



PYTORCH



Jyothi Nookula

Product Manager

Facebook AI



@FIBBONACI



AGENDA

01

PYTORCH

02

ECOSYSTEM



PYTORCH

An open source machine learning framework that accelerates the path from research prototyping to production deployment.

MACHINE LEARNING FRAMEWORK

- Primitives, activation and loss functions, optimizers

RESEARCH PROTOTYPING

- Models are Python code, Autograd and 'eager' mode for dynamic model architectures

PRODUCTION DEPLOYMENT

- TorchScript, TorchServe, quantization

PyTorch Open Source Community

- 1700+ contributors (+50% YoY)
- 37k users in discussion forums
- 57k+ GitHub downstream projects

PyTorch for Research

- Cited in 70% of NeurIPS'19 framework papers
- +127% YoY increase on Arxiv papers (Jun'19 to Jun'20)

Source: GitHub

Source: <http://horace.io/pytorch-vs-tensorflow>

Source: arxiv.org



The PyTorch Community

PYTORCH IN PRODUCTION

PyTorch in Production

- AstraZeneca, Disney, Datarock, Blue River Technology, Uber ATG, and more.

PyTorch in Production at Facebook

- Over 93% of models are trained on PyTorch
- Over 1700 inference models in production
- 600+ active users of PyTorch



Production
Usage

WHO MAKES PYTORCH: THE FULL TIMERS (60%)

Nvidia	Intel	Quansight
Microsoft	AMD	IBM
Preferred Networks	Google	Amazon

WHO MAKES PYTORCH: EVERYBODY ELSE (40%)

Operators	Python/C++ API parity	ONNX
Data Loading	Complex Numbers	VSX support
ROCm	Linear Algebra	Many, many more

WHO MAKES PYTORCH: EVERYBODY ELSE (40%)

Kshiteej Kalambarkar
and Shawn Zhong
(operators)

meganset
and Vinod Kumar
(Python/C++ API parity)

Ksenija Stanojevic
(ONNX)

Simon Wang
(data loading and
many misc bug fixes)

13%

Nylar P. Palko
(complex numbers)

quickwritereader
(VSX support)

of PRs from people who have
only contributed once this year!

Thomas Viemann
(ROCm)

Vishwak Srinivasan
(linear algebra)

Many, many more
@jeongukjae, @cloudhan, @cyyever,
@Enealor, @RockingJavaBean,
@Flamefire, @choidongyeon,
@muthuArivoli, @bpstark, @Kiyosora,
@gaurav1086, @cdluminare, ...



THE SCOPE OF PYTORCH ECOSYSTEM

AdverTorch 0.7.0

A toolbox for adversarial robustness research. It contains modules for generating adversarial examples and defending against attacks.

Abumentations 0.4.0

Fast and extensible image augmentation library for different CV tasks like classification, segmentation, object detection and pose estimation.

ELF 0.1.0

ELF is a platform for game research that allows developers to train and test their algorithms in various game environments.

fastai 0.14.0

fastai is a library that simplifies training fast and accurate neural nets using modern best practices.

Allegro Trains 0.1.0

Allegro Trains is a full-system ML/DL experiment manager, versioning and ML-ops solution.

AllenNLP 0.9.0

AllenNLP is an open-source research library built on PyTorch for designing and evaluating deep learning models for NLP.

Flair 0.10

Flair is a very simple framework for state-of-the-art natural language processing (NLP).

Glow 0.2.0

Glow is a ML compiler that accelerates the performance of deep learning frameworks on different hardware platforms.

baal 0.207

baal (Bayesian active learning) aims to implement active learning using metrics of uncertainty derived from approximations of Bayesian posteriors in neural networks.

BoTorch 0.1.0

BoTorch is a library for Bayesian Optimization. It provides a modular, extensible interface for composing Bayesian optimization primitives.

GPyTorch 0.2.0

GPyTorch is a Gaussian process library implemented using PyTorch, designed for creating scalable, flexible Gaussian process models.

higher 0.107

higher is a library which facilitates the implementation of arbitrarily complex gradient-based meta-learning algorithms and nested optimization loops with near-vanilla PyTorch.

Captum 0.1.0

Captum ("comprehension" @ LADR) is an open source, extensible library for model interpretability built on PyTorch.

Catalyst 0.2.0

Catalyst helps you write compact, but full-featured deep learning and reinforcement learning pipelines with a few lines of code.

Horovod 0.10.2

Horovod is a distributed training library for deep learning frameworks. Horovod aims to make distributed DL fast and easy to use.

Hummingbird 0.1.0

Hummingbird compiles trained ML models into tensor computation for faster inference.

CrypTen 0.2.0

CrypTen is a framework for Privacy Preserving ML. Its goal is to make secure computing techniques accessible to ML practitioners.

Detectron2 0.1.0

Detectron2 is FAIR's next-generation platform for object detection and segmentation.

Hydra 0.10

A framework for elegantly configuring complex applications.

ignite 0.1.0

ignite is a high-level library for training neural networks in PyTorch. It helps with writing compact, but full-featured training loops.

Determined 0.9.0

Determined is a platform that helps deep learning teams train models more quickly, easily share GPU resources, and effectively collaborate.

DGL 0.5.0

Deep Graph Library (DGL) is a Python package built for easy implementation of graph neural network model family, on top of PyTorch and other frameworks.

Kornia 0.3.0

Kornia is a differentiable computer vision library that consists of a set of routines and differentiable modules to solve generic CV problems.

MMF 0.3.0

A modular framework for vision & language multimodal research from Facebook AI Research (FAIR).

MONAI 0.4.0

MONAI provides domain-optimized foundational capabilities for developing healthcare imaging training workflows.

NeMo 0.2.0

NeMo is a toolkit for conversational AI.

PyTorch-NLP 0.1.0

Basic Utilities for PyTorch Natural Language Processing (NLP).

PyTorch-Points3d 0.7.0

A PyTorch framework for deep learning on point clouds.

Optuna 0.1.0

An open source hyperparameter optimization framework to automate hyperparameter search.

Parti 0.1.0

Parti is a unified platform for storing, training, and evaluating distill models across many tasks.

PyTorch3D 0.1.0

PyTorch3D provides efficient, reusable components for 3D-Computer Vision research with PyTorch.

Ray 0.13.0

Ray is a fast and simple framework for building and running distributed applications.

PennyLane 0

PennyLane is a library for quantum ML, automatic differentiation, and optimization of hybrid quantum-classical computations.

PFRL 0.2.0

PFRL is a deep reinforcement learning library that implements various state-of-the-art deep reinforcement algorithms in Python using PyTorch.

skorch 0.1.0

skorch is a high-level library for PyTorch that provides full scikit-learn compatibility.

