浦发银行
SPD BANK

GOPS 2022
shenzhen

浦发银行MLOps落地实践

个人简介

浦发银行 SPD BANK

GOPS 2022 shenzhen

郭林海

浦发银行总行信息科技部创新实验室
区块链&数字人团队负责人

近20年金融行业信息系统建设经验，牵头负责浦发银行区块链、人工智能、数字人民币等相关前沿创新技术应用研究探索工作。

浦发银行 SPD BANK

GOPS 2022 shenzhen

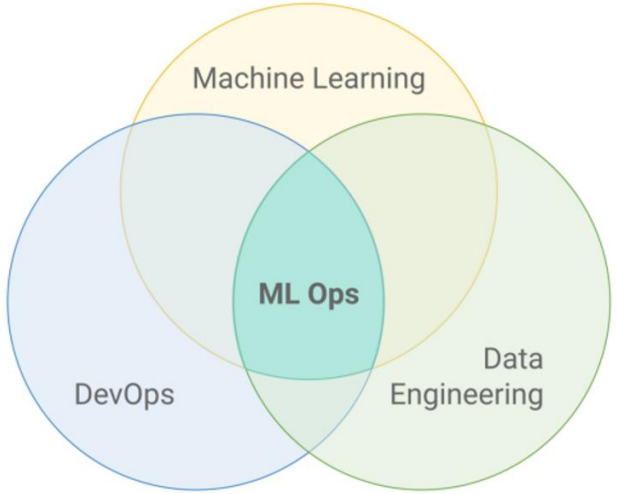
目录 Contents

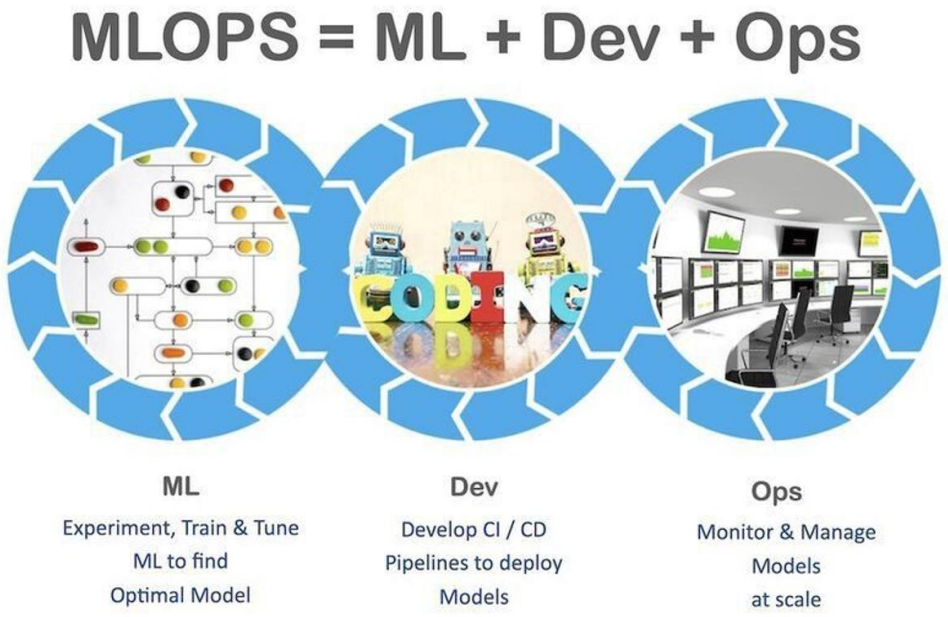
浦发银行
SPD BANK

01

我们对MLOps理解

MLOps：AI工程化的基石





MLOps优化了开发、部署和管理

- MLOps具有自动化的管道、流程和工具，可以简化模型构建的所有步骤。 通过持续的开发、测试、部署、监控和再序列，MLOps可以改善团队之间的协作，缩短开发生命周期，从而**实现更快、更可靠、更有效的模型部署、操作和维护。**

