Introduction to Elyra

Upkar Lidder, @lidderupk
Developer Advocate

Patrick Titzler, @ptitzler
Developer Advocate
Center for Open Source Data and AI Technologies

IBM Developer

IBM

IBM Developer

IBM Developer @ibmdeveloper

Hi there! 👋 We are Developer Advocates 🙇‍♀️🙇‍♂️ from 🌎🌍🌏. We are here to help you and provide developer education 🎓. Check out all of our developer resources, tools, and content at https://developer.ibm.com/

Planet Earth 🌍🌎🌏

Following

CALL FOR CODE
open source projects need you

Join the cutting-edge community building open source projects to fight back against the most pressing issues of our time. See your code deployed to help those in need.

Learn how to get started →

Courses and Programs taught by Upkar Lidder

Introduction to Cloud Development with HTML5, CSS3, and Java...
IBM

Course

Introduction to Containers, Kubernetes and OpenShift
IBM

Course

Developing Cloud Applications with Node.js and React
IBM

Course

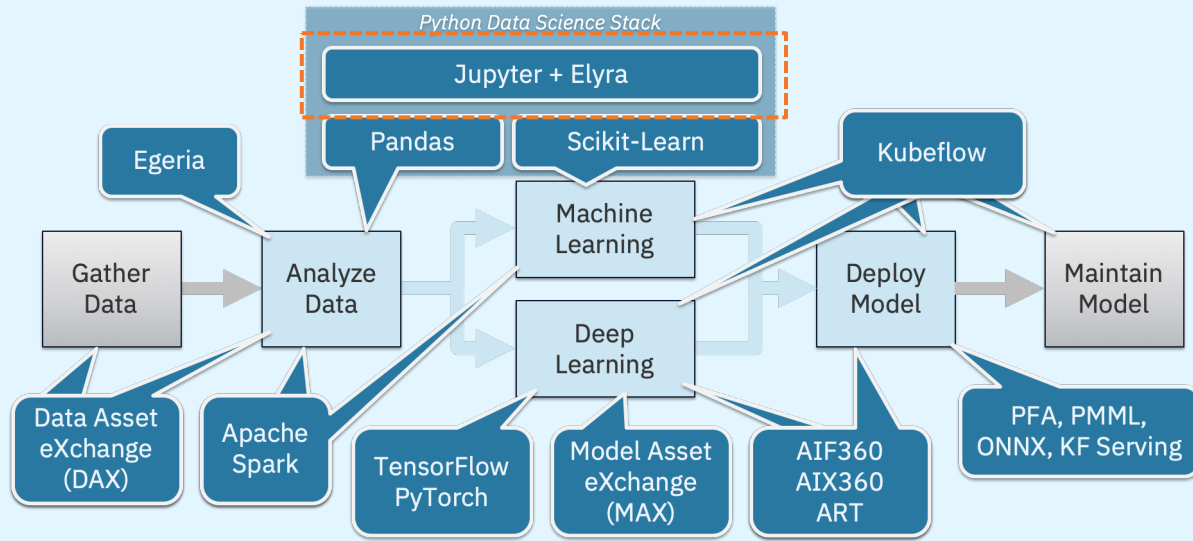### Developing Cloud Native Applications
IBM

Course

- CODAIT aims to make AI solutions dramatically easier to create, deploy, and manage in the enterprise.
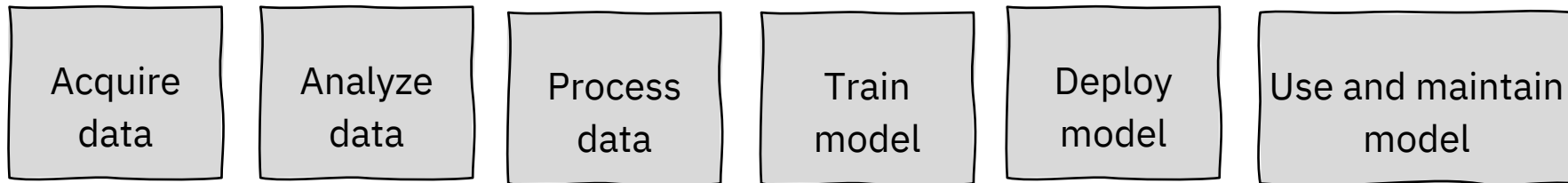
- 40+ developers/data scientists

- We contribute to and advocate for the open-source technologies that are foundational to IBM's AI offerings.