Tensorly 0.10

Tensorly is a high-level API for tensor methods and deep tensorized neural networks in Python that aims to make tensor learning simple.

Polyaxon 0.2.0

Polyaxon is a platform for building, training, and monitoring large-scale deep learning applications.

Proton 0.1.0

Proton is a keras-like framework for PyTorch and handles much of the boilerplate code needed to train neural networks.

Translate 0.47

Translate is an open source project based on Facebook's machine translation systems.

Pyro 0.4.0

Pyro is a universal probabilistic programming language (PPL) written in Python and supported by PyTorch on the backend.

PySyft 0.4.0

PySyft is a Python library for encrypted, privacy-preserving deep learning.

PyTorch Geometric 0.3.0

PyTorch Geometric is a library for deep learning on irregular input data such as graphs, point clouds, and manifolds.

PyTorch Lightning 0.10

PyTorch Lightning is a Keras-like ML library for PyTorch. It leaves core training and validation logic to you and automates the rest.

WHAT DROVE ADOPTION?

1. Usability



Andrej Karpathy  @karpathy · May 26, 2017

I've been using PyTorch a few months now and I've never felt better. I have more energy. My skin is clearer. My eye sight has improved.

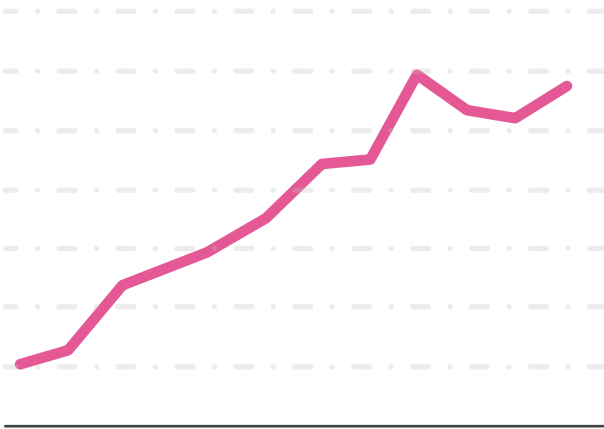
 35

 489

 1.7K

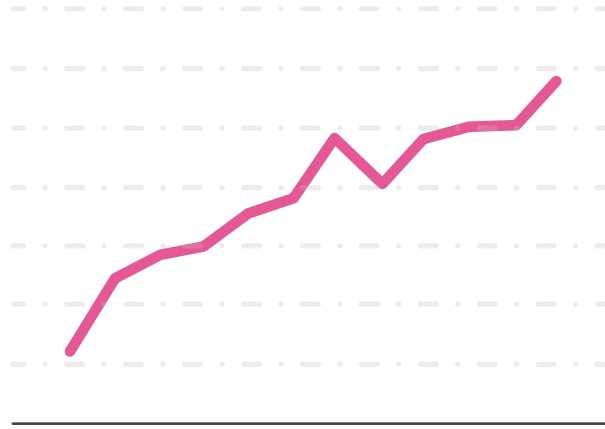


2. Production Usage



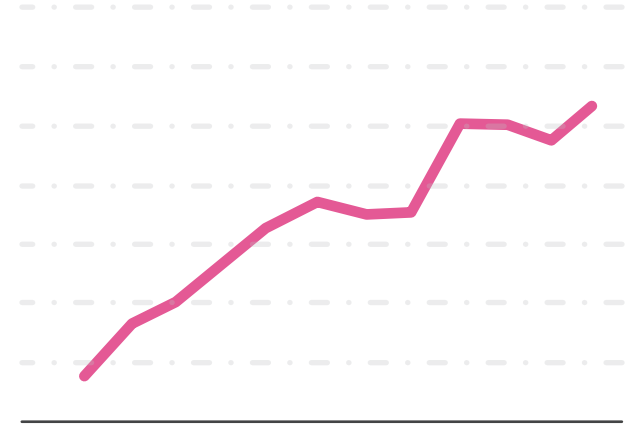
2X

EMPLOYEES TRAINING
MODELS



5X

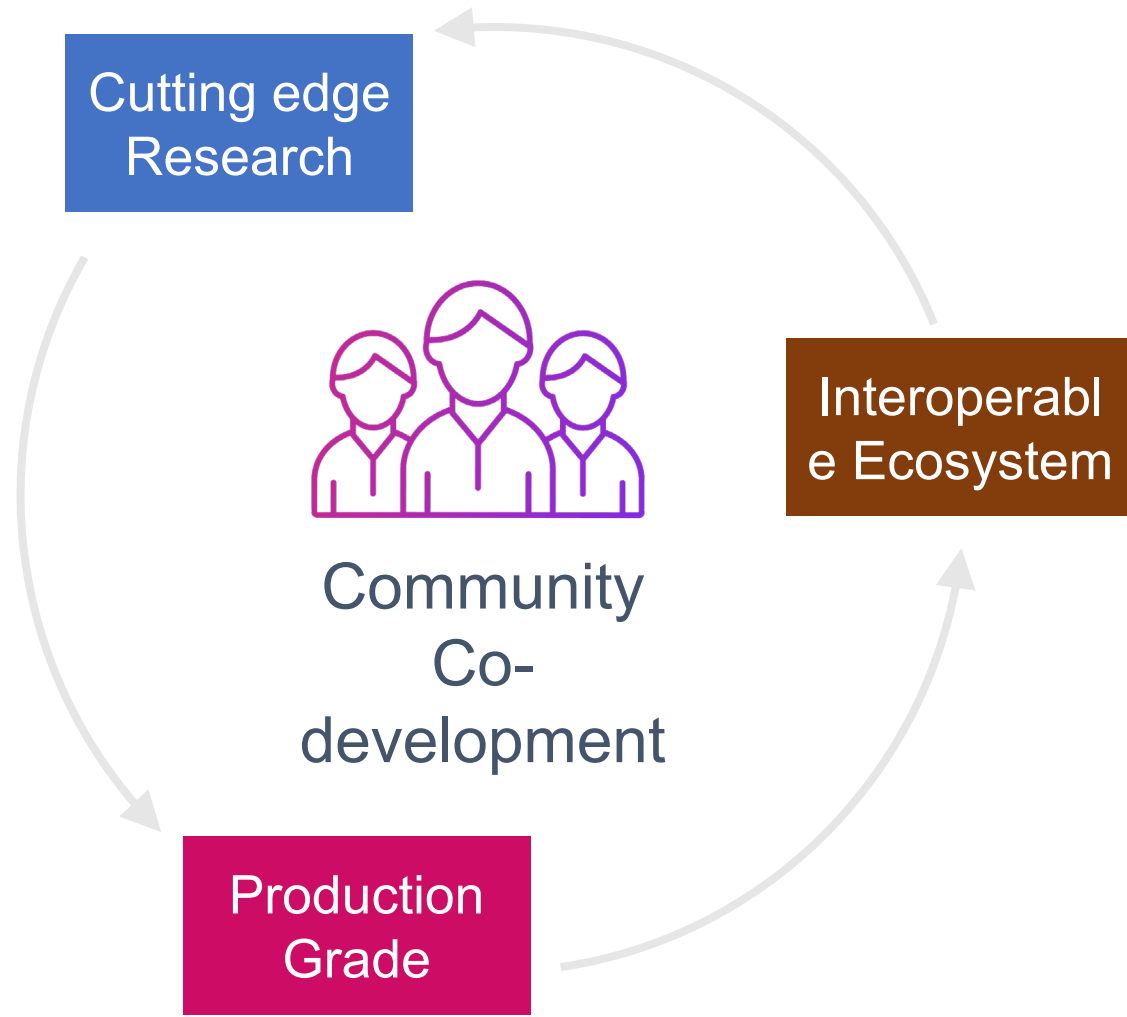
MODELS TRAINED



8X

COMPUTATIONAL
RESOURCES

INTERNAL FB TRENDS OVER 18 MONTHS



