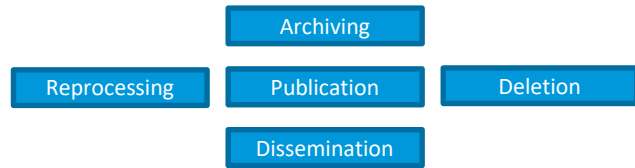
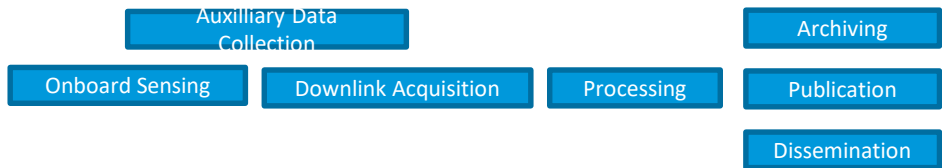


Data Flow Operations Concepts – Product Lifecycle Evolution

December 2018

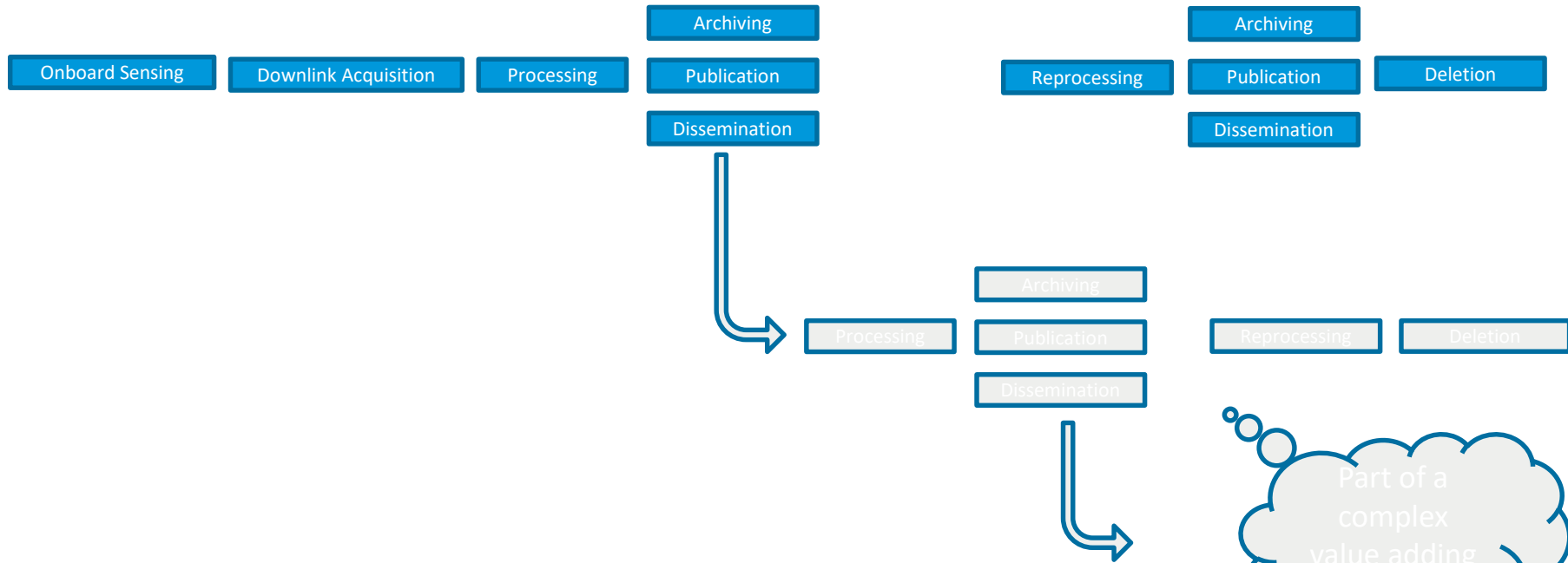
EO Product Lifecycle – Systematic Production