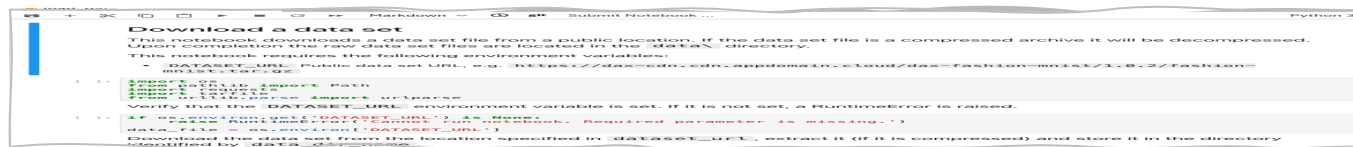Improving the Enterprise AI Lifecycle in Open Source
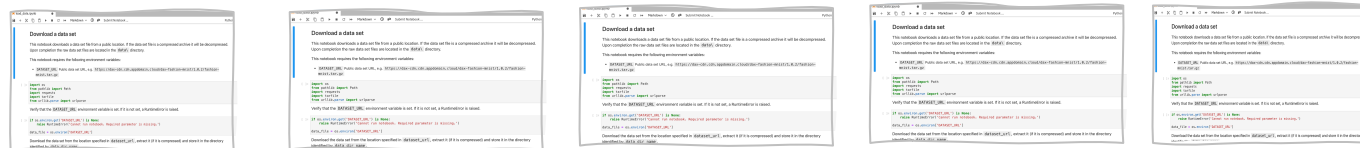
Machine Learning (ML) Workflows

- Typical workflow tasks

Acquire data	Analyze data	Process data	Train model	Deploy model	Use and maintain model

- Many tasks comprise of sub-tasks and are performed iteratively
- Jupyter notebooks are frequently used

(monolithic – does many things)

(modular)

