# Responsible AI, FEAT & Veritas

Dr. Li Xuchun
AI Development Office,
FinTech & Innovation Group, MAS

# Agenda

1. Responsible Use of AI and DA, FEAT
2. Veritas Background
3. Veritas Phase One
4. Veritas Phase Two Plan

# 1. Responsible use of AI and DA, FEAT

# Why are we concerned with the responsible use of AI and DA?

What all of us have to do is to make sure
we are using AI in a way that is
for the benefit of humanity,
not to the detriment of humanity.”

- *Tim Cook, CEO, Apple*

# FEAT Principles

## Fairness

With data, there are infinite possibilities of what we **can** do. But, ***should*** we?

## Ethics

Are there are certain **lines** that should not be crossed?

## Accountability

Do we understand how data-driven decisions are made? Who bears the **responsibility** and **risks** of these decisions?

## Transparency

To what extent should **explanations** be given for data-driven decisions?

# Examples of the FEAT Principles

## Fairness

- Ensure that individuals or groups are not systemically disadvantaged through AIDA-driven decisions unless these decisions can be justified
- Regular review of models and AIDA-driven decisions to minimize bias and ensure models

## Ethics

- Use of AIDA is aligned with firm's ethical standards, values and codes of conduct
- AIDA-driven decisions held to same ethical standards as human-driven decisions

## Accountability

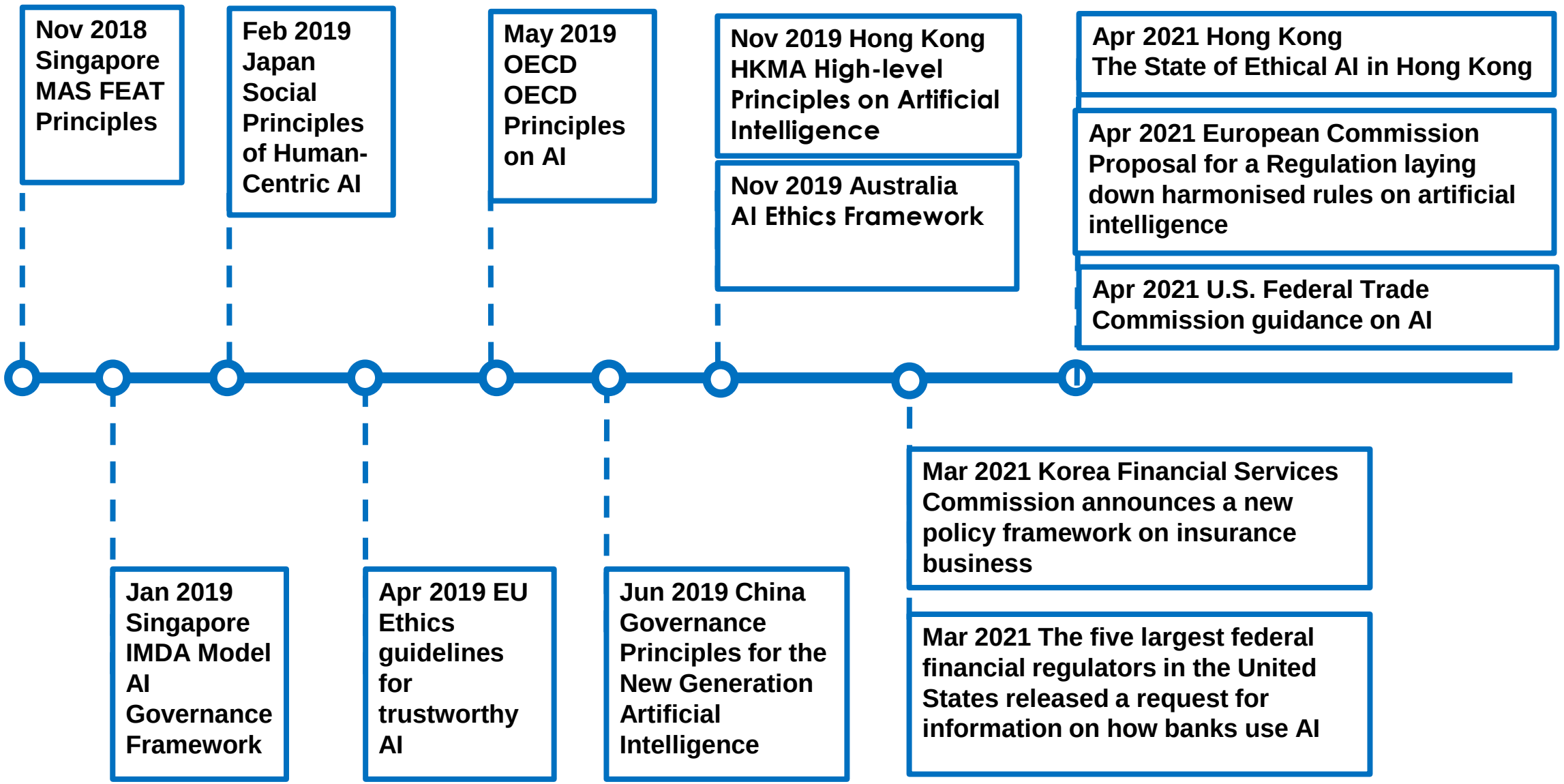
- Internal accountability to the firm's stakeholders (management, Board)
- External accountability to data subjects