Flexibility for
fast ideation
and
breakthrough

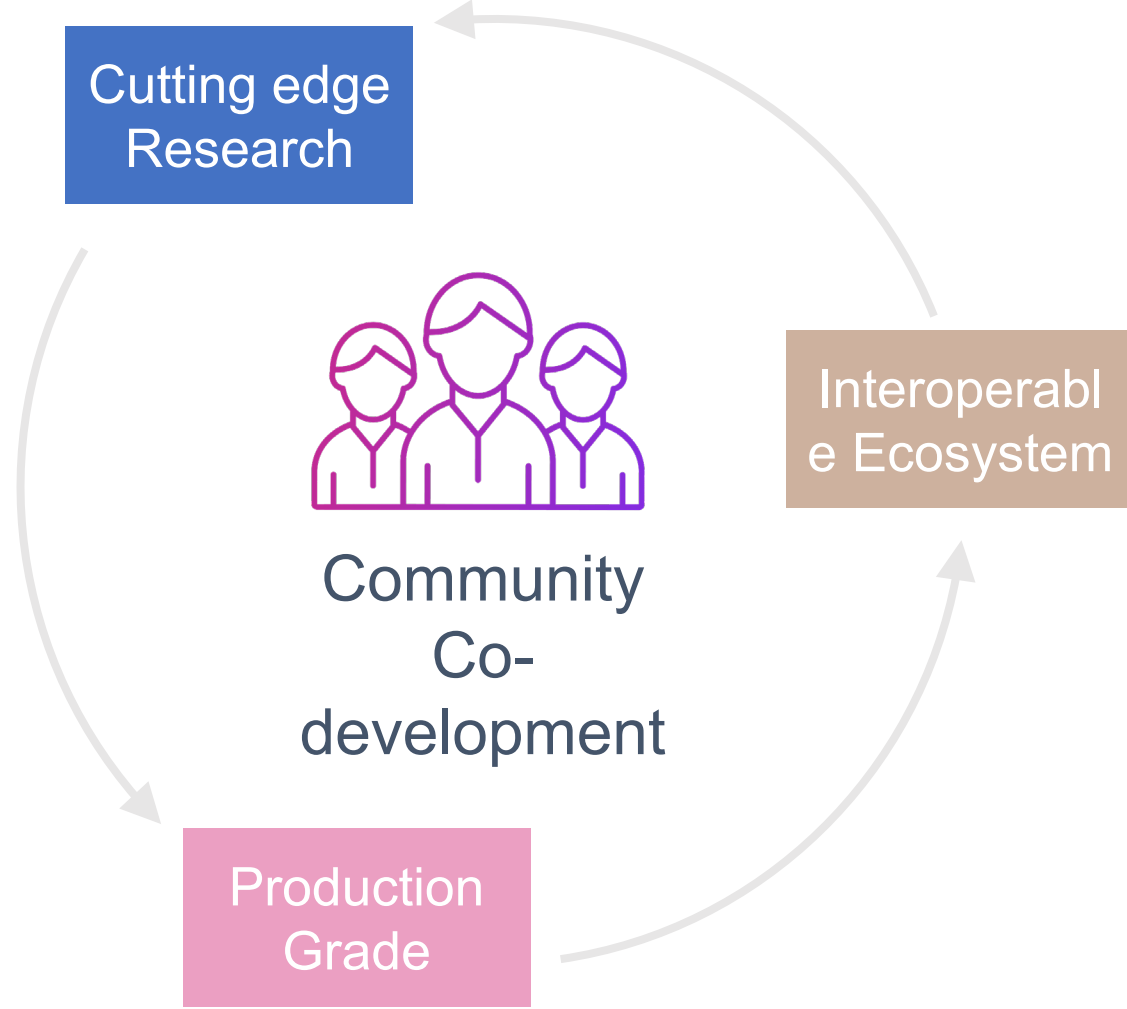
Cutting edge
Research



Community
Co-
development

Interoperabl
e Ecosystem

Production
Grade



Flexibility for
fast ideation
and
breakthrough

Cutting edge
Research

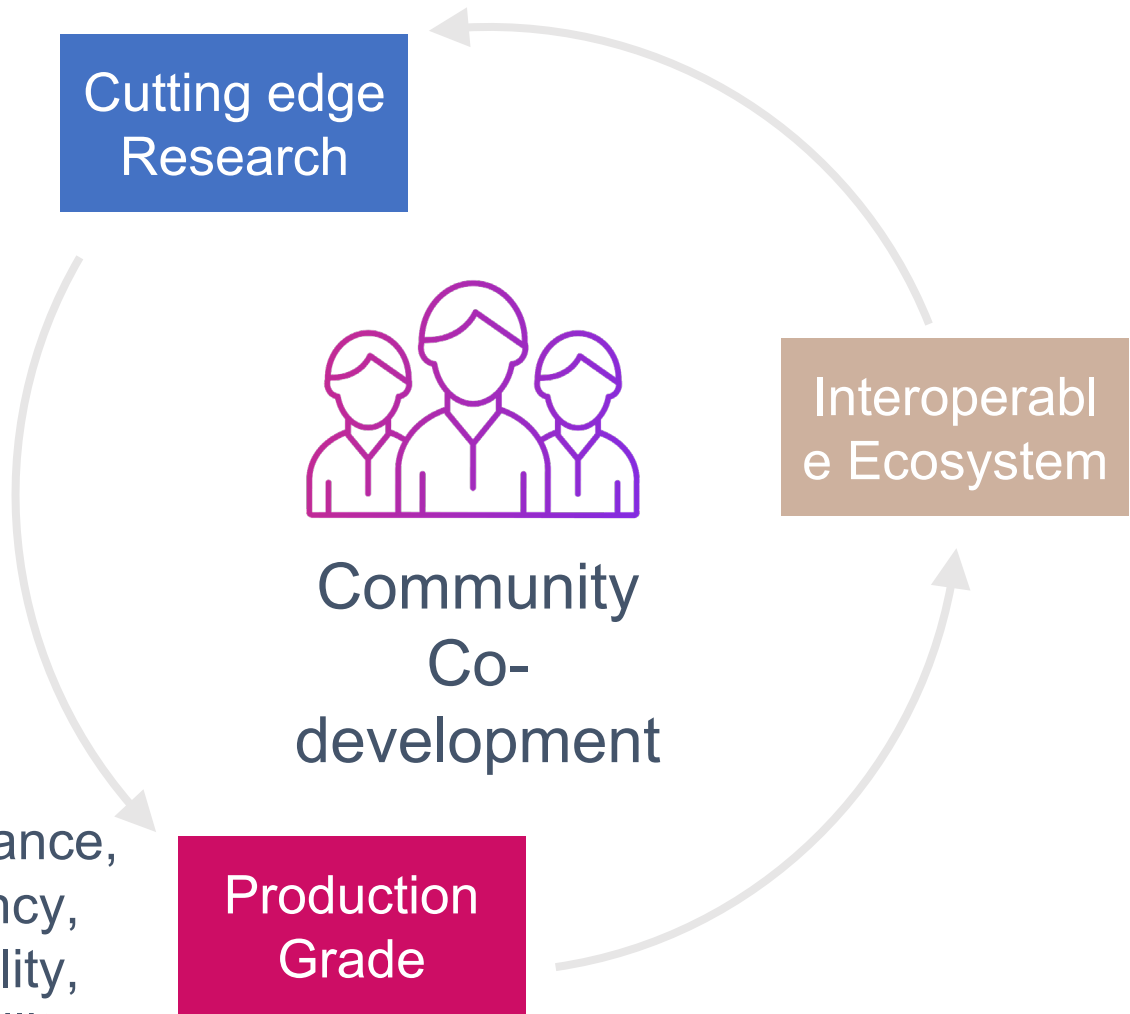


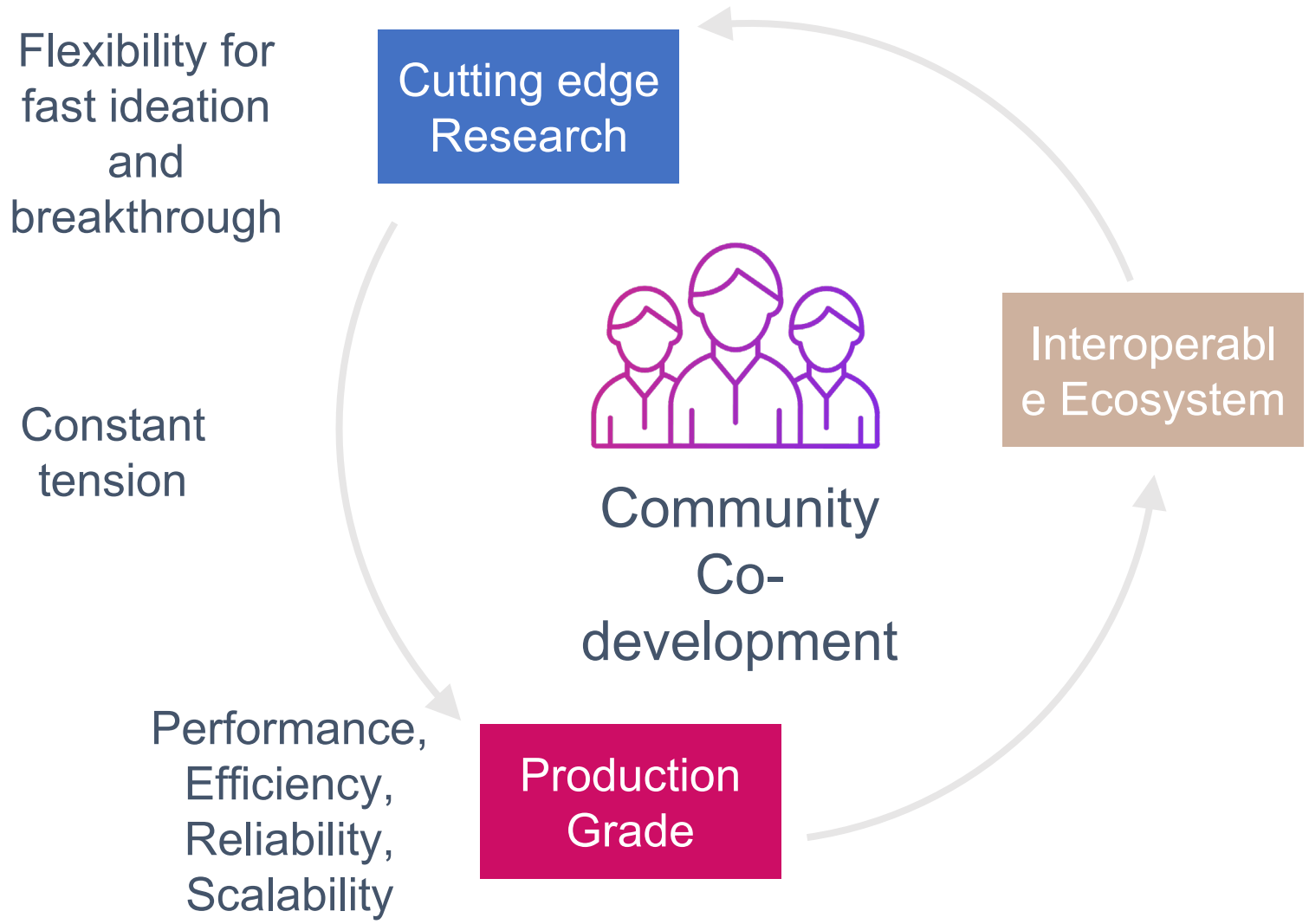
Community
Co-
development

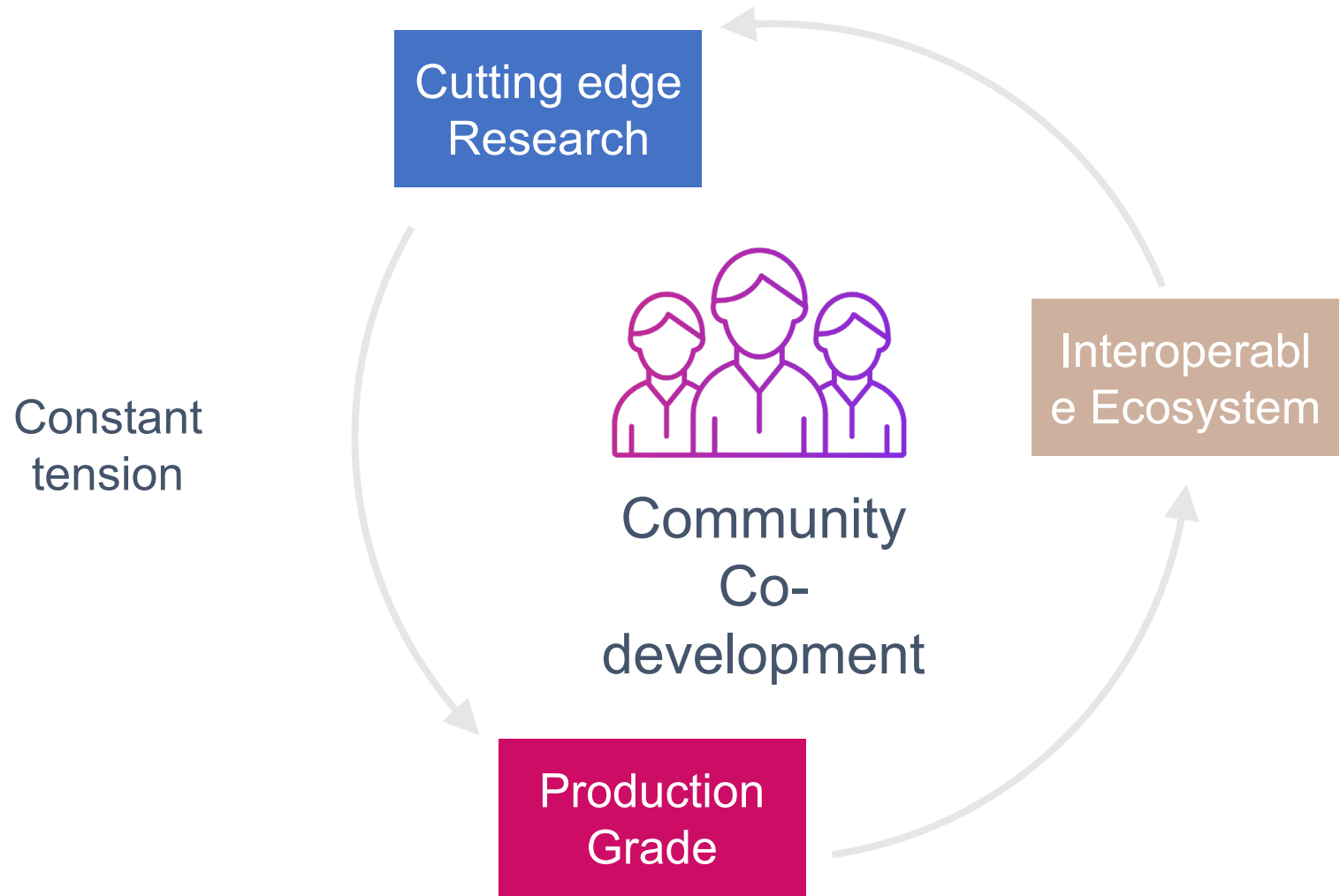
Interoperabl
e Ecosystem

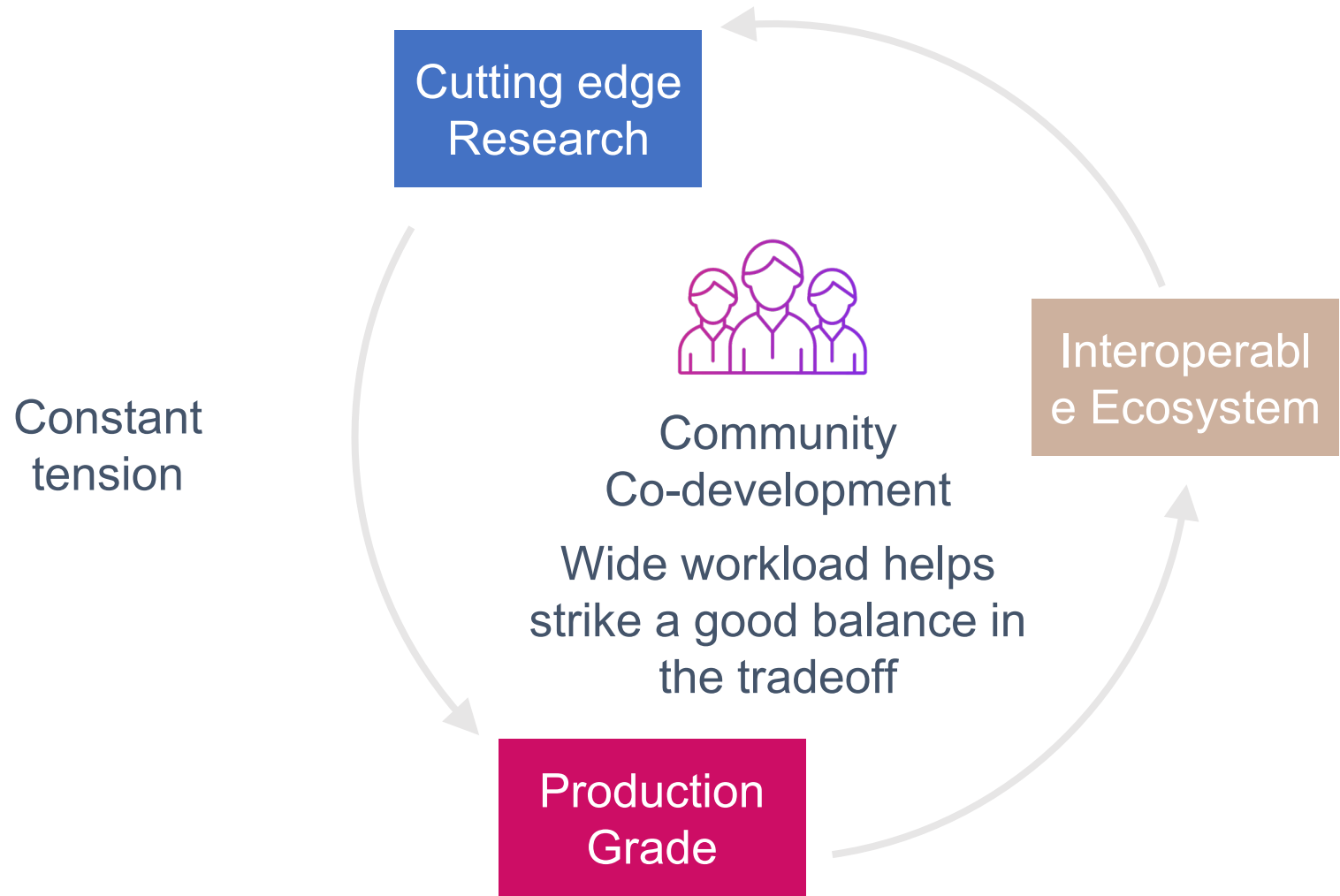
Performance,
Efficiency,
Reliability,
Scalability

