bp OSDU Data Governance- Key Principles & Standards

Entitlements & Legal Tag Management (Security & Compliance) (v1)

Commented [HC1]: @Kerr, Andrew,my comments are below. Thanks for the doc.

Document Ownership

Owner	Andrew Flack
Contributors	Andrew Kerr
	Chris Hough
	Prasanna Balakrishnan
	Paul Stapleton
	Thibault Uzu
	Greg Harris
	Chris Rose

Commented [HC2]: Broad statement, shouldn't the owner be a role, rather than a person?

Commented [AB3R2]: Discussed on call - fully agree, but for time being, the roles that might own these documents don't yet exist/havent yet been agreed ☹

Document Revisions – Change log

Date	Name	Change Comments
16/07/2024	Andrew Kerr	Document creation
17/07/2024	Andrew Kerr	Added links to OSDU Forum Standards Documentation
23/07/2024	Andrew Kerr	Further details added throughout
26/072024	Chris Hough	Comments/feedback provided & incorporated
13/08/2024	Prasanna Balakrishnan	Minor edits, additions to Entitlements Deployment, Data Partition sections
13/02/2025	Andrew Kerr	Added OSDU Production implementation framework

Document Overview

This document outlines the processes and standards required for managing entitlements and legal tags within the OSDU data platform. The goal is to ensure the right people have the right access to the right data, in line with bp's regulatory and compliance obligations, supporting bp's right to operate.

This document includes detailed definitions and recommendations for implementing and managing entitlements and legal tags within the OSDU data platform - Data must be access-controlled, governed for right of use, and have clear ownership. Every data record must include a set of 'Entitlements' (owners, users, viewers) and at least one 'Legal Tag'.

- **Entitlements Usage and Management**:
  - Groups should be named following a standards convention
  - Entitlements should be defined at the platform level & leveraged by consumers
  - Aligned with exiting bp policies, contractual obligations & country data restrictions
- **Legal Tags Usage and Management**:
  - Valid Legal Tags are a requirement for any data to be ingested to the data platform
  - Legal Tags must be managed to ensure compliance with relevant policies
  - Legal Tags may have an implication on derived data (thought ancestry/lineage)

Contents

1. Introduction

It has been recognized that we need a **managed set of processes/standards to support OSDU entitlements & legal tag management**. The intent of this standard is to ensure the **right data is available to the right people**. And, in addition, to leverage the potential value in OSDU Legal tags, providing **greater control over regulatory / compliance data obligations (supporting bp's right to operate).**

The data held within OSDU data platform must be managed securely and responsibly, such that:

- Data is Access-Controlled
- Data is Governed for Right of Use
- Data has Clear Ownership

Within the OSDU data platform, every data record that exists **must** contain a set of "Entitlements" (Owners, Users, Viewers) and at least one "Legal Tag" (definitions explained in following section).

In order to manage Entitlements & Legal Tags effectively, we will require the below functionality / capability – the technology chosen to deliver this capability is not defined as part of this standard:

- Manage Users & Groups
- Manage Permissions
- Manage, apply & maintain Legal Tags and their usage across all data
- View data associated with Entitlements & Legal Tags
- Link Entitlements & Legal tags to the relevant contractual, regulatory or legislative restrictions associated to the data
    - Which, in-turn, enables potential access management automation by leveraging Policy-Based Access Control (PBAC) over Role-Based Access Control (RBAC)

2. Definitions

2.1 Entitlements Definitions

Entitlements (otherwise known as access control tags, or "acl"s) associated with any OSDU entity are included by the SystemProperties "acl" (schema fragment) and form part of **all** OSDU records.

"AbstractAccessControlList" format:

- owners[] - List of Owners
    - The acl/group /list of owners of this data record, formatted as an email
- viewers[] - List of Viewers
    - The acl/group/list of viewers to which this data record is accessible / visible / discoverable, formatted as an email

Link to OSDU Community Gitlab

Commented [HC4]: Is this process part of a set of processes?