Elyra:Set of AI-centric extensions to JupyterLab

re-use code

Code snippets

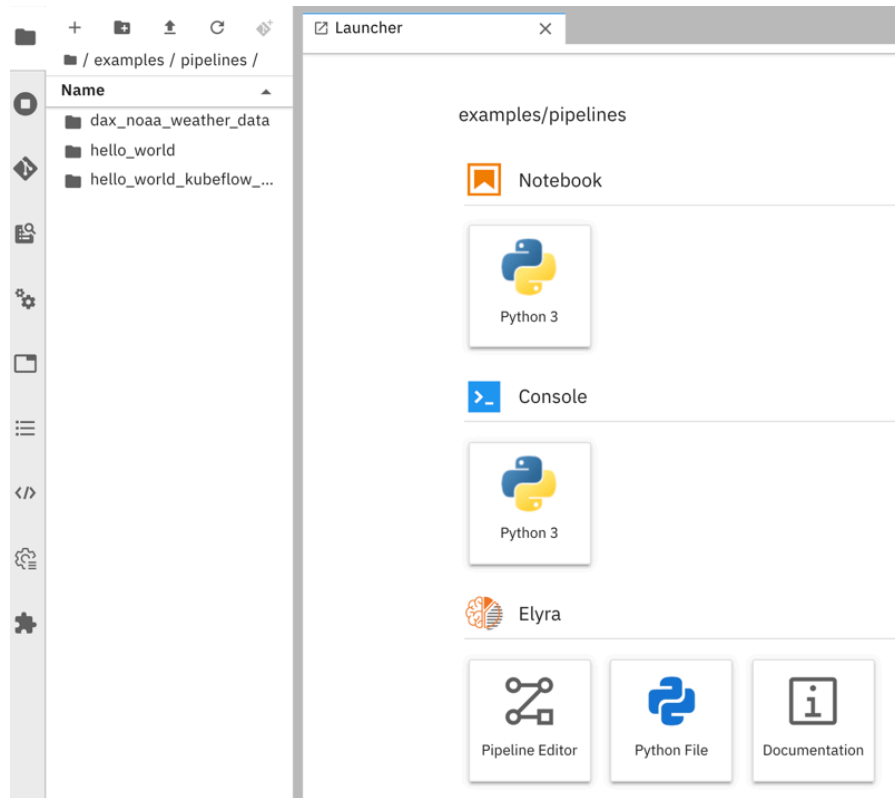source control

Git integration

run remotely

Python scripts

run batch

Notebooks

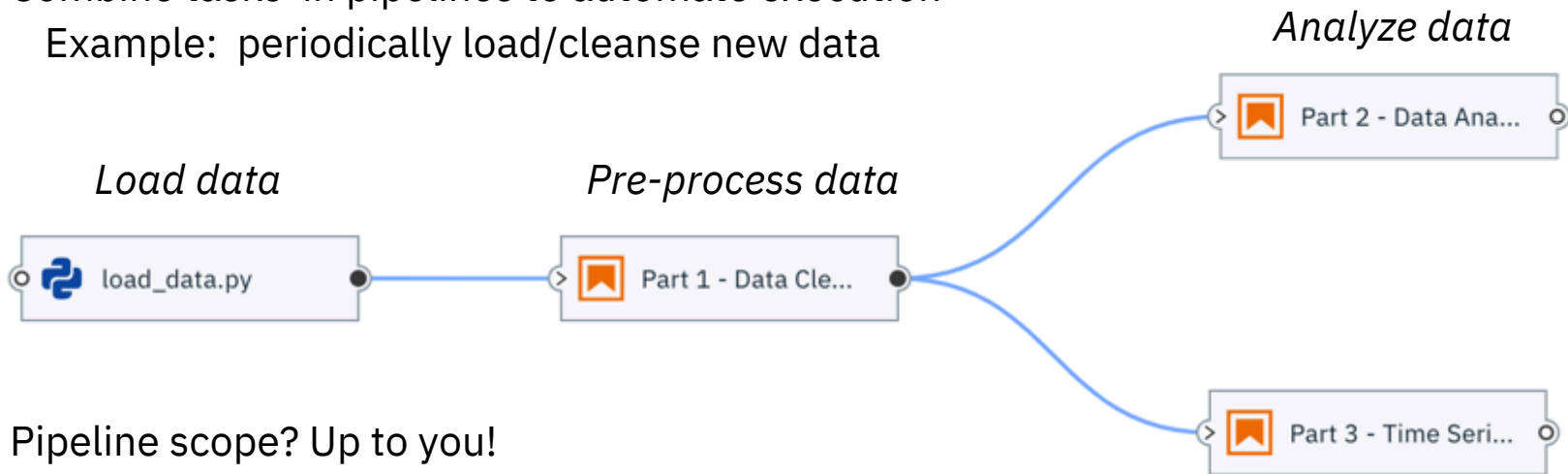ML workflow

Pipelines

/ examples / pipelines /

Name

📁 dax_noaa_weather_data
📁 hello_world
📁 hello_world_kubeflow_...

Launcher ✕

examples/pipelines

Notebook

Python 3

Console

Python 3

Elyra

Pipeline Editor   Python File   Documentation
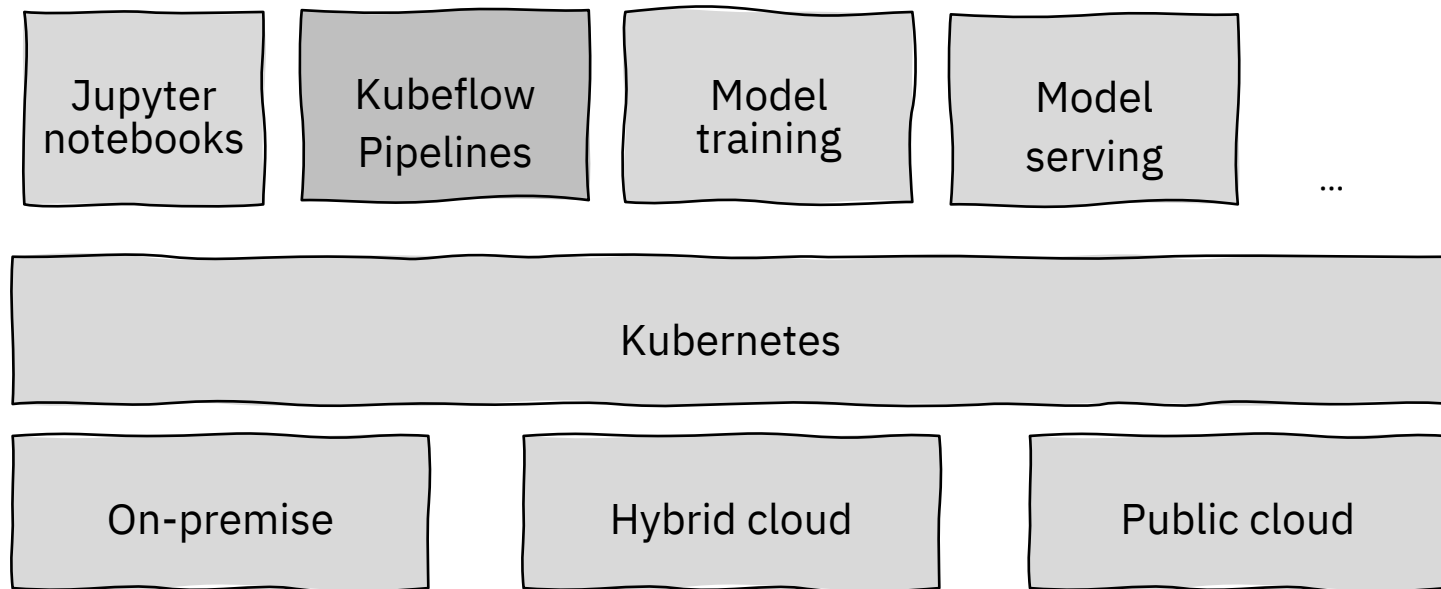
Implementing ML workflows using pipelines