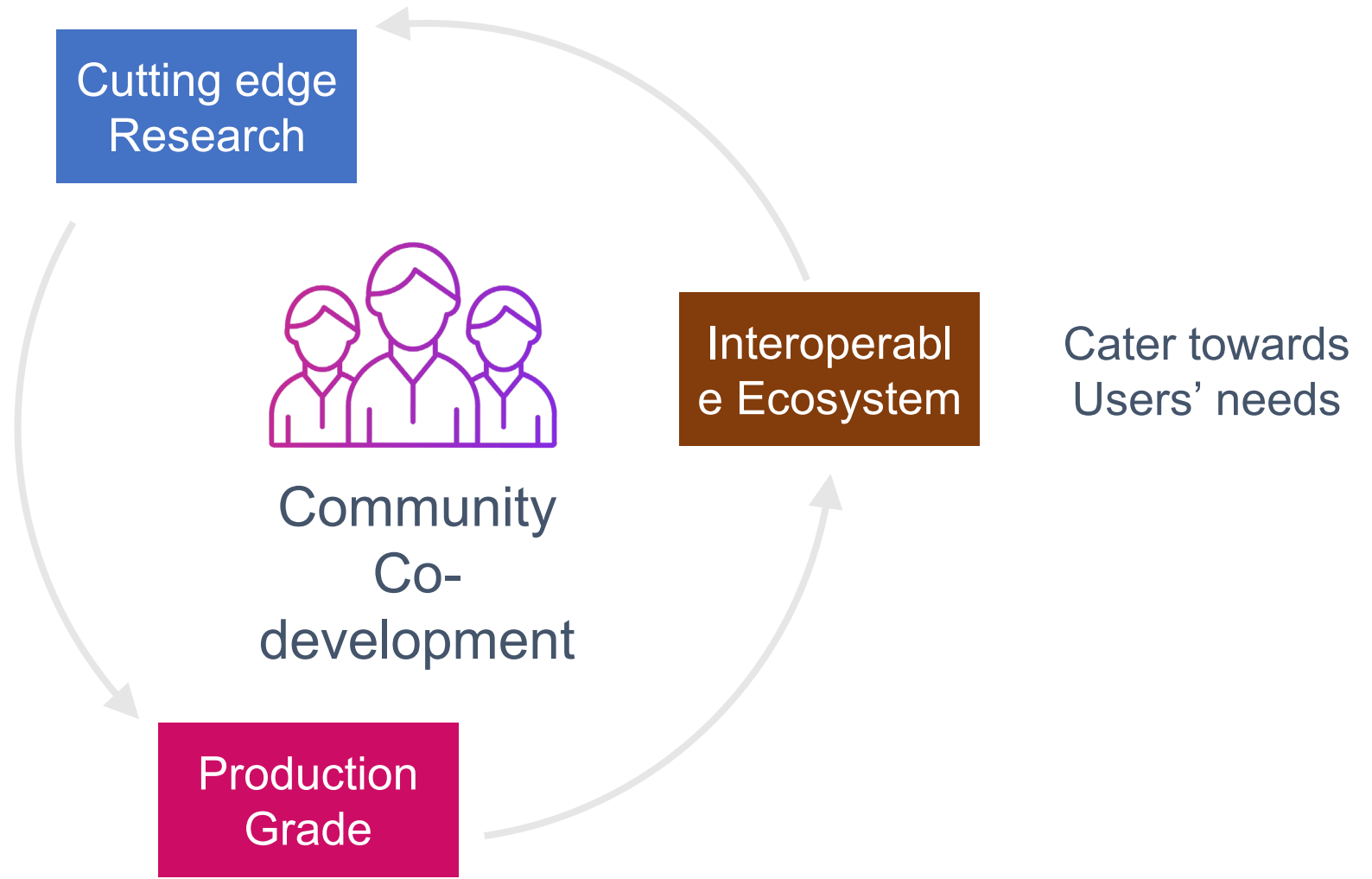
Production
Grade











Cutting edge
Research

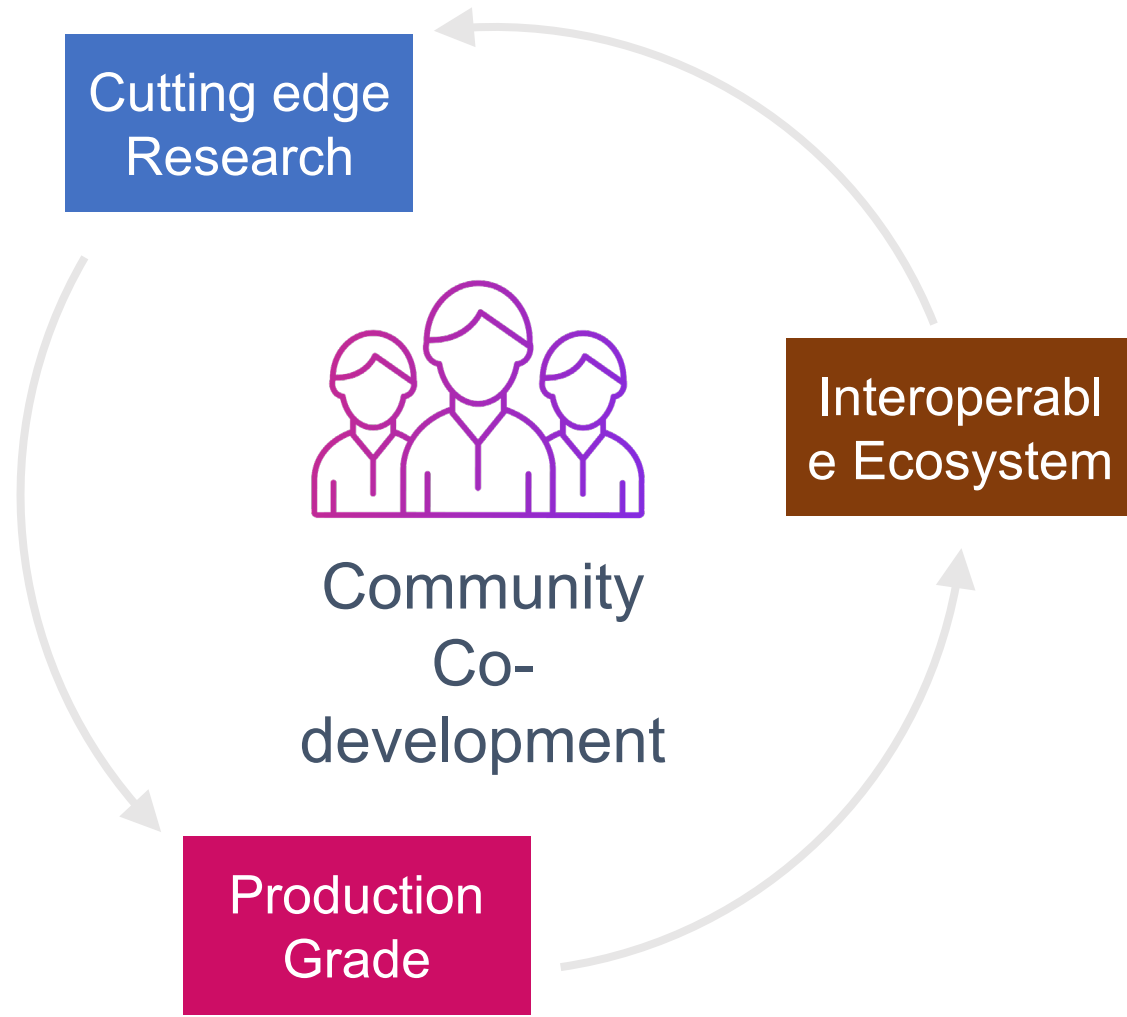


Community
Co-
development