Commented [HC5]: OSDU by itself is not a thing. The OSDU Data Platform is

Commented [HC6]: Data Platform

Commented [HC7]: The what and the what? :)

Commented [HC8]: This is a specific schema fragment, maybe be more general or specific, if that makes sense

OSDU facilitates 2 categories of access control implementation:

- Role-Based Access Control (RBAC): Authorization based on pre-defined static roles
- Policy-Based Access Control (PBAC): Authorization determined dynamically based on policies

Microsoft's **Azure Data Manager for Energy (ADME)**, bp's PaaS solution of choice to deliver the OSDU platform, **currently only supports RBAC** (PBAC is not available – as of August 2024)

- Longer-term, Microsoft is planning to build Integrated Access Management (IAM) capabilities via "Microsoft Entra ID" (formerly "Azure AD") for entitlements management in ADME

Commented [HC9]: Maybe define what ADME is

2.2 Legal Tag Definitions

A Legal Tag is the entity that represents the legal status of data in the Data Platform. It is a collection of properties that govern how the data can be consumed and ingested.

Commented [HC10]: Should this use of the term have a space between?

Legal Tags associated with any OSDU entity are included by the SystemProperties "Legal" (schema fragment) and form part of **all** OSDU records.

Each data record in OSDU must refer to associated Legal Tags through the "AbstractLegalTags" schema (format shown below):

Commented [HC11]: Maybe drop the data modelling specific term. Most readers don't need to know or will understand. It's just part of every schema / data object / data type loaded into the Data Platform

- LegalTags[]
  - The list of legal tags, which resolve to legal properties (like country of origin, export classification code, etc.) and rules with the help of the Compliance Service.
- otherRelevantDataCountries[]
  - The list of other relevant data countries as an array of two-letter country codes
- Status
  - The legal status. Set by the system after evaluation against the compliance rules associated with the "Legal Tags" using the Compliance Service.

Note - The Legal Tag **name** needs to be between 3 and 100 characters and only alphanumeric characters and hyphens are allowed – **the Legal Tag name acts as the unique ID for the Legal Tag**

Commented [HC12]: And this is the unique identifier

Link to OSDU Community Gitlab

While Legal Tags are primarily used to support compliance and entitlements, they also describes some aspects of ***provenance*** related to the entities that originated a data object – for example the two properties within the Legal Tag schema, shown below:

- DataType – the class of ownership of the data relative to the company's data platform; shows whether the data originated internally or externally
- Originator – the company that owns the data, and is a proxy for the source organization

Commented [HC13]: Maybe be clearer that these are specific properties

3. Entitlements Usage & Management

3.1 Entitlements Deployment Recommendations

The below recommendations are aligned with "Architecture OSDU KDD (2024)":

1. Use RBAC for initial OSDU data platform deployment: leveraging the existing AAD groups from dsbp
    - Start with a **simple, open model** then refine over time as needed per data type (driven by defined Use-Cases)
    - **Tailor access model over time** in-line with data type requirements or country/asset-specific access control requirements (e.g. access can be restricted to regional/asset users, as required, based on data type and/or regional data regulation, legislation or contractual obligation)
        - **Current Stratus Access Model (and potentially existing AAD Groups) could be used as initial basis for an OSDU Access Model**
        - **In the absence of PBAC capabilities in ADME, there may be a need to review and restructure the existing AAD groups for fine-grain control**
    - **Refine access model maturity** over time (e.g. leverage OSDU entitlements control to enable separation between data discovery – via the metadata catalogue – vs access to the underlying data content – via storage/DDMS
    - ADME supports Single-Sign-On (SSO) using existing AAD groups. Application integration with ADME should leverage the OSDU entitlement's model to **control Application user access** to datasets that reside within ADME.
2. Test and review PBAC capabilities when available in ADME
    - In order to usefully leverage a PBAC approach, we will be dependent on leveraging both the Entitlements ***and*** OSDU Legal Tags (which should align with internal data policies and honour any regional data regulation, legislation or contractual obligation)

3.2 Summary of key differences between RBAC & PBAC approach

	RBAC (Role Based Access Control)	**PBAC (Policy Based Access Control)**
Permission depends on	Predefined groups with the set of permissions across each group.	Sets of specific policies that determine relevant attributes and permissions.
Types of access control	Coarse-grained	Fine-grained
Real-time parameters	No	Yes
Authorisation	Limited to role-based authorization, does not take into account other security controls, e.g., user location or time of the day	Authorization based on device, location, time, and other security controls possible
Ease of compliance	Assignment of a group may give certain users unwanted access to sensitive information. Compliance is difficult to ensure.	Reliance on policies allow business managers to set permissions directly, ensuring high compliance.

