



Introduction to Cask Hydrator and Cask Tracker

The **Introduction to Hydrator and Tracker** course will cover concepts and background for Cask Hydrator and Cask Tracker. You will learn to build and manage custom Hydrator Plugins. You will also learn to explore metadata, audits and lineage of your data.

Audience & Prerequisites

This course is designed for developers and data scientists who are looking manage data ingestion and data processing in a code free manner, and manage the metadata and lineage of their data. Participants should have completed the Introduction to CDAP course as a prerequisite.

Materials Required

- Laptop with Mac, Windows or Linux on 64-bit OS.
- Virtual Box installed (<https://www.virtualbox.org/wiki/Downloads>)
- CDAP Standalone VM installed (Instructions provided separately)

Course Overview

- Introduction to Cask Hydrator
- The Hydrator Studio
- Creating and managing Hydrator pipelines
- Batch & Real-time pipelines
- Writing custom Hydrator Plugins
- Hydrator APIs and managing plugins
- Introduction to Cask Tracker
- Metadata, Dataset Lineage and Usage Metrics
- Cloudera Navigator Integration



Course Outline

<p>1. Hydrator Overview</p> <ul style="list-style-type: none">• Need for Hydrator• Hydrator capabilities• Hydrator applications <p>2. Introduction to Hydrator Studio</p> <ul style="list-style-type: none">• Overview of Hydrator Studio capabilities• Walk through building pipelines <p>3. Batch Pipeline</p> <ul style="list-style-type: none">• Overview of Batch pipeline• Plugins for batch plugin• Advanced options• Post actions <p>4. Real-time pipelines</p> <ul style="list-style-type: none">• Overview of Real-time pipelines• Plugins for real-time pipelines	<p>5. Joins, Macros and Actions in Hydrator</p> <ul style="list-style-type: none">• Joiner plugin• Macro-substitutable properties• Macro functions• Control flow in Hydrator using Action plugins <p>6. Introduction to artifacts and plugins</p> <ul style="list-style-type: none">• Overview of artifacts• Creating applications from artifacts• Writing customer Hydrator plugins <p>7. Hydrator APIs and Packaging</p> <ul style="list-style-type: none">• Plugin APIs• Packaging custom plugins <p>8. Introduction to Cask Tracker</p> <ul style="list-style-type: none">• What is Cask Tracker• Metadata• Dataset Lineage, Audit logs• Data Preview and Usage Metrics• Cloudera Navigator Integration
---	---