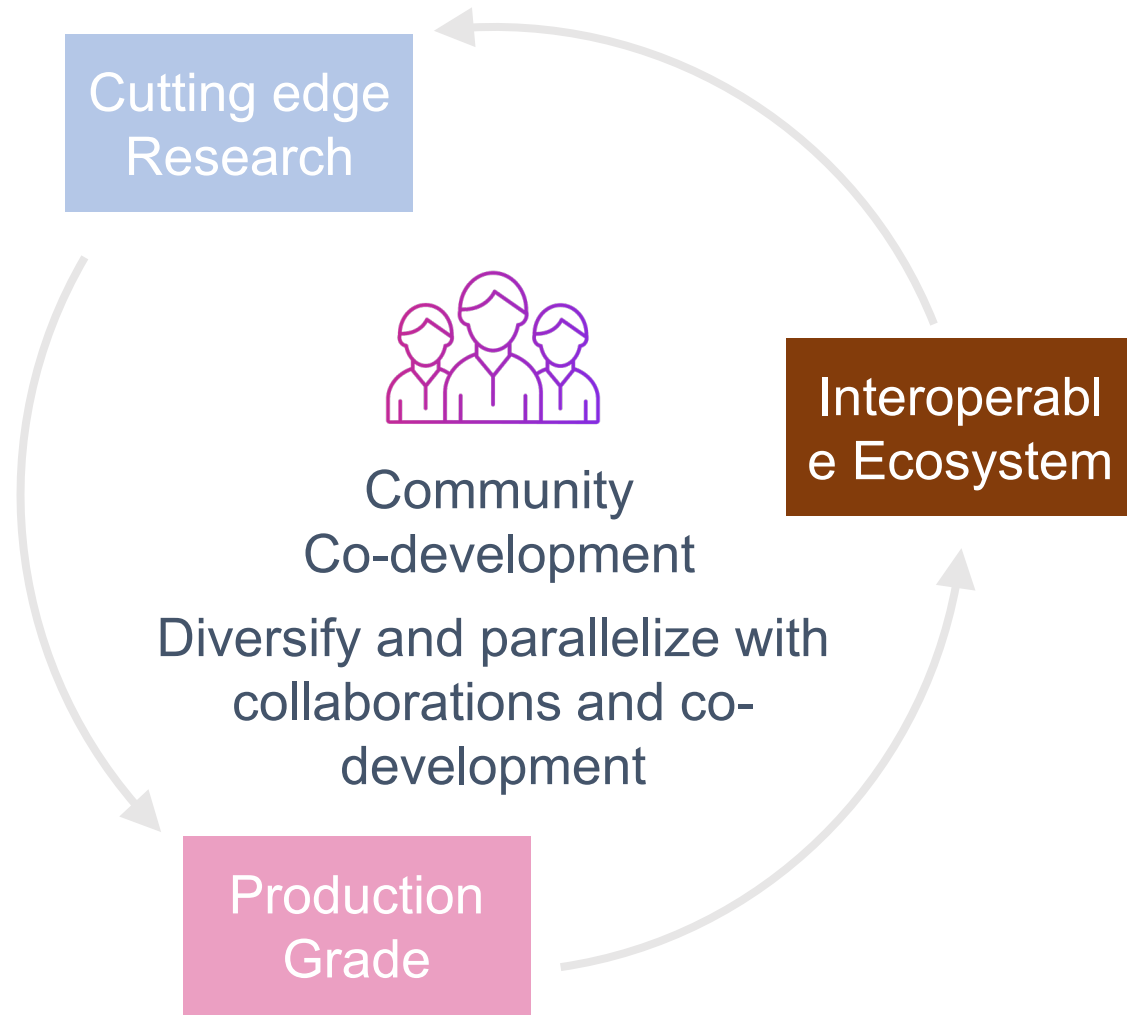
Interoperabl
e Ecosystem

Cater towards
Users' needs

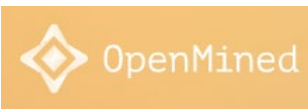
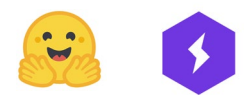