Commented [HC14]: Link to the document in the DMS?

Commented [AB15R14]: "DMS" lol - but yeah meant to provide a link, added now!

Commented [HC16]: There are different groups for the metadata access, compared to the content (for WPC). I just feel some, mostly P&O folks, would be concerned with full open access to the data, which isn't the case.

Commented [AB17R16]: Original wording taken from architecture KDD - will re-word more appropriately here - good spot!

Ease of modification	Fairly easy to create new groups or revoke certain permissions for a particular group, but the risk of role explosion is high.	Relatively easier to define new policies using high-level policy language.

3.3 Entitlements Management (& Ownership)

Management & Ownership Summary

While bp are still following the RBAC approach, OSDU entitlements (i.e. ACLs or AAD Security Groups) **must** be actively managed and maintained. ACL members should be regularly reviewed and kept current (in-line with bp data policies & guidance – see below).

In the **current** bp digital landscape, access and permissions (entitlements) have been historically managed by the asset, data product, or digital product teams (with respect to their specific needs).

In an **OSDU-centric digital landscape**, bp has the opportunity to govern access more efficiently at the **data platform level**. Therefore, entitlements should be managed consistently across data types/kinds, with implementation owned by the relevant data owner/data product owner. These entitlements should (ideally) then **be honoured by any consuming application or tool**.

bp Data Policies & Guidance

Refer to existing work done to summarize regional data restrictions –

- Global Legal Guidance – Data Policies - Power BI Report (may be out of date)
- Country-Level Legal Guidance (may be out of date)

Entitlements Service Overview

The entitlements service is used to create and manage access & permissions for the data held within the OSDU data platform, in the form of "groups". A group **name** defines a **permission**. Users who are added to that group **obtain that permission**:

Commented [HC18]: "Data Platform"

- **Data** groups are used for **data authorization** e.g. data.welldb.viewers, data.welldb.owners
- **Service** groups are used for **service authorization** e.g. service.storage.user, service.storage.admin
- **User** groups are used for hierarchical grouping of user and service identities e.g. users.datalake.viewers, users.datalake.editors

For each group you can either be added as an OWNER or a MEMBER. The only difference being if you are an OWNER of a group, then you can manage the members of that group.

Group Naming Convention

ACLs/Groups must follow the OSDU naming convention – each group's name should start with:

- "data." for data groups
- "service." for service groups
- "users." for user groups

Note - a group names **are** case-insensitive.

As such, all group IDs must be in the of form of an email, following the below convention:

{groupType}.{serviceName|resourceName}.{permission}@{partition}.{domain}.com

with:

- groupType {'data', 'service', 'users'}
- serviceName {'storage', 'search', 'entitlements', ...}
- resourceName {'welldb', 'npd', 'datalake', 'public', ...}
- permission {'viewers', 'editors', 'admins' ...}
- data-partition-id {'bp-ADME', 'common', ...}
- domain {bp.com' ...}

e.g.

- Data groups (data authorization):
  data.{resourceName}.{permission}@{data-partition-id}.{domain}.com
- Service groups (service authorization):
  service.{serviceName}.{permission}@{data-partition-id}.{domain}.com
- User groups (hierarchical grouping of user and service identities):
  users.{serviceName}.{permission}@{data-partition-id}.{domain}.com

For an **RBAC implementation** - It is recommended that *individual* users are not added directly to OSDU groups, but instead bp should **leverage existing, managed AAD user groups/security groups** which, in-turn, should be **nested within the appropriate OSDU entitlement group**. These AAD groups should then be used for individual user access.

Commented [HC19]: Okay, so not individual users. Maybe say that?

Impact of OSDU Data Partitions

A group is unique to each **data partition**. This means that access is defined on a **per data partition basis** i.e. giving a service permission in one data partition **does not give that user service permission in another data partition**.

Data Partitions need to be used in conjunction with the Entitlements model to control access at the regional/data domain levels.

3.4 Entitlements Implementation (Q1 2024 – Early Production)

It has been agreed that we should keep the access model simple initially, mirroring that of dsbp.

All users may have open “Viewer” (read-only) access to all data in OSDU, and only OSDU Admins, DMs and those supporting the OSDU programme will have “Editor” (read/write) access to all data (this may include direct application access to support workflow testing and implementation).