A Simplistic Data Provider View of Systematic Production



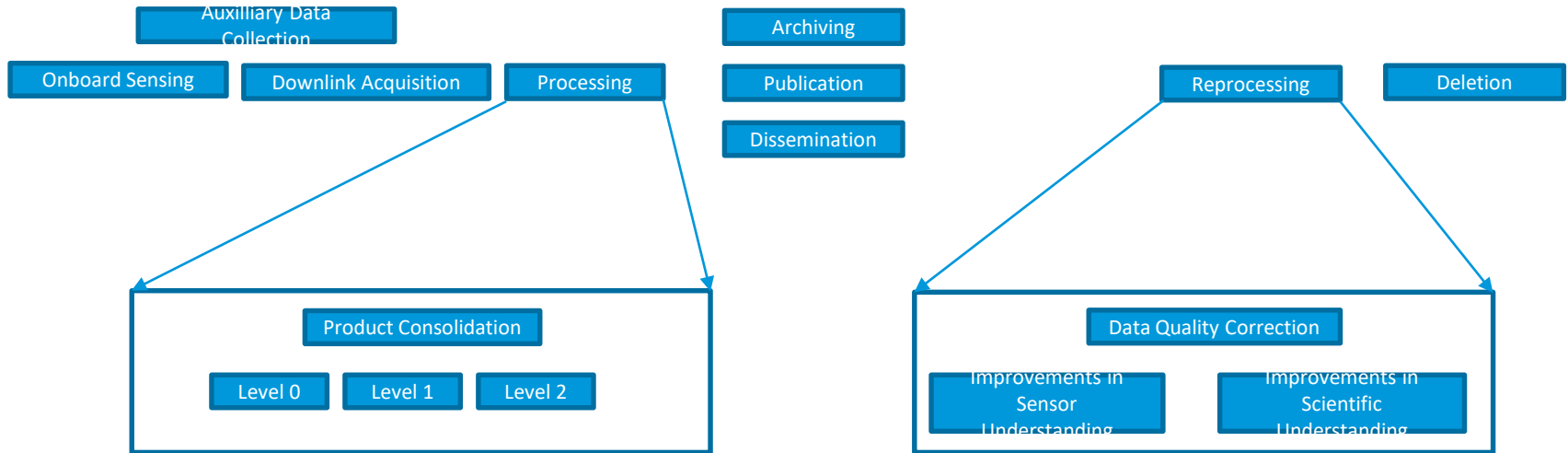
EO Product Lifecycle – Systematic Production



Part of a complex value adding chain



EO Product Lifecycle - Processing



Result of many trade off analysis and experiences
Cost / Quality / Schedule



EO Product Lifecycle – Current Production Baselines



Relying on the correct and reliable configuration of complete processing chain

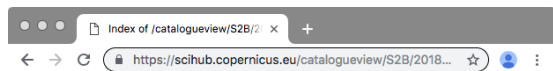
EO Product Lifecycle – Current Implementation of Catalogue Completeness - APIs and lists