MLOps可以提升试验和交付速率，帮助企业实现机器学习的工程化

- MLOps帮助组织监控模型性能和管理模型漂移的预测误差，**通过借助标准化流程来保持AI模型与不断变化的业务的一致性。**

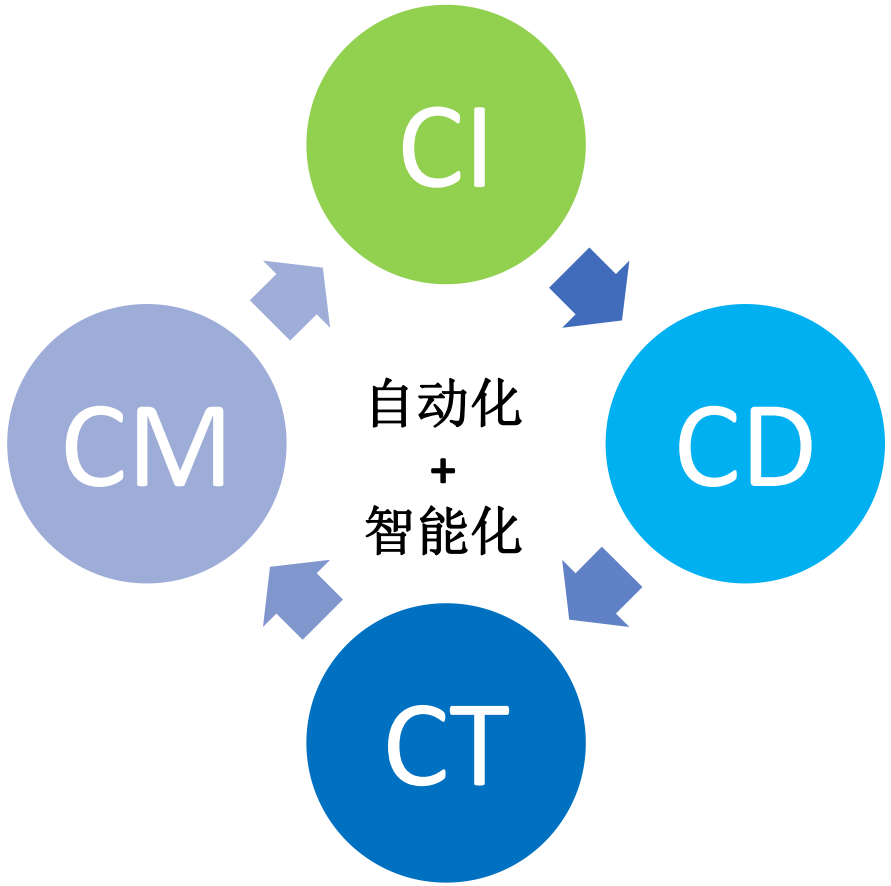
MLOps让专业开发可以横向扩展

- MLOps通过开发团队和生产团队之间的深入协作让更大更专业的团队一起以标准化的方式进行有效的工作；通过有效的工具如autoML，加速数据科学家模型开发的速度，**使得端到端的模型开发变得可扩展、更高效、更快捷。**

MLOps有助于解决与数据使用相关的新挑战

- MLOps工具可以自动记录和存储关于数据如何使用、何时部署和重新校准模型、由谁进行以及为什么进行更改的信息；**并帮助企业解决在软件开发中通常不会遇到的问责、透明度、监管、合规以及AI伦理问题。**