As for Legal Tags it was agreed that we would create and put in place “placeholder” legal tags based as a minimum on Country, Asset and tight/restricted wells. These would be generic legal tags with no expiry date, but would allow us to (1) update tags at a later data with appropriate legal information, and (2) allow us to query all data based on a clear set of country/asset/tight legal tags such that we may **refine the entitlements model in bulk at a later date** based on these criteria (should we wish to).

The below assumes that **correct metadata** has been applied in dsbp to denote wells restricted as bp Confidential, associated field or asset details, and country/location (and that this metadata has been mapped and brought across to OSDU during migration).

Initial Entitlements Implementation Plan

Apply the below Entitlements – applicable to ALL records loaded to OSDU Prod

"acl":{
"viewers: [
data.default.viewers@admedevdp.dataservices.energy (NOTE – Naming standard to be agreed)
],

- Mirror dsbp (Viewer access to all data products for all bp)
- Nest within this OSDU group the appropriate bp AAD Security Groups (all bp)
- Same applies for Trajectories & Reference data (non-dsbp sources)
- May opt to create new viewer group, specific to data type/asset OR rename the resourceGroup element from "default" to something like "OSDU" or "AllData"?

"owners": [
data.default.owners@admedevdp.dataservices.energy (NOTE – Naming standard to be agreed)
],

- OSDU "Admin" only access to write or edit data within OSDU
- This would include the creation of & nesting of the below groups
  - OSDU Admin AAD group (to include platform DMs, admins and OSDU Program DMs)
  - Any existing Technology AAD groups from each data product team
- Note - access may also include direct application access, to allow for workflow testing, etc.
- May opt to create new viewer group, specific to data type/asset OR rename the resourceGroup element from "default" to something like "OSDU" or "AllData"?

Mature Production

The intent is to re-evaluate BOTH the Entitlements model and Legal Tag model at a later date, and refine it over time (e.g. provide access control based by country, asset or data type, or to restrict access to sensitive data AND/OR provide more controlled asset and country-related legal tags).

The following will be considered at a future date:

- Country (Region) or Asset-specific entitlements model, mirroring that of the existing Stratus S: Drive and/or Studio access management model
- Service model supporting access control and management (via the likes of MySubsurface)
- Application access control & write-back principles (what apps can write what data kinds to ADME)
- Framework to support Application inheritance of ADME entitlements (e.g. ensure entitlements defined within ADME are inherited/honoured by consuming applications)

3.5 Full OSDU Entitlements Service Documentation

See OSDU API Documentation for current details on Entitlements creation & management:

https://osdu.pages.opengroup.org/platform/security-and-compliance/entitlements/api/#group-creation-guidelines

4. Legal Tag Usage & Management

4.1 Legal Tag Usage

Overview

A legal tag is **required** for data ingestion. Therefore, when data is ingested to the OSDU data platform, the ingesting app/pipeline/process must assign an appropriate/valid Legal Tag to be assigned to the record, prior to ingestion. The validity of this Legal Tag is then verified by the ingestion service.

More than one Legal Tag may be assigned to a single data record (e.g. a data record may need to conform to country-related data regulation *in addition* to contractual limitations on the use of the data itself). However, should *either* Legal Tag expire, the data will be soft-deleted from the platform (see more on this in following sections).

If there isn't an appropriate, valid Legal Tag available to support data ingestion, the ingestion service will **reject** the record and it will not be ingested to the platform. A correct/valid Legal Tag should instead be used OR, a new Legal Tag must 1st be **created in the platform (prior to ingestion** of the record).

The creation & management of Legal Tags should be **a governed process, managed centrally across the OSDU data platform**. Ideally, Legal Tags should be **managed from a single place** (i.e. a designated Data Management Tool); however, there ***may*** be future use-cases where specific applications or data workflows may be granted the authority to create new Legal Tags (in which case, these should be **governed by appropriate business rules**).

Note - When generating a new Legal Tag, the Legal Tag name is the unique identifier that is used for reference by any associated data record.

What happens when data is ingested to the data platform?