Products() -> uuid, md5, id/filename ... metadata

DeletedProducts() -> uuid, md5, id/filename ... deletionCause

CatalogueView: .csv lists



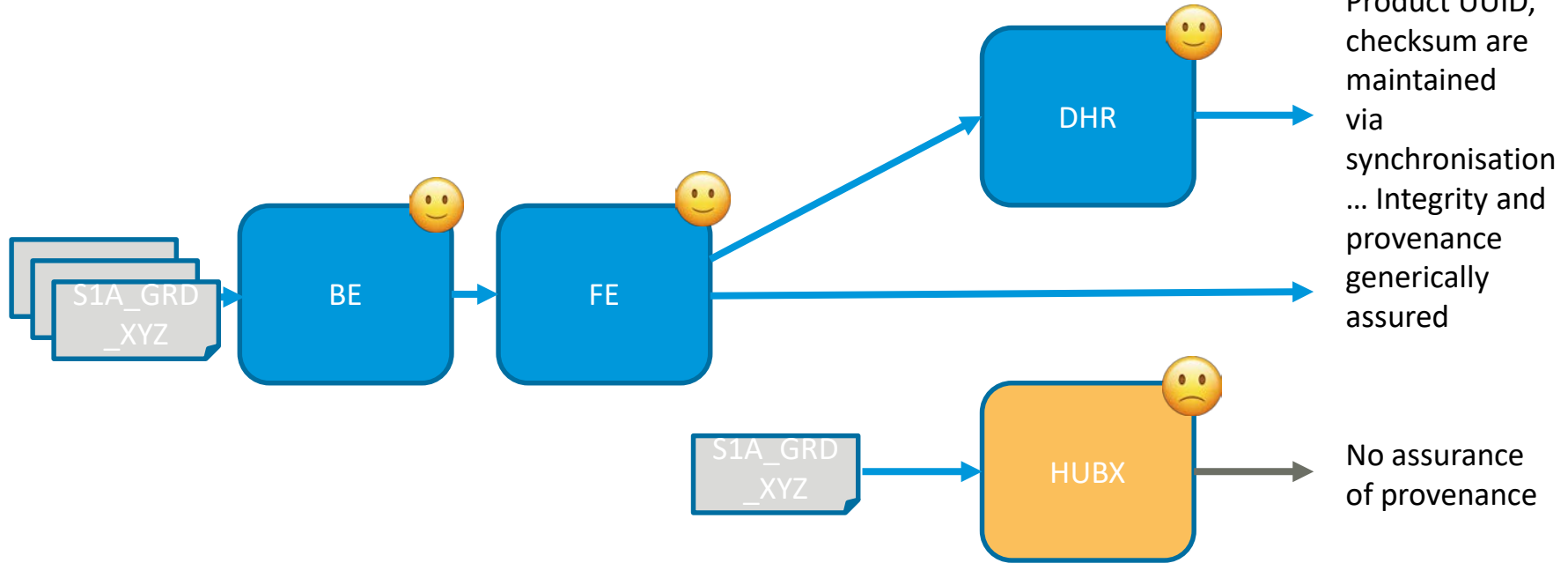
Index of /catalogueview/S2B/2018/09

| Name | Last modified | Size |
|---|------------------|------|
| Parent Directory | - | - |
| S2B_20180901_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180902_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180903_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180904_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180905_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180906_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180907_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180908_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180909_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180910_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180911_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |
| S2B_20180912_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.3M |
| S2B_20180913_OPENHUB_catalogue_20180930235959.csv | 2018-10-02 14:31 | 1.4M |

