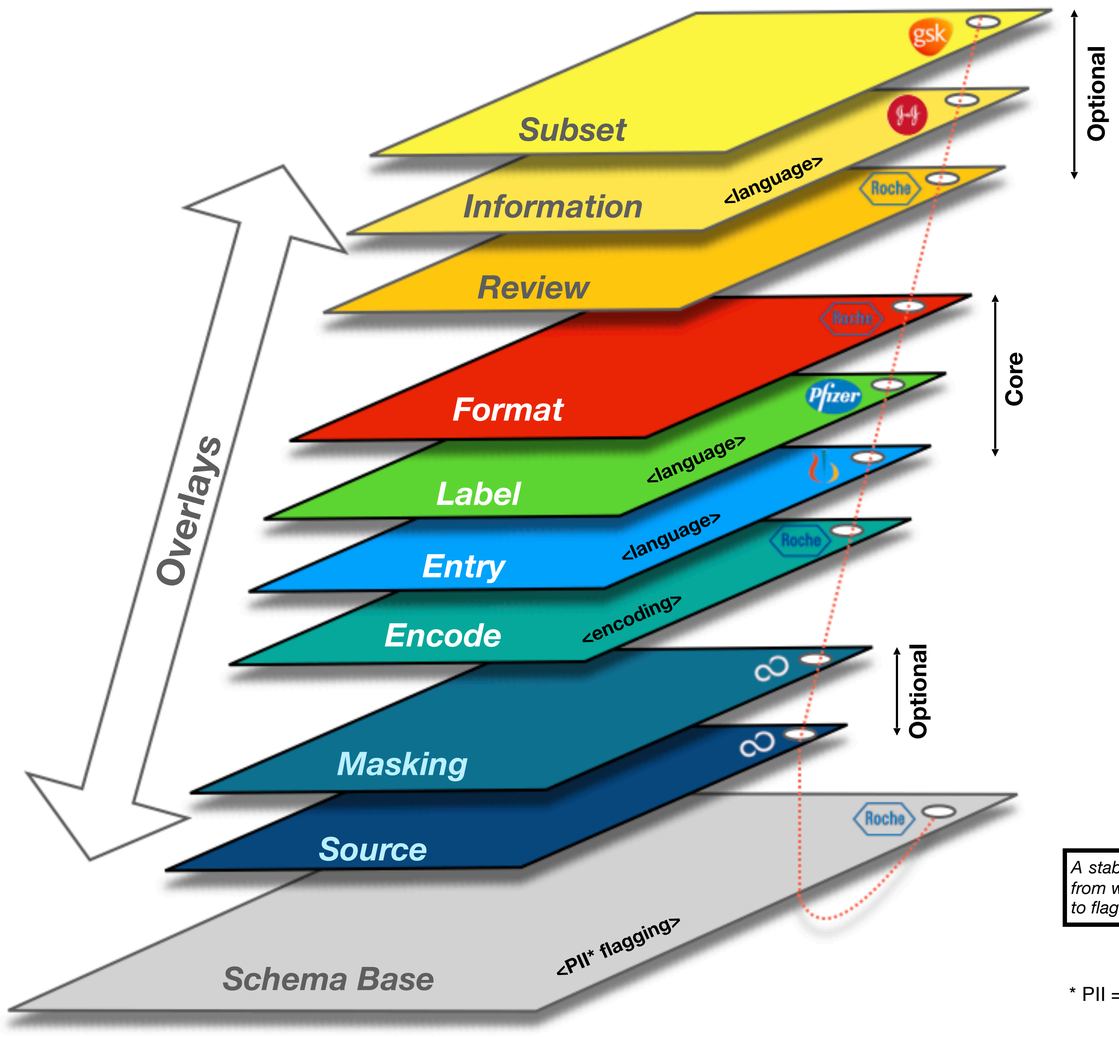
Overlays Capture Architecture (OCA)

Providing a standardised global solution for data capture and exchange

Paul Knowles, Human Colossus Foundation / paul.knowles@humancolossus.org / www.humancolossus.foundation



*A **Subset Overlay** is used to create a schema subset.*

*An **Information Overlay** is used to add instructional, informational or legal prose to assist the data entry process.*

*A **Review Overlay** is used to enable third-party comments to initial field entries (e.g. for auditors, reviewers, compliance officers, etc.).*

*A **Format Overlay** is used to add formats to schema attributes (incl. field lengths).*

*A **Label Overlay** is used to add labels to schema attributes (incl. category labels).*

*An **Entry Overlay** is used to add predefined field values to schema attributes.*

*An **Encode Overlay** is used to define character set encoding (e.g. UTF-8, ISO-8859-1, Windows-1251, Base58Check, etc.).*

*A **Source Overlay** is used to specify endpoints where dynamic data or linked attachments can be located.*

*A **Masking Overlay** is used to add data masks to schema attributes that have been flagged as PII* in the Schema Base.*

*A stable base object that defines a single set of data in its purest form thus providing a standard base from which to decentralise data. A **Schema Base** facilitates a PII* schema object which allows the issuer to flag personally identifiable information.*

* PII = personally identifiable information

Why is OCA useful?

Data pooling. Decoupling can occur at any time as overlays are linked objects. With all coloration stored in the overlays, combining data from related sources becomes much easier. Overlays can be removed from the base objects before the data merging process begins and reapplied to ensure consistent coloration post data pooling.