When data is ingested, it **must** have an assigned Legal Tag name -this name is checked for a corresponding valid Legal Tag in the system:

- A ***valid*** Legal Tag means it **exists and has not expired**
- If a Legal Tag is ***invalid***, it means the chosen Legal Tag either **does not exist or has expired**, and the data is **rejected**
    - e.g. we may not allow ingestion of data from a certain country due to regulatory reasons, or we may not allow consumption of data that has an expired contract, or for an asset that has been sold, etc.
    - In the case where no Legal Tag exists – it must be created within the data platform prior to ingestion

Upon ingestion of data, it's the ***responsibility of the app/pipeline/process ingesting that data to use the appropriate Legal Tag, conforming to agreed bp business rules and processes.***

What happens when a legal tag expires for a record already in the data platform?

In the same manner as above, the existing data will be **invalidated (soft-deleted)** when the legal tag expires, as it would **no longer be compliant**.

Therefore, it is imperative that Legal Tags are actively managed, maintained and kept in-line with the relevant policies, contracts, regulation and legislation for which they relate.

Commented [HC20]: "Data Platform"

Commented [HC21]: Isn't that part of the ingestion service?

Commented [HC22]: Space?

Commented [HC23]: Before ingestion. I'm pretty sure the service will knock it back if it's not there

Commented [HC24]: Hm, should apps be able to do it? A philosophical question indeed!

Commented [HC25]: yeh

Commented [HC26]: I think it's just the word "Process" that throws me off this sentence

Commented [AB27R26]: I think in this case I'm thinking of like "business process or workflow" - something like ADI, where a new well is generated as part of a digital workflow?

4.2 Legal Tag Management (& Ownership)

Management & Ownership Summary

Legal Tags in OSDU data platform form a crucial part of ***all*** data stored in the platform. They govern the legal rights to access and use the data held in the platform, along with critical information to support appropriate governance of that data.

As such it is recommended that the Legal Tags in use within bp OSDU are actively managed by a dedicated team(s), such that the Legal Tags are kept current & accurate. This may be a combined effort, bringing together Asset data owners (country-, regional- or asset-specific regulation, legislation, or contractual obligation) and data product owners (data-type-specific regulation or contractual obligations surrounding global 3rd party data usage).

Note – see section "3.1 Entitlements Deployment Recommendations" for dependency on Legal Tags to enable effective PBAC entitlements implementation. Legal Tags can be used to ensure Policy Based Access Controls honour any regional data regulation, legislation or contractual obligation captured in the within the Legal Tags.

bp Data Policies & Guidance

Refer to existing work done to summarize regional data restrictions –

- Global Legal Guidance – Data Policies - Power BI Report (may be out of date)
- Country-Level Legal Guidance (may be out of date)

Legal Tag Properties

Below are details of the properties you can supply when creating a Legal Tag along with the values that can be used (which may be specific to each data partition). All values are mandatory unless otherwise stated.

- **Country of Origin**

  Definition - the country from where the data **originated**, NOT from where it was sent

  - Array of ISO Alpha-2 country codes
    - *Note – unlike most other reference data, this reference list is managed by the Legal Tag Service, not by data definitions*
  - Property is case-sensitive

- **Contract Id**

  Definition - the Contract Id associated with the data, if applicable

  - Must be between 3 and 40 characters and only include alphanumeric values and hyphens
  - 'Unknown' or 'No Contract Related' may be used where no relevant contract ID

- **Expiration Date**

  Definition - sets the inclusive date when the Legal Tag expires and the data it relates to is no longer usable in the Data Platform. This normally is taken from the physical contracts expiration date i.e. when you supply a contract ID.

  - Any date in the future, in the format yyyy-MM-dd (e.g. 2099-12-25)
  - This is a non-mandatory field & can be left empty

Commented [HC28]: The OSDU Data Platform

Commented [KA29]: @Hough, Chris - note on difference in managing ref data via legal services etc

Commented [HC30R29]: Yes, this list is managed for the Legal tag service, not by data definitions

        - If left empty it will be auto-populated with the value 9999-12-31

    - Contract expiry *should* be mandated for certain types of data e.g. 3rd party/licensed data, where a relevant time-bound contract ID exists (however, this is not enforced by the platform & is reliant on company-specific business rule implementation)
        - If the provided contract ID is "Unknown" or "No Contract Related", then the expiration date may be empty. However, if a date is provided, then will be honored in validating the legal tag and the associated data, even if there's no contract is provided.