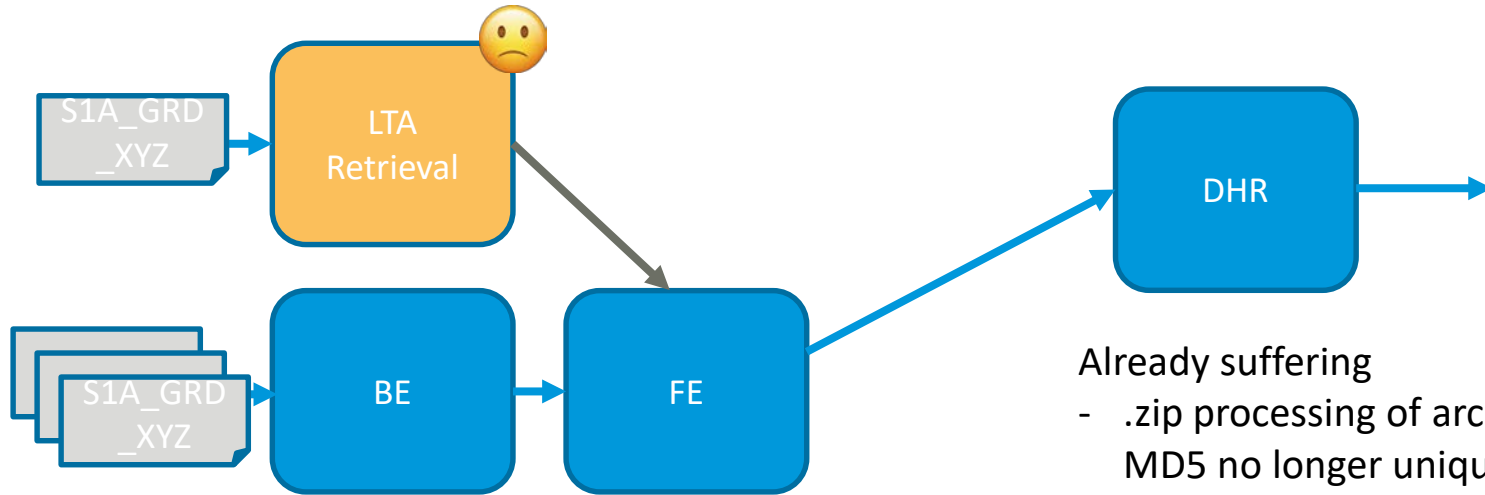
| id | Name | ContentLength | IngestionDate | ContentDateStart | ContentDateEnd | Checksum | Algorithm |
|---------------------------------------|---|---------------|-------------------------|-------------------------|-------------------------|----------|-----------|
| 0007994a-807e-4d75-beb3-80d8f548557d | S2B_MSILC_20180901102519_ND00_ND_180T1_201809011070048 | 810779453 | 2018-09-01T11:38:49.885 | 2018-09-01T08:25:29.024 | 2018-09-01T08:25:29.024 | MDS | |
| 0007007a-48aa-4c0e-97b3-0e931619304 | S2B_MSILC_20180901114639_ND00_ND_180T1_20180901130611 | 124627787 | 2018-09-02T05:37:57.664 | 2018-09-02T11:46:39.024 | 2018-09-02T11:46:39.024 | MDS | |
| 0010087f-43ab-4e38-99c0-c8661107652 | S2B_MSILC_201809011014639_ND00_ND_180T1_20180901108205 | 596674673 | 2018-09-01T07:54:42.017 | 2018-09-01T01:46:39.024 | 2018-09-01T01:46:39.024 | MDS | |
| 000876f9-6449-4b41-8099-908a1e0e077d | S2B_MSILC_20180901151149_ND00_ND_180T1_20180901120924 | 554466283 | 2018-09-02T08:55:45.532 | 2018-09-02T11:46:39.024 | 2018-09-02T11:46:39.024 | MDS | |
| 001870b1-13d0-47ba-8184-a0980a084e0 | S2B_MSILC_20180901100049_ND00_ND_180T1_20180901106000 | 624448840 | 2018-09-01T08:52:36.461 | 2018-09-01T02:04:39.024 | 2018-09-01T02:04:39.024 | MDS | |
| 001e4e40-7163-404e-a3e7-8aa8a5a277f3 | S2B_MSILC_20180901115348_ND00_ND_180T1_20180901130242 | 699400039 | 2018-09-02T05:46:03.141 | 2018-09-02T11:53:49.024 | 2018-09-02T11:53:49.024 | MDS | |
| 005e9e3c-022b-461b-8c5c-7eaa8884e7e4 | S2B_MSILC_20180901122529_ND00_ND_180T1_20180901101178 | 543688116 | 2018-09-02T12:17:13.588 | 2018-09-02T22:25:29.024 | 2018-09-02T22:25:29.024 | MDS | |
| 000e9fde-8e72-4d5c-9d7f-1c680a0b37b | S2B_MSILC_20180901100049_ND00_ND_180T1_20180901106000 | 806748406 | 2018-09-01T08:52:36.461 | 2018-09-01T02:04:39.024 | 2018-09-01T02:04:39.024 | MDS | |
| 0010087f-43ab-4e38-99c0-c8661107652 | S2B_MSILC_20180901102519_ND00_ND_180T1_201809011070048 | 710832002 | 2018-09-01T08:46:56.957 | 2018-09-01T08:25:29.024 | 2018-09-01T08:25:29.024 | MDS | |
| 0006923a-0247-4919-430f-9f9c56680904 | S2B_MSILC_201809011014639_ND00_ND_180T1_20180901108205 | 290987931 | 2018-09-01T07:04:29.046 | 2018-09-01T01:46:39.024 | 2018-09-01T01:46:39.024 | MDS | |
| 101707eae-9445-4a42-1e1b-4e6707782626 | S2B_MSILC_20180901100049_ND00_ND_180T1_20180901106000 | 423140208 | 2018-09-01T07:47:29.243 | 2018-09-01T05:40:49.024 | 2018-09-01T05:40:49.024 | MDS | |
| 000876f9-6449-4b41-8099-908a1e0e077d | S2B_MSILC_20180901100049_ND00_ND_180T1_20180901106000 | 602734170 | 2018-09-01T08:40:39.822 | 2018-09-01T00:05:29.024 | 2018-09-01T00:05:29.024 | MDS | |
| 000c9e46-4365-40e4-e7ef-13178e0e7040 | S2B_MSILC_20180901123249_ND00_ND_180T1_201809011023740 | 731189537 | 2018-09-02T11:34:42.304 | 2018-09-02T22:25:29.024 | 2018-09-02T22:25:29.024 | MDS | |
| 10082112-1e7e-4e17-88e1-37a09099 | S2B_MSILC_20180901102519_ND00_ND_180T1_201809011070048 | 636369533 | 2018-09-02T11:44:50.059 | 2018-09-02T18:25:29.024 | 2018-09-02T18:25:29.024 | MDS | |
| 000c6e5b-4846-4041-99ef-23a70841838 | S2B_MSILC_20180901101019_ND00_ND_180T1_201809011161700 | 755900554 | 2018-09-02T08:59:29.240 | 2018-09-02T10:10:10.024 | 2018-09-02T10:10:10.024 | MDS | |
| 00092172-76a2-4b11-8e00-386a0a08868 | S2B_MSILC_20180901182909_ND00_ND_180T1_20180901154956 | 803281643 | 2018-09-02T18:42:26.500 | 2018-09-02T18:29:09.024 | 2018-09-02T18:29:09.024 | MDS | |
| 100-424203-344e-4d7f-af6a-7f61e158420 | S2B_MSILC_20180901182909_ND00_ND_180T1_20180901154956 | 352700764 | 2018-09-02T18:56:08.004 | 2018-09-02T18:29:09.024 | 2018-09-02T18:29:09.024 | MDS | |
| 000b3344-43ee-430e-4a0e-5c76a76405d0 | S2B_MSILC_20180901100049_ND00_ND_180T1_20180901106000 | 344860993 | 2018-09-01T08:08:31.257 | 2018-09-01T02:04:39.024 | 2018-09-01T02:04:39.024 | MDS | |
| 000f9e9c-9394-438b-8884-43a21186c1c3 | S2B_MSILC_2018090110701159_ND00_ND_180T1_20180901103909 | 325493707 | 2018-09-02T18:54:15.375 | 2018-09-02T17:15:59.024 | 2018-09-02T17:15:59.024 | MDS | |
| 0010087f-43ab-4e38-99c0-c8661107652 | S2B_MSILC_20180901102519_ND00_ND_180T1_201809011070048 | 715468046 | 2018-09-02T12:22:15.738 | 2018-09-02T22:25:29.024 | 2018-09-02T22:25:29.024 | MDS | |