- Modular notebooks (or Python scripts) allow for re-use in other projects
  - Example: load data from data source (database, cloud storage, …)

- Combine tasks in pipelines to automate execution
  - Example: periodically load/cleanse new data

Analyze data

Load data

Pre-process data

load_data.py

Part 1 - Data Cle...

Part 2 - Data Ana...

Part 3 - Time Seri...

- Pipeline scope? Up to you!

Kubeflow in a Nutshell

- Scalable, portable, distributed machine learning platform that runs on Kubernetes

Jupyter notebooks	Kubeflow Pipelines	Model training	Model serving

...

Kubernetes

On-premise	Hybrid cloud	Public cloud

- More info: https://www.kubeflow.org/

Kubeflow Pipelines in a Nutshell

- Platform for building and deploying portable, scalable machine learning workflows

- SDK/ DSL Python is used to define pipeline artifacts

| Pipeline artifacts	→	Workflow specification	→	Pipeline	→	Pipeline run

define   compile   upload   run

- More info: https://www.kubeflow.org/docs/pipelines/overview/pipelines-overview/
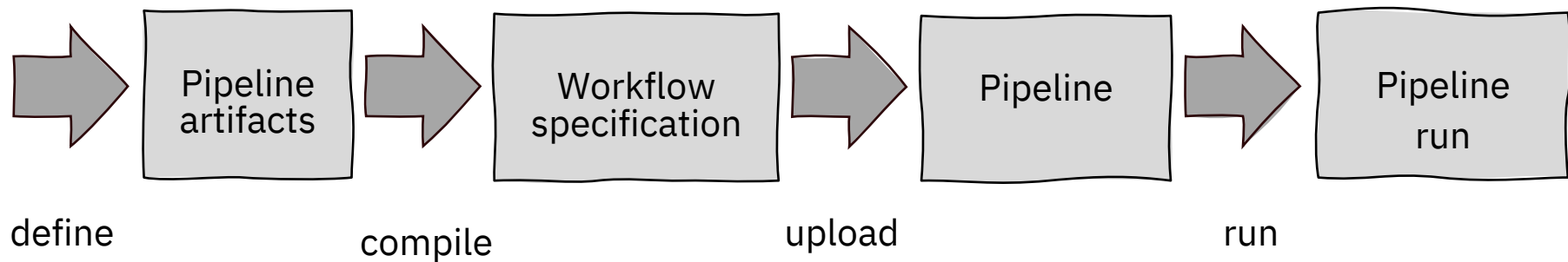
Building Pipelines with Elyra

- Use Visual Pipeline Editor to assemble pipelines from notebooks or Python scripts

Elyra
pipeline

Pipeline
run

assemble   run
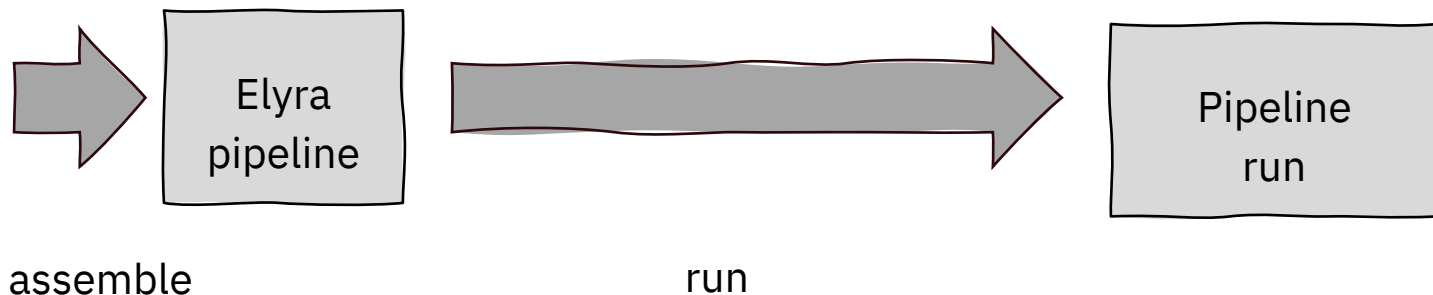
- Pipelines comprise of one or more [notebook/script] nodes