    Note – when a Legal Tag expires, ALL associated data will be soft-deleted from OSDU (therefore will not be returned by search & storage APIs)

- **Originator**

  Definition - the name of the client, organisation or supplier from which the data originated (from bp's perspective)
    - Property is case-sensitive
- **Data type**

  Definition – the "type" of data in reference to ownership. This may be used in a ***governance*** context, to determine whether we (bp) should expect a contract to be present and assigned to the data held against this Legal Tag (*examples* shown below)

  Note – bp can define what data types are allowed within any one partition
    - Public Domain Data (e.g. public data) - no contract ID required
    - First Party Data (e.g. bp proprietary data) - no contract ID required
    - Second Party Data (e.g. client data) - contract ID must be present
    - Third Party Data (e.g. licensed data) - contract ID must be present
- **Security classification**

  Definition - the standard security classification for the data

  Note - Secret data is not permitted to be stored in the OSDU data platform
    - Public
    - Private
    - Confidential

  This property is NOT case sensitive

Commented [HC31]: That's a long time in the future!

Commented [HC32]: This is really important!
And OSDU Data Platform, of course :)

- **Export classification**

  Definition – Export Control Classification Number (ECCN) identifying dual-use items for export control purposes (in this case 'dual-use technology' *may* include data).

  Note - only data with the ECCN classification 'EAR99' and '0A998' are permitted in OSDU

  - EAR99[1]
  - 0A998 (0 as Zero)[2]
  - Not - Technical Data
  - No License Required

  This property is NOT case sensitive

- **Personal data**

  Definition – identifies whether data is personally identifiable or not, to support conformance with GDPR (and other global regulations)

  Note – OSDU does **not** currently allow data that is considered **Sensitive Personal Information**

  - Personally Identifiable
  - No Personal Data

  This property is NOT case sensitive

Example LegalTag json record

```
"LegalTags": [
    {
      "name": "osdu-thirdparty-public",
      "description": "",
      "properties": {
        "countryOfOrigin": [
          "US"
        ],
        "contractId": "A1234",
        "expirationDate": "2099-01-25",
        "originator": "OSDU",
        "dataType": "Third Party Data",
        "securityClassification": "Public",
        "personalData": "No Personal Data",
        "exportClassification": "EAR99"
```

Commented [HC33]: Can you include definition for these export classifications. I did google it once upon a time

[1] **ECCN EAR99:** data that can generally be exported without a license; however, careful due diligence is required to ensure the data is not going to an embargoed or sanctioned country, a prohibited end-user, or used for a prohibited end-use (as this may then require a license)

[2] **ECCN 0A998:** This is a new ECCN that covers certain specified oil and gas exploration equipment and software. It encompasses oil and gas exploration data, e.g., seismic analysis data, and certain hydraulic fracturing (commonly known as "fracking") items, including hydraulic fracturing design and analysis software and data. "Data" is a new item category that is now controlled by the EAR and subject to a license requirement, and "data," as used in ECCN 0A998, is not the same as "technology" as defined in the EAR.

Legal Tag Expiry

One of the main cases where a Legal Tag can become invalid is if a contract expiration date passes. **This makes both the Legal Tag invalid and all data associated with that Legal Tag *(potentially including derivatives)*.**

When a Legal Tag is deemed to be invalid, ALL data associated with that Legal Tag will be soft-deleted from OSDU (therefore will not be returned by search & storage APIs).

Updating a Legal Tag

Existing/invalid Legal Tags **CAN be updated** to make them valid again, and so make all associated data re-accessible.

The properties that can be updated are:

- description
- contract ID
- expiration date
- extensionProperties properties

Note – you cannot update the Legal Tag Name (as this would break existing references – due to the fact that the name is used as the unique identifier for the Legal Tag)

Commented [HC34]: As it is the unique identifier

4.3 Data Derivatives - Inheritance / Ancestry of Legal Tags

What are Derivatives?

In the context of the OSDU data platform, the term "derivative data" is data that has been derived from primary data sources.

Commented [HC35]: Data Platform

Often when ingesting data into the platform, it is the raw data itself. In these scenarios, bp ***may*** associate a single Legal Tag with this data.

However, in the case when the data to be ingested comes from multiple sources, it is the case of derivative data.

Example use-cases:

- A bp practitioner uses multiple OSDU records to create a new interpreted record
- A bp practitioner runs an algorithm over some seismic data to process it and create a new attribute volume

Commented [HC36]: This would be the more common use case. Another common term maybe processed or interpreted

How do derivatives relate to Legal Tags?