MLOps：自动化和智能化的结合



数据工程

自动化数据分析和探索，**节省60%-80%**数据分析人力成本

自动化特征工程与特征库，**节省80%**人力成本

成熟完备的数据治理方法和规范

模型开发和迭代

高性能AI加速引擎，模型训练**提速500%-800%**

通过AutoML进行智能选择网络和参数，**节省90%**参数调优

完备的实验管理机制，加速模型迭代

持续交付

自动打包，一键发布

数据/代码/模型任一维度的变化，将自动触发持续交付的流程

模型监控

自动化模型效果管控以及自动重训练

整体提效

CI/CD/CT/CM全流程自动化

模型迭代时长：**1~2个月->一周**

全流程可追溯，提升上线安全性

MLOps七个基本原则

浦发银行 SPD BANK

GOPS 2022 shenzhen

MLOps实践要点

涉及整个MLOps流程设计全生命周期所有环节，如数据管理，模型训练，服务运维等

- 流程/制度/岗位管理等

规范化

- 数据/代码/配置/模型/超参/流水线

版本管理

- 模型训练/特征探查/数据标注/资源调度

智能化

- 全流程自动化 CI/CD/CT/CM

自动化

- 训练数据/特征/环境/代码

可复现

- 云端部署/边端云部署

易部署

- 数据漂移/模型指标/服务性能

可监控

浦发银行
SPD BANK

02

MLOps在浦发的应用和实践

深度学习平台建设原则

