

Capgemini's data masking architecture, organization and process



A proven approach lets financial institutions use sensitive production data flexibly yet with no risk to security or privacy

Organizations often need to use production data safely outside of production processes: for example to provide realistic test data, or for support purposes. Data masking enables production data to be used securely and in compliance with privacy requirements. A variety of masking techniques are available for different situations.

With extensive experience of building and managing enterprise-wide data masking solutions for global banks, Capgemini offers a robust metadata-based architectural approach, together with a process and organization that rapidly discovers what data needs to be masked and then implements and runs the masking process. This work can be carried out in a federated model or by a single center of excellence (CoE) for the entire enterprise and its ecosystem.

Why data masking is needed

Organizations need to use production data to:

- **Facilitate testing:** Adapting production data for testing purposes provides realistic, high-quality data, minimizing the chances of defects getting through testing and enabling seamless testing across all applications.
- **Facilitate operational and production management:** Access to the production database may be needed for technical support, data input, or customer service.

At the same time, organizations must safeguard the security and privacy of sensitive data, complying with information security guidelines and regulations. More than 50

international privacy laws require financial firms to provide granular auditing (these include the US's Gramm-Leach-Bliley and Sarbanes-Oxley Acts, and the UK's Data Protection Act). In addition, many countries have "sovereign soil" requirements for clients' personal data to be processed, accessed and stored locally.

If organizations fail to comply, they lay themselves open to regulatory breaches and audit violations, which can be expensive both in terms of immediate financial impact (fines can reach half a million dollars per incident) and in terms of brand damage and loss of customer loyalty.

An important part of the solution is data masking, which transforms or selectively replaces production data so that it can be used without jeopardizing security or compliance.

How data masking works

Data masking must be implemented robustly, following certain principles. For example, when data is masked for testing purposes, the process must not be reversible; it must be repeatable and, as far as possible, automated. The masked data must be representative of the source data, and retain its referential integrity.

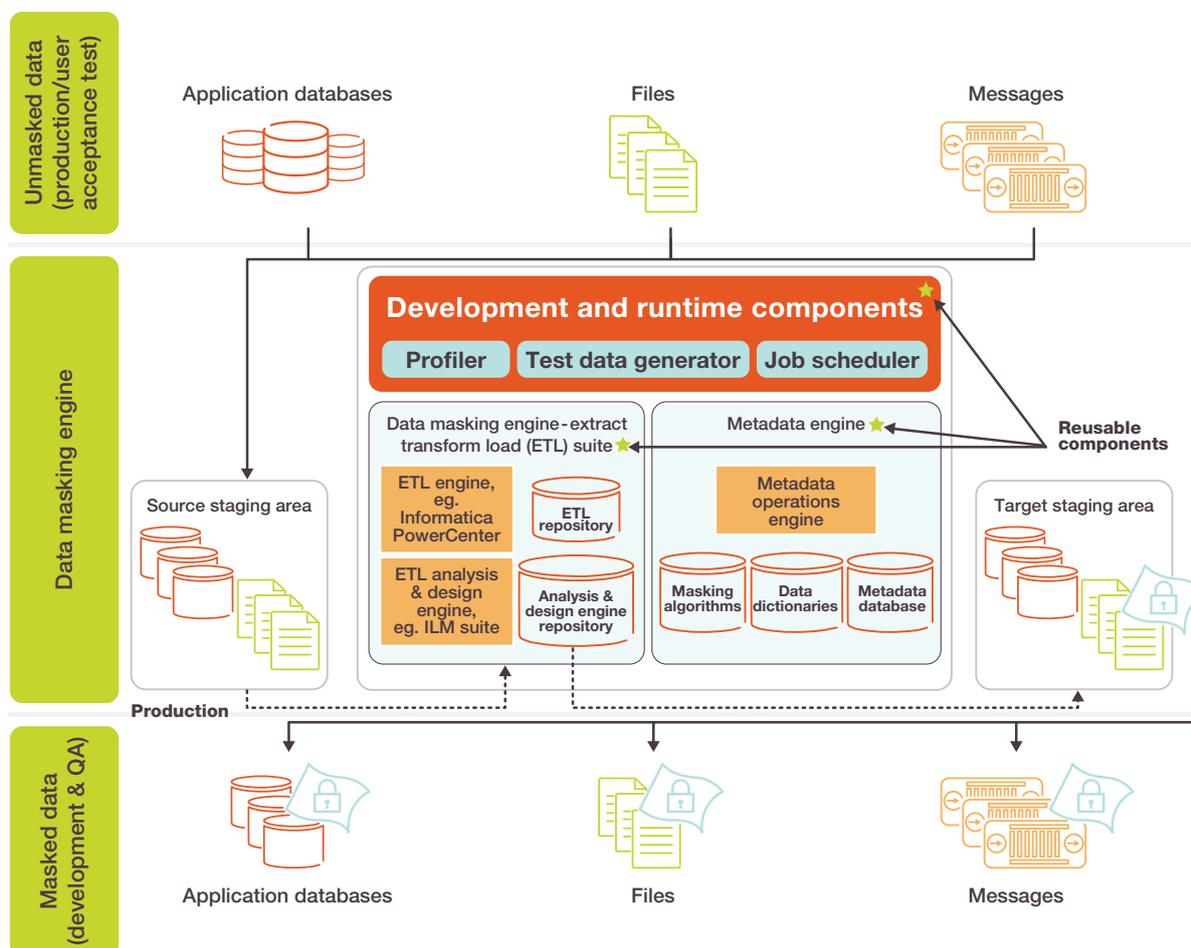
Three key techniques are used:

Static data masking desensitizes data in non-production environments such as testing. Data is desensitized while retaining its "look and feel".

Dynamic data masking is used in production environments to provide selective access to live data. People see only the data fields that they are authorized to see; the rest are scrambled or hidden. Masking is performed in real time, either by intercepting the back-end database or on the front-end user interface (UI). Dynamic data masking at UI level can be combined with tokenization (described below) for compliance with sovereign soil requirements.

Tokenization is a highly secure way to protect data in the production environment. At the data entry point, real production values are exchanged for tokens that are then used in production systems, eliminating the risk of data misuse. The tokens are exchanged back to provide real values for the systems that are authorized to access them. Tokenization is often implemented with encryption for advanced protection.

Figure 1: Data masking architecture



Capgemini's approach

Capgemini advocates a centralized approach that standardizes data masking across the client's ecosystem. Our unique **metadata-focused architectural framework** (figure 1) facilitates management and reuse, and ensures scalability

This tool-agnostic framework adapts to each client's environment and technology standards. It is designed to be flexible so that, for example, masking can take place in memory or through the use of the staging area. It supports a wide range of relational databases, files and real-time messaging in both mainframe and distributed environments.

Masking is not just about technology, and so our proven **process and organization** provides a structured approach to discovery and fulfillment of masking needs. We help clients embed strong governance to maintain consistent business rules for masking and to establish policies and processes to ensure consistency across the enterprise.

Our approach covers all the requirements of a masking program including introduction of proven methods of discovery and demand management, on-boarding of applications, and execution of the masking process.

Let's take a closer look at two of the strengths of our approach: discovery and the Center of Excellence (CoE).

Discovery

A Capgemini tool automatically identifies and reports potentially sensitive elements within application data by examining metadata. Risk scoring, coupled with questionnaires, determines the degree of sensitivity of each type of data, and each data lifecycle. This information can be used to prioritize data for masking and hence shape the on-boarding plan.