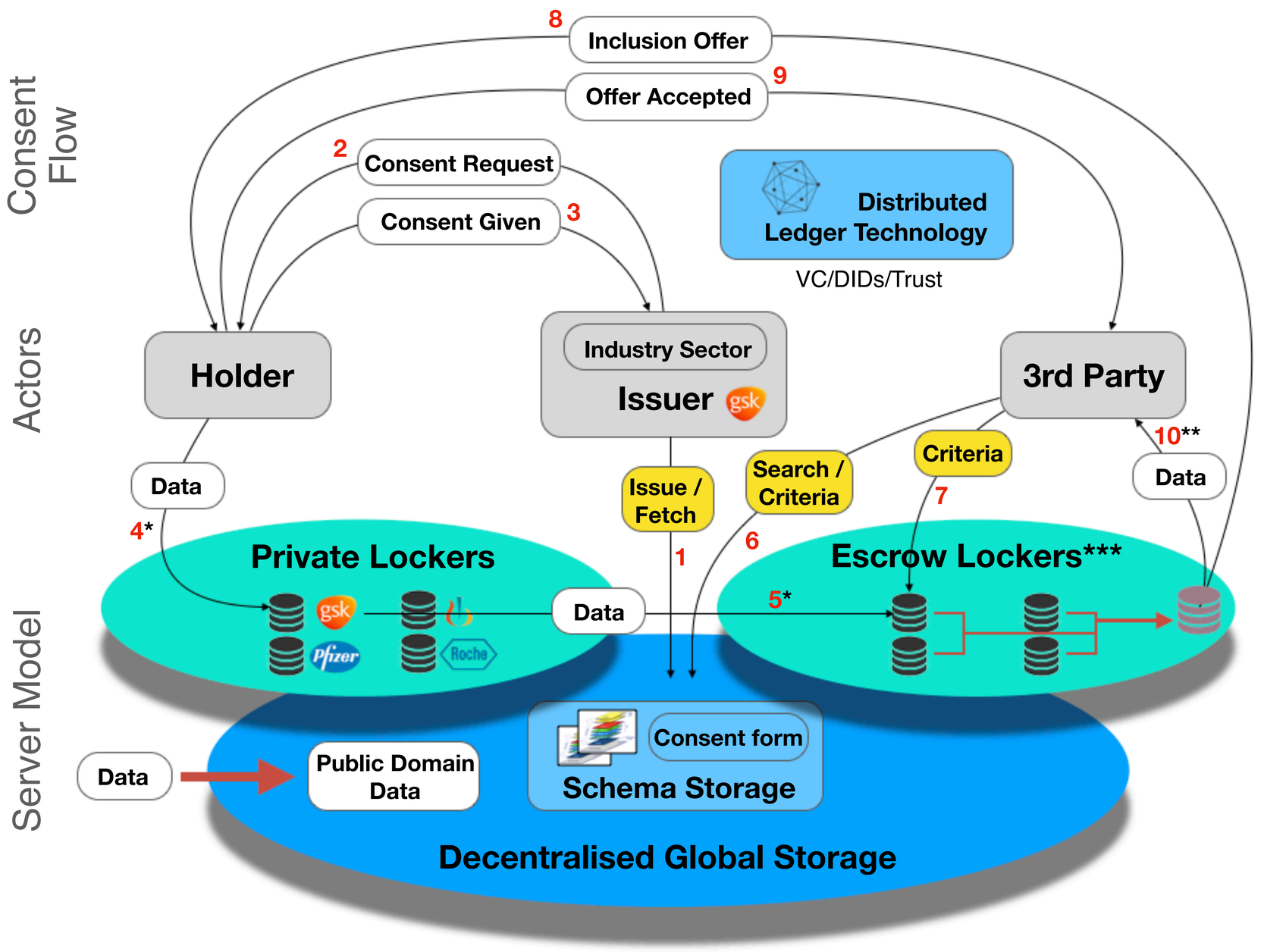
Flagged attributes for encryption. Using the Blinding Identity Taxonomy (BIT) as a reference, issuers can flag attributes in the schema base that could potentially unblind the identity of a governing entity. With attributes flagged at the base object layer, all corresponding data can be treated as sensitive throughout the data lifecycle and encrypted or removed at any stage making associated entity identification impossible.

Internationalisation. As character set encoding definitions are captured in a separate linked data object, a single report definition can contain different attribute forms for different languages available to users, based on a user's locale and other language preferences.

Stable schema bases. Most schema updates tend to be done at the application stage. In the case of OCA, all extension, coloration, and functionality definitions are applied in the overlays. This enables issuers to edit one or more of the linked objects to create simple updates rather than having to reissue schema bases on an ongoing basis.

Data decentralisation. Schema base definitions can remain in their purest form thus providing a standard base from which to decentralise data. Once the data holder has given adequate consent, data controllers can contribute anonymous data to decentralised data lakes upon which 3rd parties can trigger accurate criteria searches for matched data. This eliminates the need for data silos and encourages consented data sharing. The data holder is empowered by self-determination regarding secondary use of their personal data.

OCA facilitates a unified data language



* Handling of PII data is specified in the initial *Consent/Access Request*

** All PII data is anonymised prior to leaving an escrow locker

*** Escrow lockers can be broker services which handle data on behalf of the holder as a personal agent

HUMAN COLOSSUS FOUNDATION