firoze Salim**: Continue work on the development of a logical model and provide updates to the group.
4	**Review of ESDA and Data Ownership BETA** John Olatunji and Firoze Salim
	John Olatunji led the presentation on the **ESDA Initiative**, which provided an overview of the initiative's background, strategic context, and key data and submission statistics.

Key Discussion Points:

Introduction to the ESDA Initiative:

- John began by explaining the strategic context of the ESDA initiative, which aligns with the government's Mission 3: Better Data to Power Decision Making under the Transforming for a Digital Future Roadmap. The initiative aims to establish a single data ownership model and make all critical data assets available for use across government.
- John highlighted that after co-developing artefacts with a wide range of cross-government stakeholders and completing an Alpha phase, the initiative transitioned to its Beta phase. This phase focuses on refining the ESDA process and gathering feedback to improve the final guidance.

Timeline and Outcomes of the ESDA Beta Phase:

- The initial submission deadline for ESDAs was set for January 2024, but due to resource constraints and competing demands, many departments faced challenges in meeting the deadline. The final return date was subsequently moved to April 2024.
- John shared statistics indicating that 29% of central government departments submitted their initial ESDA returns by the January deadline, while 20% met the April deadline for final returns. However, 80% of the submissions were incomplete, reflecting ongoing challenges in meeting the required metadata standards.

Feedback from Departments:

- Departments provided valuable feedback on the ESDA definition and guidance materials, noting a lack of clarity around the types of assets included within the scope, particularly data services.
- Some participants suggested adding a new category for data of high public interest to complement the existing five essential processes or purposes in the definition.
- The term "Essential Shared Data Asset" was not found to be particularly engaging, and departments requested more relatable and practical language in the guidance.

Implementation Challenges:

- John pointed out that many departments did not engage fully with data/information asset owners during initial returns, which limited the completeness and accuracy of their submissions.
- The submission process was rushed due to late notification of expectations, and departments faced delays in sign-off processes from Data Owners.
- To bridge these gaps, the CDDO team conducted bilateral meetings and shared examples of completed templates to support departments in refining their final returns.

Support and Communication Successes:

- Cross-government information sessions and drop-in sessions were considered highly valuable by departments. Continuous communication and support from the

CDDO team were recognised as essential in navigating the ESDA submission process.
- Recommendations for future guidance include providing more real-world examples and practical steps to make the process more relatable and actionable.

Key Findings:

Successful Engagement and Identification of High-Value Data Assets:

- The ESDA initiative achieved its core objectives by successfully engaging departments across the government and identifying high-value data assets that are essential for public service delivery and operational efficiency.
- Firoze highlighted that while there was strong engagement overall, departments varied in their understanding and application of the ESDA guidance, particularly when it came to the provision of metadata for the ESDAs identified."

Challenges in Stakeholder Buy-In and Understanding of Objectives:

- Although departments were able to provide metadata for key data assets, there were concerns around the clarity and perceived benefits of the ESDA framework, especially from data owners and stewards.
- Firoze emphasised that further articulation of the benefits and rationale for identifying and managing ESDAs is necessary to secure stronger buy-in from stakeholders. In particular, there was a need to communicate the strategic value of ESDAs more effectively to departments that were less engaged.

Inconsistencies in Metadata Quality and Submission Processes:

- While many departments submitted metadata for their identified ESDAs, the quality and completeness of submissions varied significantly. Approximately 80% of submissions were incomplete or did not meet the defined metadata standards, indicating an ongoing need for support and refinement of guidance.
- Firoze pointed out that the discrepancies in submissions were largely due to resource constraints, lack of technical expertise, and in some instances a rushed submission process where ALBs were contacted late in the process by their sponsoring department.

Role of Information Asset Owners (IAOs) and Data Owners:

- The findings revealed that there is still some confusion regarding the roles and responsibilities of IAOs and Data Owners within the ESDA framework. Departments with well-established IAO roles expressed resistance to adopting the Data Ownership Model due to additional accountabilities that traditionally fall under other business areas.
- The need to clearly define and align these roles is critical for ensuring successful implementation of both the ESDA and Data Ownership Models.

Resource Pressures and the Impact on Adoption:

- Departments highlighted significant resource pressures as a key barrier to

successful adoption of the ESDA framework. CDDO was advised to explore options for reducing these burdens, such as through automation of metadata ingestion and providing additional support to departments with limited resources.
- In some cases, departments opted to focus on preparing a comprehensive final return by April 2024, rather than submitting an initial return by the January deadline, due to limited staff availability and competing priorities.

Recommendations:

Refining ESDA Guidance and Definitions:

- The ESDA guidance materials should be revised to include clearer definitions and practical examples of asset designation and metadata submission. This will help departments better understand what constitutes an ESDA and how to approach its identification and management.
- It was suggested to add a new category for data of high public interest, which would complement the existing essential processes or purposes within the ESDA definition.

Automating Metadata Submission and Ingestion:

- CDDO should explore the potential for automating the metadata submission process to reduce the manual effort required by departments. This would not only alleviate the resource burden but also improve the accuracy and consistency of submitted metadata.

Enhancing Communication and Support:

- Departments requested more real-world examples in the guidance, particularly around defining and designating ESDAs. Incorporating practical scenarios and case studies can make the guidance more relatable and actionable.
- Continuous engagement with departments is essential to address any concerns about data discoverability, access rights, and resource implications related to the Data Marketplace.

Expanding the Scope of the ESDA Initiative:

- There is an opportunity to expand the implementation of the ESDA framework beyond central government departments to include Arm's Length Bodies (ALBs), local authorities, educational institutions, and health sectors. A phased approach should be adopted, prioritising assets that align with government missions.

Developing a Flexible Data Ownership Model:

- The Data Ownership Model should be refined to clarify the distinction between data and information ownership and ensure that the responsibilities of IAOs and data owners are complementary.

Next Steps and Timeline:

- The draft Beta review document has been shared with PRG members and

	members are asked to provide any comments on its findings and recommendations by 16 October. ● The next phase of the ESDA initiative will focus on refining the guidance materials based on feedback from departments and early adopters. The updated artefacts will be shared with PRG members for review, and the goal is to present the revised documents to the Data Standards Authority (DSA) Steering Board within a month. ● Firoze emphasised the importance of continued engagement and collaboration to ensure that the final versions of the ESDA and Data Ownership models are practical, adoptable, and meet the needs of all departments.
5	**AOB and Close:** Jenny Brooker, Chair
	● **Additional Updates by Jenny Brooker:** ○ Jenny provided a brief update on the broader data sharing strategy and how all the work fits into the overarching data foundations being developed by CDDO under the Department for Science, Innovation, and Technology (DSIT). ● **Conclusion and AOB:** ○ Jenny thanked everyone for their attendance and contributions to the discussion. ○ Jenny reiterated the need for continued collaboration and active participation from all departments to ensure the success of all with work presented in PRG ○ The meeting concluded with a reminder that meeting notes and presentations will be shared with participants.