The term **Provenance** is referred to as

> ‘The beginning of something’s existence; something’s origin’

A record of ownership of a work of art or an antique, used as a guide to authenticity or quality.[1]

Provenance Tracking is crucial in scientific studies where workflows have emerged as an exemplar approach to mechanize data-intensive analyses.

Changes in workflow design and integration alter the data input and output which results in the need for persistent versions of the workflows. However, this results in data versioning issues and difficulties in tracking changes to input data sources and metadata. Two key problems are encountered in managing scientific data and workflow provenance:

- **Data Versioning**: How to manage versions of data?
- **Workflow Versioning**: How to manage versions of workflows?

**Overview of CWL and RO**

**COMMON WORKFLOW LANGUAGE**

- **Declarative approach** to workflow declaration, make minimal internal assumptions about precise software environment, base software dependencies, configuration settings, alteration of parameters and software versions.
- **Promotes collaborative research** by providing a standard to represent analysis.
- **Provides an extensible open source solution** to develop interoperable research.
- **Supports use of containers** such as **Galaxy, Taverna or Arvados**.
- **Interoperability** and **portability**

**RESEARCH OBJECT**

- **Aggregates and organizes the workflow outputs, workflow definition, inputs used for workflow run, description of execution environment and attribution details**.
- **Provides JSON-LD format manifest** to describe the research object, listing inputs, outputs, workflow run details and external URIs if any.
- **Abstract representation of workflow specification** is described using **wfdesc**.
- **Attribution details** annotated by using unique identifiers to establish credit and attribution for publication to facilitate the re-use of experiment data.
- **Domain-neutral** hence flexible in allowing researchers to use vocabularies to facilitate the representation of domain-specific information including title, hypothesis and conclusion.
- **Records evolution** of a research object by utilizing ‘roevo’ ontology.

**Provenance**

The term **Provenance** is referred to as

‘The beginning of something’s existence; something’s origin’

A record of ownership of a work of art or an antique, used as a guide to authenticity or quality.[1]