EO Product Lifecycle – Current Implementation - Ingestion vs. Synchronisation in the DHR Network



EO Product Lifecycle – Current Implementation – Already Suffering



Already suffering

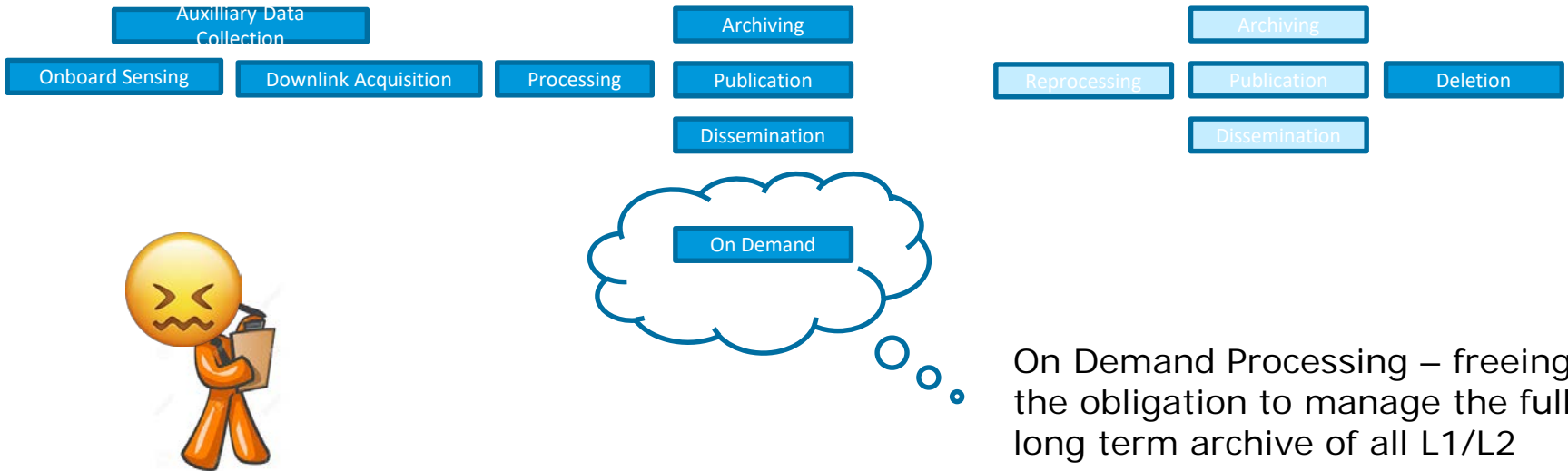
- .zip processing of archive data differs, MD5 no longer unique