In these scenarios, we ***may*** opt to assign Legal Tags to this new data which is the ***union*** of the Legal Tags associated to ***all source data records*** from which it was derived.

Commented [HC37]: It's a very basic form of data lineage

(e.g. I have Data A associated with Legal Tag 1, and Data B associated with Legal Tag 2. If I create Data C from Data A and Data B, then I ***may*** need to associate both LegaltTag 1 and Legal Tag 2 to Data C)

We ***may*** opt to ***inherit the Legal Tag(s)*** associated with the ***parent "master data" record*** (e.g. a newly-created well log may inherit the Legal Tag from the associated Wellbore) or those ***associated with the data from which the new record was derived*** (e.g. a derived composite well log may inherit the Legal Tag from the parent well log(s) from which the interpretation was made).

OSDU facilitates this inheritance through the "**AbstractLegalParentList**"

The "AbstractLegalParentList" is an array of none, one or many entity references of 'direct parents' in the data platform, which mark the current record as a derivative (note the source record version is required – so 'id:version')

During record creation or update the ancestry.parents[] relationships ***may*** be used to collect the legal tags from the sources and aggregate them in the legal.Legal Tags[] array.

Link to OSDU Community Gitlab

NOTE – OSDU documentation mandates that the parent Legal Tags *MUST* be included against any derived record in OSDU; however, bp has changed this terminology to *MAY*. This is because not all Legal Tags on a record will necessarily extend to the data derived *from* them – e.g.

- what if bp has limited rights to use partner or 3rd party data, but full rights and ownership to any derived interpretation data made based on this partner or 3rd party data? Therefore the Legal Tags may need to differ between parent and child records.

However, we will have many use-cases (particularly for bp-owned data) where we will want to ensure that Legal Tags *DO* inherit from their parent records (which is the simplest solution). In this case, it is important to note that, if one or more of the legal tags of the source data expires, the access to the derivatives is also terminated.

See additional guidance here: Lineage & Ancestry Capture (incl. Activity Model & Business Decisions)

Creating derivative Records

When creating Records that represent derivative data, the following must be assigned:

- The Record Id and version of all the Records that are the ***direct parents*** of the new derivative (added to the ancestry section)
- The Alpha-2 country code of where the derivative was ***created***

If populating, the Record Service should be used to populate Legal Tag values based on the parents.

e.g.

```
"legal" :{
"otherRelevantDataCountries": ["US"] },
//the physical location of where the derivative was created

"ancestry" :{
"parents": ["osdu:id:1:version", "osdu:id:2:version"] }
//the record ids and versions of the Records this derivative was created from
```

4.4 Legal Tag Implementation (Q1 2024 – Early Production)

It was agreed that we would create and put in place "placeholder" legal tags based as a minimum on Country, Asset and tight/restricted wells. These would be generic legal tags with no expiry date, but would allow us to (1) **update tags at a later data** with appropriate legal information, and (2) allow us to query all data based on a clear set of country/asset/tight legal tags such that we may refine the entitlements model in bulk at a later date based on these criteria (should we wish to).

The below assumes that **correct metadata** has been applied in dsbp to denote wells restricted as bp Confidential, associated field or asset details, and country/location (and that this metadata has been mapped and brought across to OSDU during migration).

Initial Legal Tag Implementation Plan

It has been agreed to provide a ***minimum*** set of Legal Tags, assigned to each data record, based on the country, asset and originator of the data record. The intent is to programmatically derive this information from the related parent Well Header for each record.

Naming Standard

Legal Tag name: ***"ISO 2-Character Country Code"-"Field/Prospect/Exploration"-"Originator"***

"ISO 2-Character Country Code" – derived from the Country defined in the Well / Well Origin record

"Field/Prospect/Exploration" – derived from the Field or Prospect defined in the Wellbore record

- IF "Field" is defined, use "Field Name"
- ELSE IF "Field" is NOT defined, use "Prospect Name"
- ELSE IF both "Field" & "Prospect" is NOT defined use "Exploration"

NOTE ON SPECIAL CHARACTERS and spaces – where field names are separated by a "space" these should be delimited by a "-" (same goes for all special characters such as ( ) / _ & ')
e.g. BPO Greater Tortue Ahmeyim field in Mauritania > "MR-Greater-Tortue-Ahmeyim-bp"