## Transparency

- Disclosure of AIDA use to data subjects
- Provision of clear explanations on data used to make AIDA-driven decisions and how it impacts them
- Provision of clear explanations of how AIDA-driven decisions impact data subjects

# 2. Veritas Background

# Global AI Governance Landscape



# Veritas

## Challenges

Practical implementation of FEAT principles: Fairness, Ethics, Accountability and Transparency.

## Objective

Creating a standardized and modular implementation **methodology** of FEAT principles and programming **open source code**

## Consortium

Formation of industry consortium to develop Veritas to validate AI models against FEAT

## Impact

inclusive financial services by using AI and Machine Learning in a responsible way

# 3. Veritas Phase One

# Scope of Veritas Phase One in 2020

**Veritas was announced at 2019 Singapore FinTech Festival as part of Singapore National AI Strategy by DPM.**

**Veritas phase one: focus on fairness principles of FEAT.**

**Veritas phase one: two banking use cases were selected which are credit risk scoring and customer marketing.**

# Phase One Deliverables : Whitepapers and Python Open Source Code

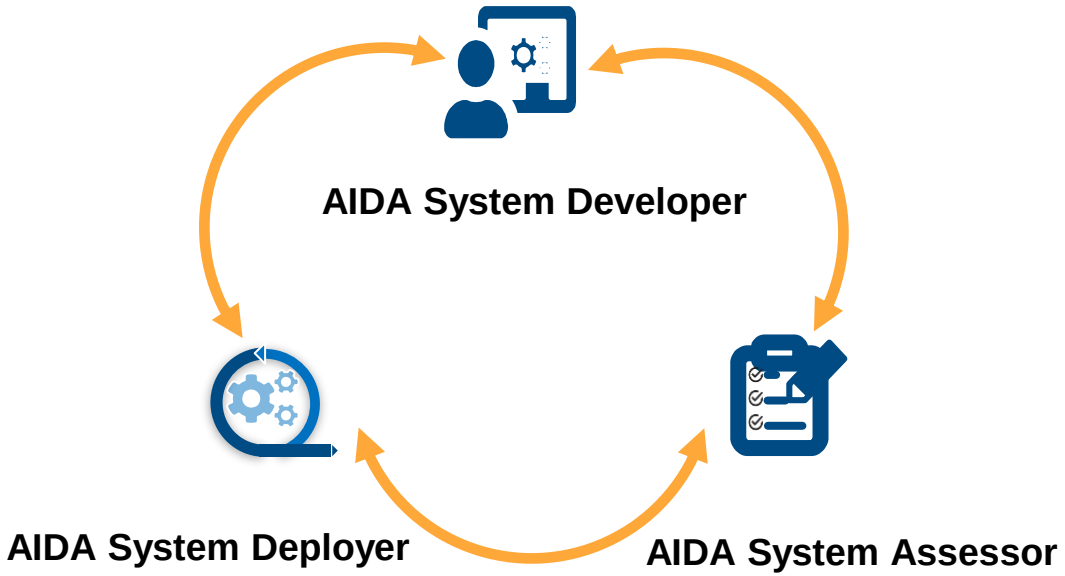
**Whitepaper to document the generic methodology to assess the fairness for AIDA system.**

