Options Summary

Information Asset Register for GDPR

Version / Date	**0.1 22/6/18**
Author	**Stephen Crowther**
Owner	**Peter Grogan**

1. Introduction

As part of the activity required for SCC to achieve compliance with the new GDPR regulations, the authority will need to create and maintain an Information Asset Register (IAR) listing those databases within the organisation that hold personal data.

There are two options under consideration:

a. Assyst

b. Azure Data Catalogue (ADC)

2. Options for an IAR

Option	Costs	Benefits
a. Assyst	None for the application itself. Resource will be required to populate the IAR manually with the information relating to applications and the data they hold and to maintain register entries on an ongoing basis.	Utilises existing application already in use within the authority. IAR fields have already been set up by the Service Desk, although these have not yet been populated.
b. Azure Data Catalogue	Part of the package of MS products already purchased by SCC, so no initial cost for the application itself. Will require some manual entry to start with, prior to automated discovery being activated. If more than 5000 data objects being monitored - cost of 0.75p per ADC user per month. Maximum anticipated users = 30, so max. cost per year of £270. It is possible to restrict usage to a smaller group if required, thereby reducing cost.	Offers automated discovery / database interrogation capability that Assyst lacks. Can provide reporting on database size / number of records etc. to support with GDPR auditing activity. Has uses beyond providing IAR that can be exploited by ICT going forward.

The Service Desk has set Assyst up with the necessary fields for an IAR, however all entries would need to be input and maintained manually on an ongoing basis for the register to be kept up to date. There is no resource currently identified to cover this commitment and it carries with it the risk that over time the register may become outdated and inaccurate. In addition, using Assyst would rely on the collective knowledge of SCC to identify databases, with the risk that some may be overlooked and omitted from the IAR. Information such as the size of a database and the number of records it holds will not be available via Assyst, necessitating manual enquiries to database administrators to access this information.

ADC will require less manual intervention than Assyst due to its automated discovery functionality and it removes the reliance on collective knowledge to identify databases within the organisation. As a result, ADC is more likely to produce an accurate register of databases, while at the same time requiring less support cost than Assyst would. ADC supports metadata extraction from a range of data sources and information on database size and number of records is readily available, speeding up responses to auditing queries. After a data source has been registered by ADC, users can annotate the record within ADC, adding descriptive metadata to the structural metadata that is registered from the data source during discovery. This means that over time a richer view can be created for those data sources that justify the effort, while a more basic record is maintained for all others.

4. Further information on ADC

ADC_Overview.pptx

ADC links:

What is Azure Data Catalogue

Supported Data Sources (PostgreSQL is supported for automatic discovery)

ADC Pricing

Solving GDPR Discovery via Azure Data Catalog (1/3)

What exactly is a GDPR Taxonomy and how can Azure Data Catalog help? (2/3)

GDPR – Implementing a Taxonomy in Azure Data Catalog Step by Step (3/3)

5. Recommendation

Subject to budget approval, the recommended option is to choose ADC as the platform for creation of the IAR. This will require some initial manual set-up to create a draft IAR covering the main applications we know to contain personal data.

Automated discovery can then be activated to further populate the register, and we can annotate metadata for individual applications as appropriate thereafter. Further uses for ADC can be developed on a case-by-case basis to exploit the capabilities of the tool.