# Data sharing

## What is it.

The data consumer that successfully negotiated a data sharing agreement with a data provider that a data transfer can/shall take place; the request is managed and controlled by the connectors of both parties, it might be interactive, or batch driven. If an issue prevents the data consumer to perform the request, there should be a way to contact the data provider for support.

Usage control restrictions apply before and during the actual data transfer, any counter or transfer metric is updated and shared with both parties. In some cases there might be usage restrictions that apply at the destination site after the actual data transfer.

The transfer and its outcome are (optionally) recorded in the EMDS observability service, the data consumer might also rate the experience and quality of service with the EMDS community (via the EMDS portal).

## Detailed overview

### Graphical representation

A diagram of a company

Description automatically generated

### Actors

|  |  |  |
| --- | --- | --- |
| Actor | Description | Notes |
| Data consumer | A participant engaged in data sharing in the International Data Spaces requesting and using data provided by a Data Provider. | Participants can assume different roles (e.g., consumer, producer, intermediary, etc.) |
| Data provider | Participant exposing Data Sources via a Connector; a Data Provider may be an enterprise or other organisation, a data marketplace, an individual, or a “smart thing”. | The DS Authority might release credentials or delegate the release to a third party. |
| Observer | A participant that implements a business observability function across the entire data space. There might be more than one observer in a data space. | Business observability is the pillar for governance enforcement. |

### Customer journey: Data sharing request

|  |  |  |  |
| --- | --- | --- | --- |
| Sub-customer journeys | Description | Examples of ISO 25010 Qualities | Review/tests |
| Request data transfer | Consumer requests the initiation of data transfer. The interface can be an API or a U/X. The data transfer might involve intermediate steps, like provisioning an endpoint, generating transfer credentials, or analyse a query. The mechanism is asynchronous; therefore, a reliable channel of communication is needed to make sure that both parties maintain control of the process. | 1. Functional completeness: The system provides a documented, programmatic interface (API) to initiate and manage a data transfer. 2. Functional Suitability - Functional completeness: the system supports multiple protocols and the ability to perform queries against the data products endpoint. 3. Reliability - Faultlessness: The system supports asynchronous or long-running data sharing executions, with variable throughput and response times. 4. Security - Integrity: the system adequately secures the data sharing request between the consumer and its connector. 5. Reliability - Recoverability: the system provides a recovery mechanism of the transferring session. 6. Performance efficiency - Scalability: The system can scale to support increasing numbers of active long-running requests. | **1: Coverage test: assess that the API is available and test that a data sharing request is properly covered: Initiate a data sharing Retrieve data sharing information and status Receive data sharing request outcome condition**  **Retrieve data sharing information of past data sharing actions.**  **The system ranks higher if the API is secured and implements common methods, like REST.**  **2: Coverage test: for each data plane available, test a minimal data sharing and identify possible inconsistencies with the original data product protocol endpoint (e.g., during data querying, we aren’t allowed to send http headers to the data source).**  3: Part of smoke test 5: implement a slow consumer.  4: Assessment: assess that the data sharing request between the consumer and the consumer’s connector is secured (see similar tests involving AAA and secure channel interfaces).  5: Smoke test: initiate a slow consumer data sharing and break the connection. Assess the state of both the producer and consumer systems and how they manage and log the fault.  6: Assessment or smoke test: assess two use cases: 1) many clients performing data sharing requests on the same data product. 2) (Optional) many clients performing many sharing requests on the same provider connector. If the system requires autoscaling through external load balancing, test how sessions are supported and rank reliability. If the system supports only vertical scaling, rank if it could undergo disruptions during scaling. |
| Request support | Upon a reported problem, the data producer is notified that an issue must be solved. The data producer should have access to an issue management system, perhaps hosted on the EMDS portal | 1. Functional suitability - Functional completeness: The system (connector) provides a support request channel via its API? 2. Security - Non-repudiation: The system (connector) provides a mechanism to return a read receipt for support requests. | 1: Assessment: rank higher if support and feedback communication are part of the connector protocol.  2: Assessment if the feature is supported. |
|  |  |  |  |