**Whitepaper of use case studies for credit risk scoring and customer marketing.**

**Python code for Credit Risk use case**

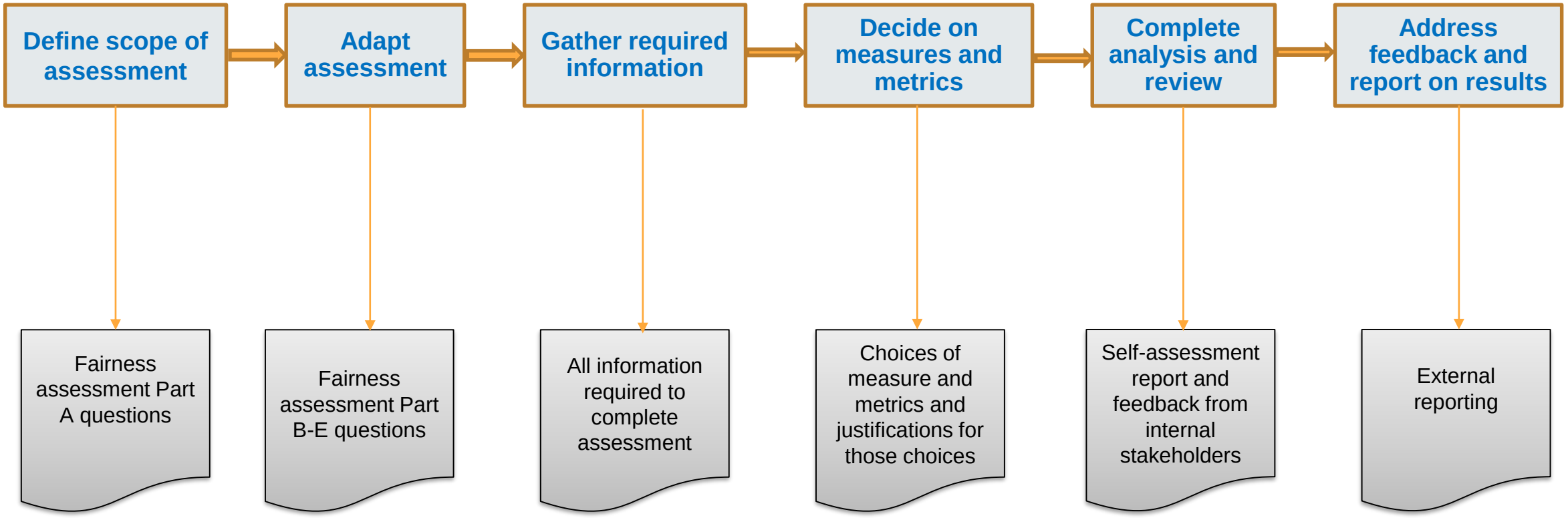
**Python code for Customer Marketing use case**

# A Comprehensive Five-part Methodology to Assess Alignment with FEAT Fairness Principles




- A Describe system objectives and context
- B Examine data and models for unintended bias
- C Measure disadvantage
- D Justify the use of personal attributes
- E Examine system monitoring

# Veritas Fairness Assessment Flow



# Veritas Programming Open Source Code

Search or jump to… / Pulls Issues Marketplace Explore

Overview | Repositories 1 | Projects | Packages

VERITAS — TRUST IN AI —

Find a repository… | Type: All | Language: All | New

**phase1** Private

Jupyter Notebook Updated 3 hours ago | Star

veritas-p

Search or jump to… / Pull requests Issues Marketplace Explore

veritas-project / **phase1** Private | Unwatch

Code | Issues | Pull requests | Actions | Projects | Security | Insights | Settings