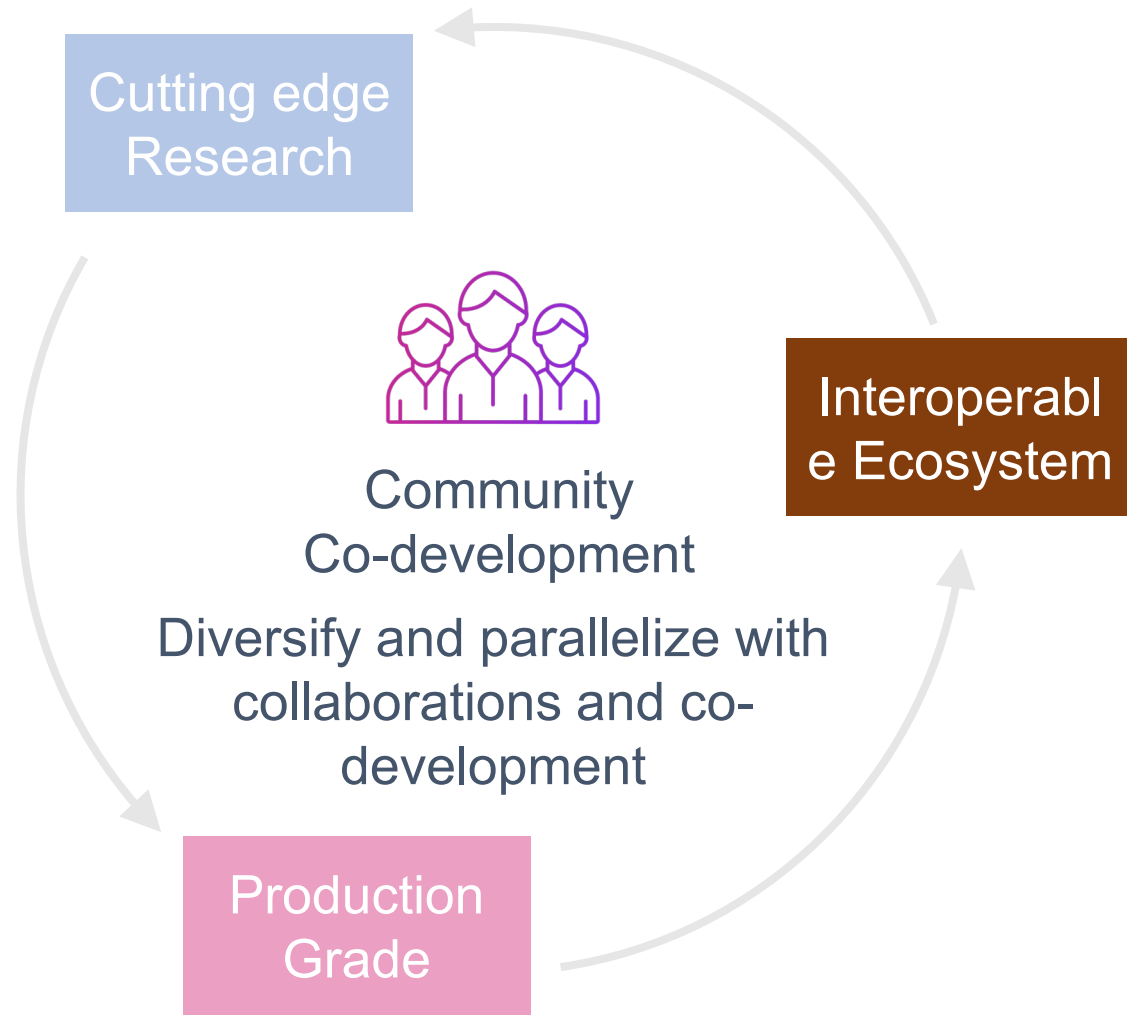
Production
Grade

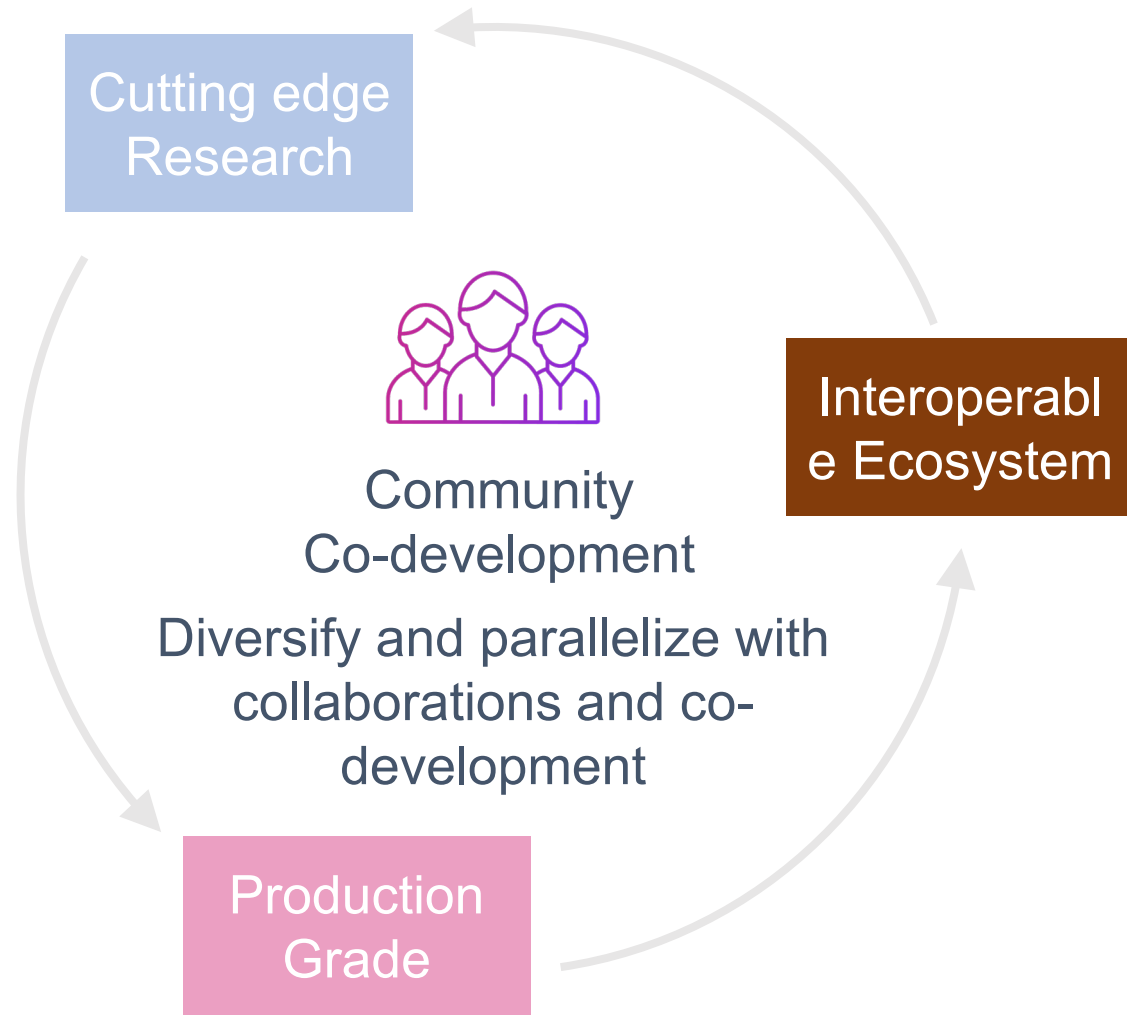


We can't enable everything ourselves



We can't enable everything ourselves







THANK YOU