### Customer journey: Sharing (data)

|  |  |  |  |
| --- | --- | --- | --- |
| Sub-customer journeys | Description | Examples of ISO 25010 Qualities | Review/tests |
| Enforce usage control | During the transfer, product-associated usage policies are enforced. Given the broad scope of such policies, the result of the sharing transaction might vary a lot. A policy could, for instance, filter data because the consumer lacks an attestation. The consumer needs to be informed of policies’ actions and if they altered the data she receives. | 1. Functional Suitability – functional completeness: the system provides a PEP function that supports usage control policies of classes: Allow-usage, Role-restricted, Connector-restricted, Purpose-restricted, Interval-restricted, Location-restricted, Log-usage-of-data. (see IDSA Position Paper on Usage Control). 2. Maintainability - Analysability: The system can be configured to persist the outcome of policy evaluation in a log, or to an observability tracing system. 3. Security - Integrity: The system provides usage control policies that verify the integrity of data (avoid data tampering). | **1: Testing: test the policies that are supported out of the box.  For the policies that are not supported, describe the effort of how to build them, and rank the system consequently (e.g.: create a plugin in a documented environment ranks better than integrating an external function that introduces dependencies and interface maintenance).**  **2: As above, but for tracing.**  3: Do not perform tests, yet. |
| Update subscription / counters / billing | When the data product is bound to a rating on consumption or any other value, the counters are updated. This information is available to the consumer so that she can modulate her requests within her budget. | 1. Functional suitability – functional completeness: The system supports counters (for a subscription mechanism) during a data transfer. 2. Compatibility - Co-existence: The system can share counters via an API to a rating engine, or incorporates such feature. | 1: Assessment: provide an analysis of counters, their coverage of measurement, and the accessibility of counters during a data transfer and afterwards.  2: Assessment: provide an analysis of rating presence out of the box or what an integration could require. |
| Incident management | Similar to “Request Support” customer journey but performed in case a security issue or data leak are detected. For instance, if the consumer is aware that her system has been compromised and the shared data has been exposed. | 1. Functional suitability - Functional completeness: The system (connector) provides an incident request channel via its API, or reuses the Request Support API. 2. Security - Non-repudiation: The system (connector) provides a mechanism to return a read receipt for incident requests. 3. Security – Confidentiality: The system handles the incident requests in a secure and private way (secure channel, only access to “need-to-know” operators). | Do not perform tests yet. |

### Customer journey: Post-sharing activities

|  |  |  |  |
| --- | --- | --- | --- |
| Sub-customer journeys | Description | Examples of ISO 25010 Qualities | Review/tests |
| Enforce post-sharing usage control policies | If usage policies of a data product define controls that go beyond the data sharing process, the data consumer is responsible to enact them and document their outcome in the observability service. | 1. Functional Suitability – functional completeness: the system provides a feedback loop that feeds a PXP function, where the consumer can send the result of the following policy classes: Purpose-restricted, Log-usage-of-data, Use-the-data-and delete-after. (see IDSA Position Paper on Usage Control). 2. Maintainability - Analysability: The system can be configured to persist the outcome of policy evaluation in a log, or to an observability tracing system. 3. Functional suitability: UC5-UC7, UC9 | 1: Testing: test the mechanism that is necessary to support post-sharing and client policies. If out of the box ranks higher.  If the mechanism is not supported, describe the effort it would require developing such function, and rank the system consequently (e.g.: create a plugin in a documented environment ranks better than integrating an external function that introduces dependencies and interface maintenance). |
| Share experience / results with community | Consumer (and producers) share data sharing experiences, advice, qualitative information over the data products they receive (transfer) and their integration. | 1. No tests required. |  |
| Log data sharing transaction and usage control assurance in the observability service | Every transaction is logged and is available for governance and support activities. | 1. Security - Confidentiality: The system provides each participant with the option to share logs and observability traces to each other or to authorities for governance and support activities. 2. Functional suitability:    1. OA1 Local logs    2. OA2 Data Space logs collector    3. OA3 Technical service logs    4. OA4 Auditable and Reviewable services    5. UC9 3. Compatibility - Interoperability: The system uses a common format for all business and technical logs. 4. Security - confidentiality: The system secures shipping and access to dataspace Log collection. 5. Func. Suit. – Functional completeness: The system allows a provider/consumer to query the logs for inconsistencies, like failed data exchange, missed SLAs, violated policies. 6. Compatibility - Interoperability: Traces and logs can be exported in a machine-analysable format, after the agent exporting them has been granted the right to do so. 7. Maintainability - Analysability.: The system provides a persistence layer for logs and traces, so they can be used for back-in-time analysis. 8. Integrity: The system persist logs and traces on an immutable storage. | **1: Assessment: analyse the mechanism that allow to switch off and on logs during a data sharing request. If a mechanism requires intervention of a connector administrator, it ranks less than an API that a user can call.**  2: No tests required.  **3: Assessment: assess if the formats can be parsed by commonly used tracing programs (see incumbent formats).**  4,5,6,7.8: Assessment: as whole these qualities can be part of a “suitability for observability” assessment test. Presence or not of these features and their maturity can be reported in a gap matrix. |
|  |  |  |  |