"Originator" – Derived from the "Operated Status" defined in the Well / Well Origin record

- IF Operated Status is "BPO" or "BPNO", use "bp"
- ELSE IF Operated Status is "NO", use "ThirdParty"

Example Legal Tags

Below are some example Legal tags that will be added to the platform and can be programmatically derived based on the described logic:

"ISO 2-Character Country Code"-"Field/Prospect/Exploration"-"Originator"

Example 1:
A Legal Tag associated with a Well log from the bp-operated Clair field in the UK would be –

"**GB-Clair-bp**"

```
"LegalTags": [
    {
      "name": " GB-Clair-bp",
      "description": "Placeholder bp Asset Legal Tag",
      "properties": {
        "countryOfOrigin": [
          "GB"
        ],
        "contractId": " No Contract Related ",
        "expirationDate": NULL,
        "originator": "bp",
        "dataType": " First Party Data",
        "securityClassification": " Confidential",
        "personalData": "No Personal Data",
        "exportClassification": " No License Required"
```

Example 2:
A Legal Tag associated with a Trajectory from a bp OBO Wildcat Exploration well in the US would be

"**US-Exploration-bp**"

```
"LegalTags": [
    {
      "name": " US-Exploration-bp",
      "description": "Placeholder bp Exploration Legal Tag ",
      "properties": {
        "countryOfOrigin": [
          "GB"
        ],
        "contractId": " No Contract Related ",
        "expirationDate": NULL,
        "originator": "bp",
        "dataType": " First Party Data",
        "securityClassification": " Confidential",
        "personalData": "No Personal Data",
        "exportClassification": " No License Required"
```

Example 3:
A Legal Tag associated with a Well Header record for a non-operated legacy well in the UK, that bp has licenced the information for, would be -

"**GB-Exploration-ThirdParty**"


```
"LegalTags": [
    {
     "name": "GB-Exploration-ThirdParty ",
     "description": " Placeholder non-bp Exploration Legal Tag ",
     "properties": {
      "countryOfOrigin": [
       "GB"
      ],
      "contractId": "Unknown",
      "expirationDate": NULL,
      "originator": "bp",
      "dataType": " Third Party Data ",
      "securityClassification": " Confidential",
      "personalData": "No Personal Data",
      "exportClassification": " No License Required"
```


Programmatic implementation flow

Mature Production

The intent is to re-evaluate BOTH the Entitlements model and Legal Tag model at a later date, and refine it over time (e.g. provide access control based by country, asset or data type, or to restrict access to sensitive data AND/OR provide more controlled asset and country-related legal tags).

The following will be considered at a future date:

- Incorporate links to Country-, Asset- or Licence-specific contract(s), legislation or regulation
- Refine "Placeholder" legal tag names and content for all "ThirdParty" legal tags to reflect bp's contractual right to the data (e.g. Legal Tag to specifically tag IHS or Wood Mac licenced data)
- Build out comprehensive legal tag framework to support PBAC access control

4.5 Related Principles & Standards

The section above relates to topics also covered in the below Principles & Standards:

- ***Lineage & Ancestry Capture (incl. Activity Model & Business Decisions) Standard***
    - Considerations around impact of Lineage & Ancestry on Legal / Contractual restrictions of any derivatives or related data
- ***Data Quality Standards***
    - "Baseline" DQ Rules may be used to govern the use of minimum required attributes

4.6 Full OSDU Legal Service Documentation

See OSDU API Documentation for current details on Legal Tag properties & creation:

https://osdu.pages.opengroup.org/platform/security-and-compliance/legal/api/#creating-a-Legal Tag

Appendix 1 – Relevant OSDU Forum Standards Documents (for Reference)

OSDU Forum Documentation URL - https://osduforum.org/getting-started/osdu-documentation/#

OSDU Entitlements Service Documentation - https://osdu.pages.opengroup.org/platform/security-and-compliance/entitlements/api/#group-creation-guidelines

OSDU Legal Service Documentation - https://osdu.pages.opengroup.org/platform/security-and-compliance/legal/api/#creating-a-Legal Tag

Key Documents (links to bp internal copies – refer to above URL to check latest updates):

- OSDU Reference Architecture
- OSDU Schema Usage Guide
- OSDU Technical Standard
- OSDU System Concept