- Run pipelines
  - Locally in JupyterLab
  - On Kubeflow Pipelines (Elyra generates the required pipeline artifacts, uploads them, and starts a run

Demo: Implementing an ML workflow using Elyra
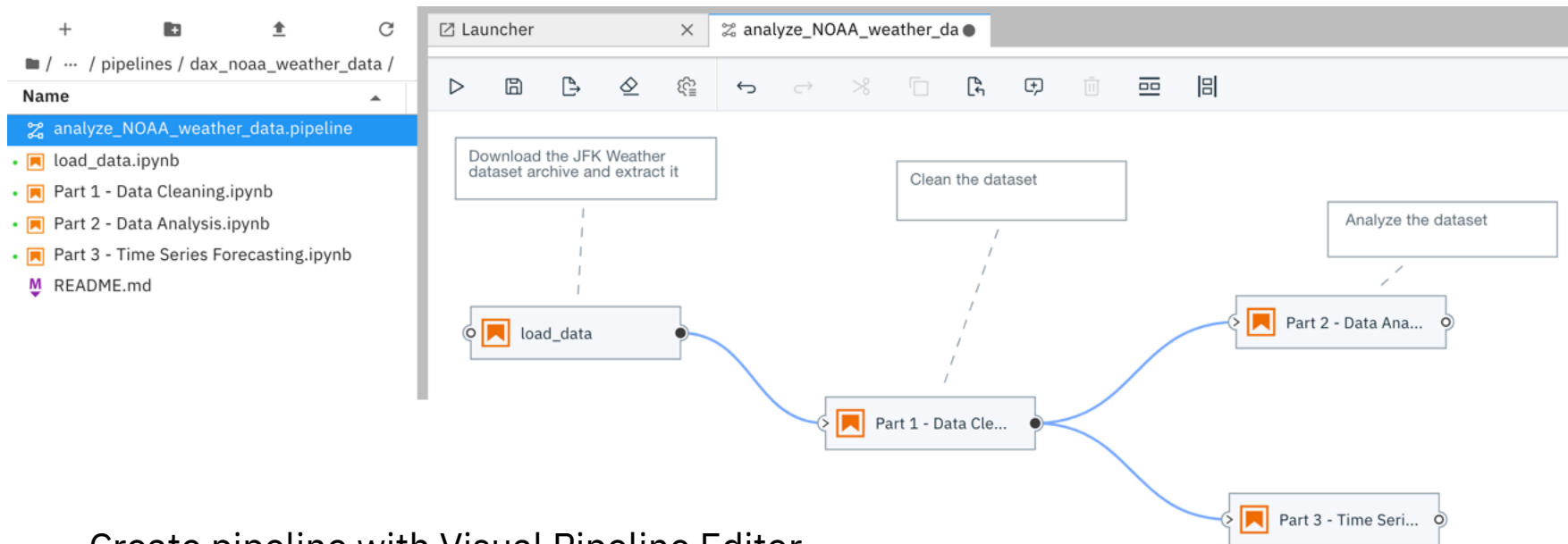
- Create pipeline with Visual Pipeline Editor
- Run pipeline locally in JupyterLab
- Run pipeline on Kubeflow Pipelines
- Tutorials

Getting Started with Elyra

[Try Elyra on Binder](#)
- No installation required - hosted on public cloud
- Nothing is persisted

[Run Elyra in a Docker container](#)
- Ready-to run images: `latest`, `x.y.z`, and `dev`

[Install Elyra](#) (requires Node.js and Python 3)
- `pip`, conda recipe, or from source code

https://elyra.readthedocs.io/en/latest/getting_started/installation.html

Elyra Community, Next Steps, and Thank You!

- Elyra community
  - https://github.com/elyra-ai/elyra
  - Weekly community meetings
  - Reach out on gitter

- Additional pipelines
  - COVID-19 (https://github.com/CODAIT/covid-notebooks)
  - Airline delay analysis (coming soon)
  - AI fairness analysis (coming soon)

- Contacts
  - http://codait.org, @codait_org
  - Patrick Titzler, @ptitzler, ptitzler@us.ibm.com
  - Upkar Lidder, @lidderupk, ulidder@us.ibm.com