main | 1 branch | 0 tags | Go to file | Add file | Code

**mebrunet** Initial upload from EAI 932acc6 3 hours ago 5 commits

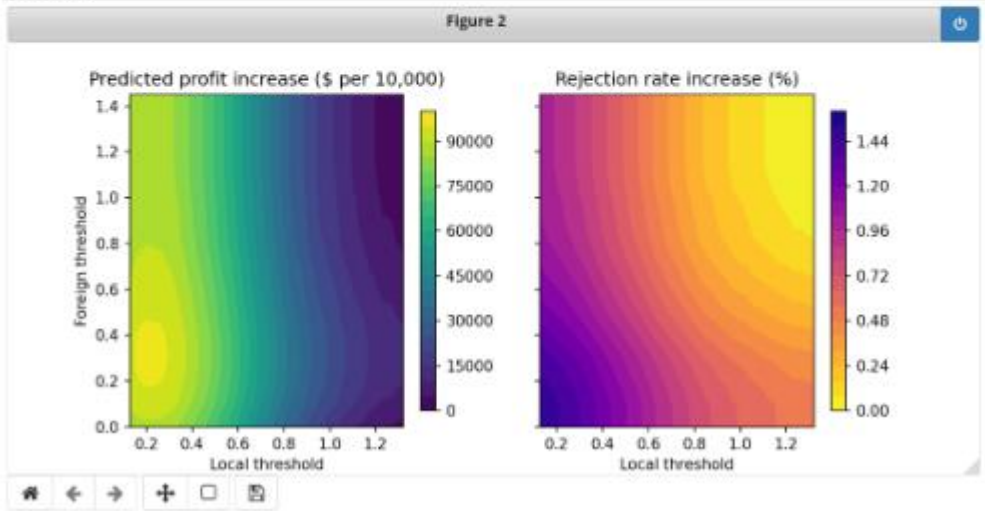
| | | |
|---|---|---|
| credit_scoring | Initial upload from EAI | 3 hours ago |
| customer_marketing | initial upload from gradient | 4 hours ago |
| .gitignore | initial upload from gradient | 4 hours ago |
| README.md | Update README.md | 4 hours ago |

```
In [16]: f, (ax1, ax2) = plt.subplots(1, 2, sharey=True, figsize=(9, 4))
         f.subplots_adjust(hspace=0.3)

         pcm1 = ax1.contourf(FT, LT, EPROFIT, levels=20)
         ax1.set_title("Predicted profit increase ($ per 10,000)")
         ax1.set_ylabel("Foreign threshold")
         ax1.set_xlabel("Local threshold")
         f.colorbar(pcm1, ax=ax1, shrink=0.9)
         ax1.axis("equal")

         pcm2 = ax2.contourf(FT, LT, ERHARM, levels=20, cmap="plasma_r")
         ax2.set_title("Rejection rate increase (%)")
         ax2.set_xlabel("Local threshold")
         f.colorbar(pcm2, ax=ax2, shrink=0.9)
         ax2.axis("equal")

         plt.show()
```



# 4. Veritas Phase Two

# Plan of the Veritas Phase Two in 2021

## SCOPE

| | Lead Members | Use Cases |
|---|---|---|
| **FAIRNESS** | Swiss Re, accenture | Insurance Underwriting |
| **ETHICS & ACCOUNTABILITY** | AXA, UOB, accenture | • Customer marketing (Banking)<br>• Fraud detection (Insurance) |
| **TRANSPARENCY** | truera, HSBC, Standard Chartered | • Credit risk scoring (Banking)<br>• Customer marketing (Banking) |

## DELIVERABLES

**01**

**CONSOLIDATED WHITEPAPER**

- **Technical Document 1** (Fairness)
- **Technical Document 2** (Ethics & Accountability)
- **Technical Document 3** (Transparency)

**02**

**TECHNICAL DELIVERABLES**

- **Source Code 1** (Fairness)
- **Source Code 2** (Transparency)

# Veritas Consortium Members in the Phase Two

## LEAD MEMBERS

**Drive the Veritas development of the use cases**

HSBC
Swiss Re
Standard Chartered
AXA
UOB
accenture
truera

## MEMBERS (FIS)

**Provide domain knowledge & feedback on the use cases**

ICBC 中国工商银行
DBS
citi
Goldman Sachs
BNP PARIBAS
The bank for a changing world
PING AN Finance · Technology
中国银行 BANK OF CHINA
MUFG
UnionBank
OCBC Bank
PRUDENTIAL
BNY MELLON
BANKOFSINGAPORE

## FACILITATORS

**Facilitate the whole program**

MAS
Monetary Authority of Singapore
accenture

## ADVISORY MEMBERS

**Provide expertise & review the whitepaper**

Microsoft
BAFT
CFA Society Singapore
aws
IBM

Thank you!