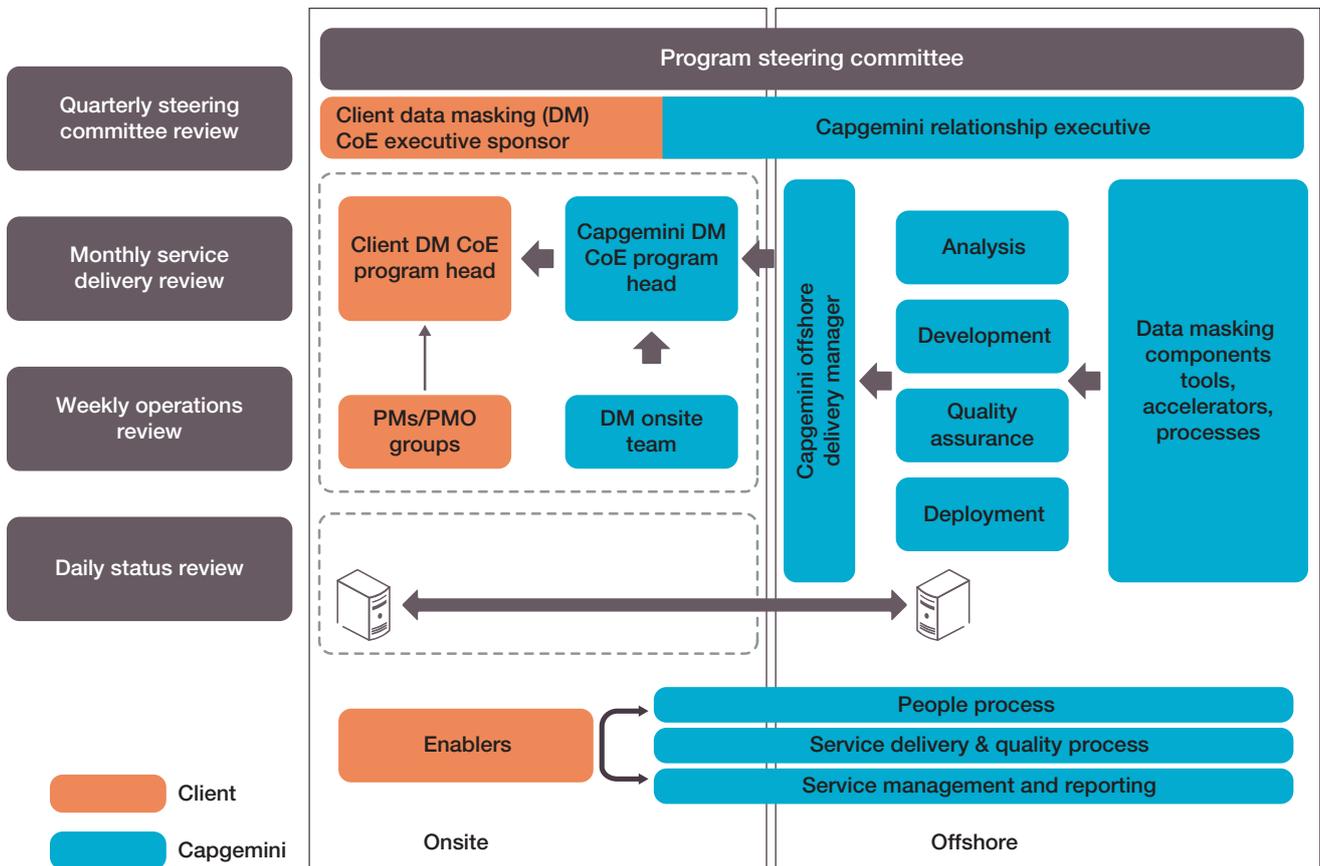
Center of Excellence

Masking can be carried out either locally in a federated model, or else centrally – the latter approach can help to achieve a consistent, efficient and reliable masking process. We can help set up a CoE with qualified teams to facilitate:

- Adoption of accelerators and best practices
- Phased implementation plan for all applications across all business units
- Setup of a common infrastructure and test environment

A typical governance structure for a data masking CoE is shown in figure 2.

Figure 2: Program and governance structure



Why Capgemini?

Our approach delivers:

- Watertight data security to avoid reputational or financial losses due to data breaches
- Flexible use of production-derived data
- Safe outsourcing of testing activities (using masked production data) if required
- Improved efficiency and effectiveness of testing and quality efforts
- A single set of test data across applications to enable seamless testing

We achieve all this cost-effectively and smoothly because of our unique metadata-driven architecture and proven approach to organization and process – all delivered by experienced managers and technicians.

Success story: masking as part of a test data management solution for a global financial services company

Partly for compliance reasons, this client needs to be confident that identifiable personal data is completely protected from unauthorized use. However, it also needs to use production-derived data in non-production environments to enable efficient and thorough testing.

We helped set up a Center of Excellence to provide repeatable, cost-effective test data management as a shared service for the entire company. A data masking capability was central, and we created it in three steps:

- **Scope definition:** analysis of regulatory and business drivers to determine what to mask, why, how and when.
- **Discovery:** identification and cataloging of sensitive information across all applications in scope, using our data discovery tool.
- **Definition and deployment:** definition of the architecture and processes for masking, followed by development and implementation of the masking solutions.

This masking capability is now helping the client to protect sensitive data and also meet stringent targets for test data management. These include a refresh time for the test data environment of less than two days, a 10% reduction in the duration of test cycles, and a data quality index of over 0.95.



About Capgemini

With almost 140,000 people in over 40 countries, Capgemini is one of the world's foremost providers of consulting, technology and outsourcing services. The Group reported 2013 global revenues of EUR 10.1 billion. Together with its clients, Capgemini creates and delivers business and technology solutions that fit their needs and drive the results they want. A deeply multicultural organization, Capgemini has developed its own way of working, the Collaborative Business Experience™, and draws on Rightshore®, its worldwide delivery model.

For further information visit www.capgemini.com/bim

or contact us at bim@capgemini.com