Possibilities are “workarounds”

- Alternative MD5’s for one UUID
- UUID generation based on core components of the product

As many drawbacks as benefits

EO Product Lifecycle – Evolution to On Demand Processing



On Demand Processing – freeing the obligation to manage the full long term archive of all L1/L2

EO Product Lifecycle – Evolution to On Demand Processing - Impacts

- ✓ To sustain the current demand the distribution function will continue to need to manage rolling archive cache of latest “fresh” data
- ✓ Full catalogue of processed data will remain available
- ~ No long term archiving of all product levels,
- ~ As upgrade of processors are introduced they will be immediately available for ad-hoc historic data requests, complete baselines no longer assured
- ? Bug fixes “versions” of data products will invalidate definitive list
- ? No “guarantee” the same production for subsequent processing using “same” input and processors (incl. aux data)

Quite a few Provenance Starting Points



PROV-Overview

An Overview of the PROV Family of Documents

W3C Working Group Note 30 April 2013

This version: <http://www.w3.org/TR/2013/NOTE-prov-overview-20130430/>
 Latest published version: <http://www.w3.org/TR/prov-overview/>
 Previous version:

e.g. Provenance Management for Modern Distributed Systems. In Proceedings of USENIX Theory and Practice of Provenance, London, England, July 9–13, 2018

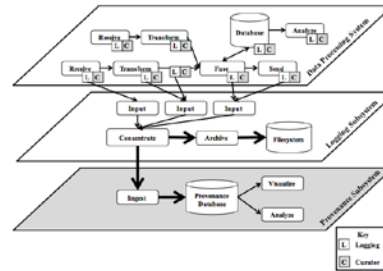
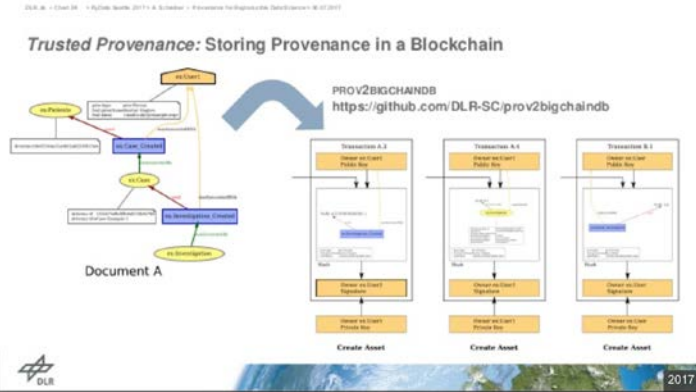


Figure 1. The Curator toolkit integrated into a microservice architecture. Some connections have been removed for legibility.

```
ProvenanceLogger logger =
    new ProvenanceLogger(Logger.getLogger("App"),
        new ProvJsonSerializer());
Entity input = new Entity();
input.setAttribute("Filename", "IMG-8942.jpg");
Activity transform = new Activity();
Used used = new Used(transform, inputData);
logger.log(inputData);
logger.log(transform);
logger.log(used);
```



Centralised vs. Decentralised Approaches

Initial Requirements

Baseline Requirements

Provide per product proof of provenance for Sentinel dataflows
a record of all processing steps (ACQ, PRO, DISSEM, LTA
..)

Provide a secure, definitive reference for valid Sentinel data
transformations

Attributable to contracted/trusted partners

Allow trusted partners to invalidate individual products

Allow trusted partners to invalidate entire product baselines

No “lock-in” for trusted partners

Allow to document reason of invalidity

Allow users to “certify” the validity of individual products

Allow users to “certify” the validity of lists of products

Allow users to identify product replacements

Fully support the Sentinel on-demand production scenarios

Distributed – Deployable in multiple facilities and cloud
infrastructures

Initial Requirements



Service Requirements

High availability, Robust to maintenance, network shortfall

Scalable - Sentinel-Scale, 100's*10⁶ products, many transformations per product

Including all components/operations necessary for management of digital signature

Capable to support data migration from initial sources (.csv etc.)

Open Source, API libraries / scriptable interfaces for integration to processor workflows

Potential Requirements

Applicable to mission planning

Allow to identify production gaps

Auditable timeline

Input for end-to-end dashboard

Extendable to per product proof of provenance for Collaborative and Downstream individual products



Way Forward



Product Lifecycle Operations Concept TN - 01/2019

For comment and review - 02/2019

Hot topic for future prototype and interoperability demonstration