浦发银行 SPD BANK

GOPS 2022 shenzhen

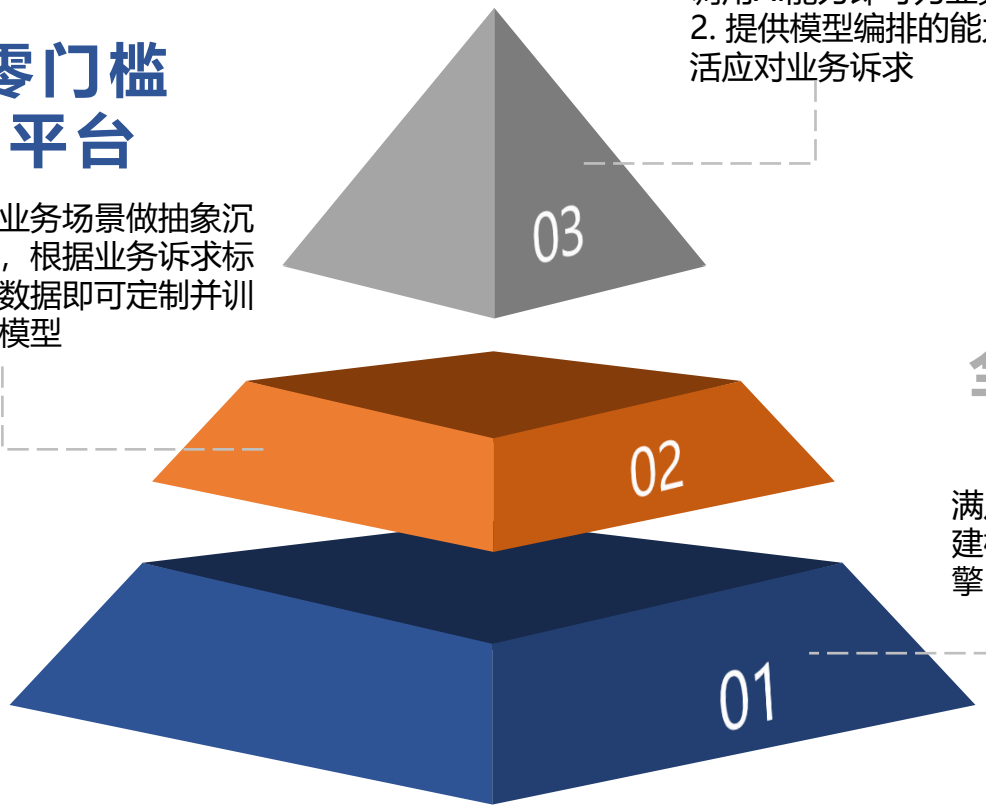
1 面向不同需求层次的技术支撑平台

开箱即用AI能力

1. 无需进行模型建模，直接调用AI能力即可为业务赋能；
2. 提供模型编排的能力，灵活应对业务诉求

零门槛平台

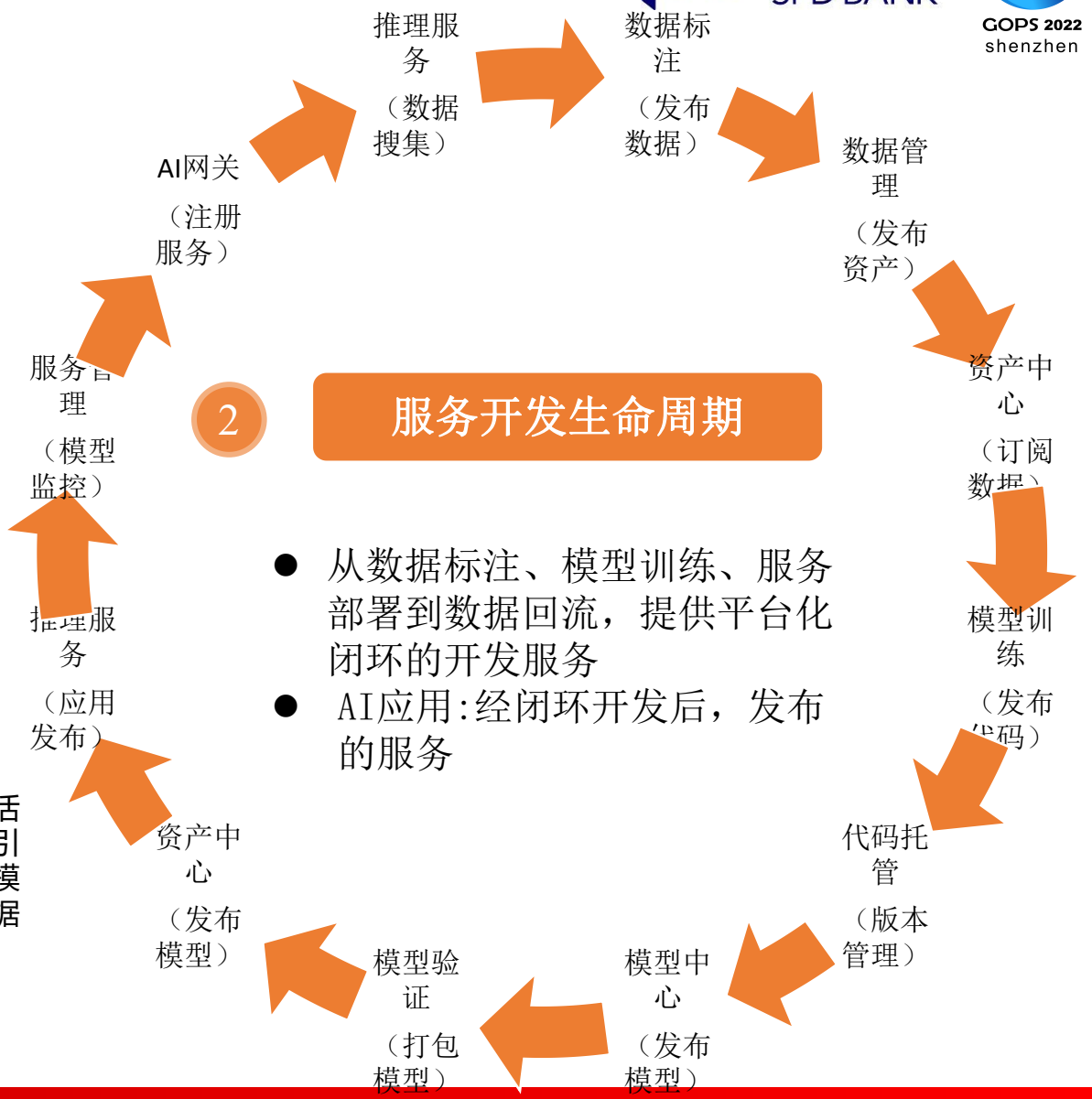
将业务场景做抽象沉淀，根据业务诉求标注数据即可定制并训练模型



全功能平台

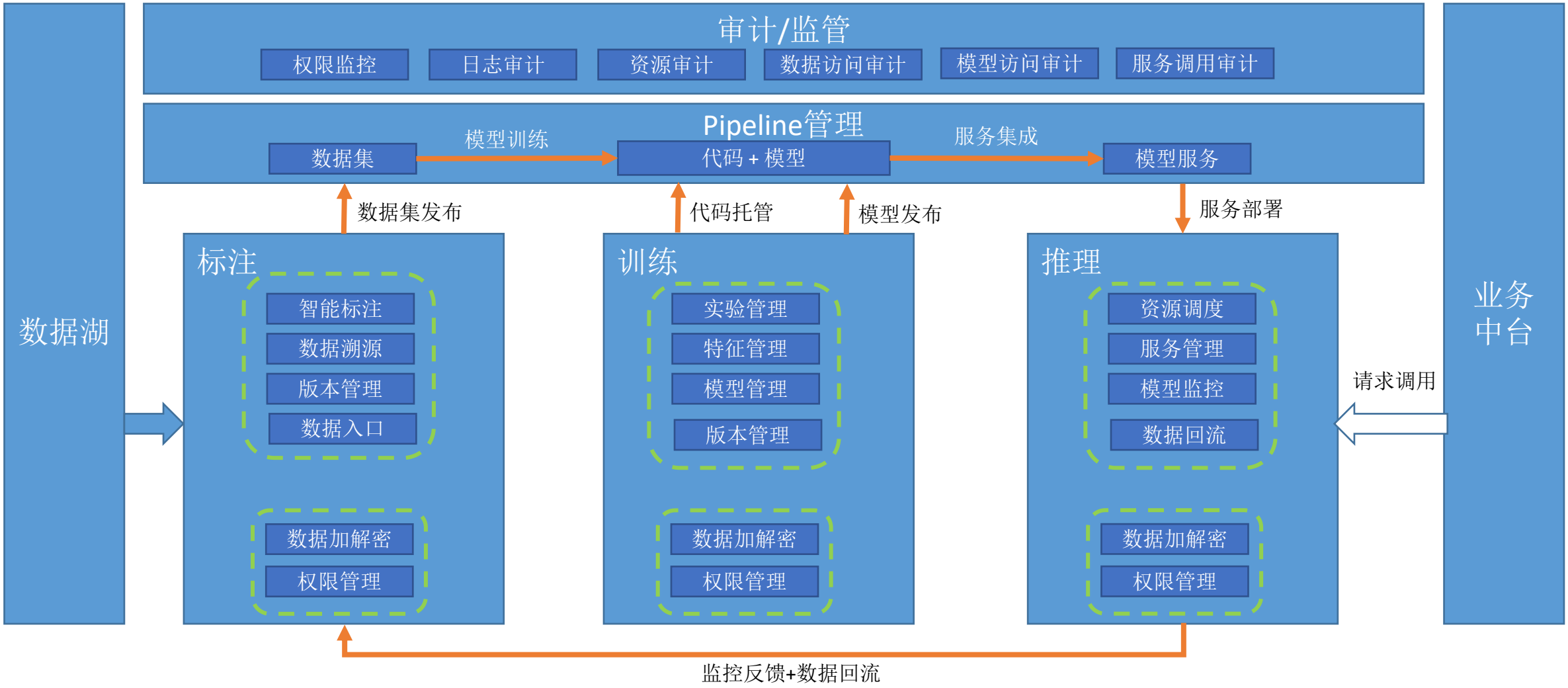
满足各种不同业务灵活建模，升级数据处理引擎，低门槛处理大规模数据

2 服务开发生命周期

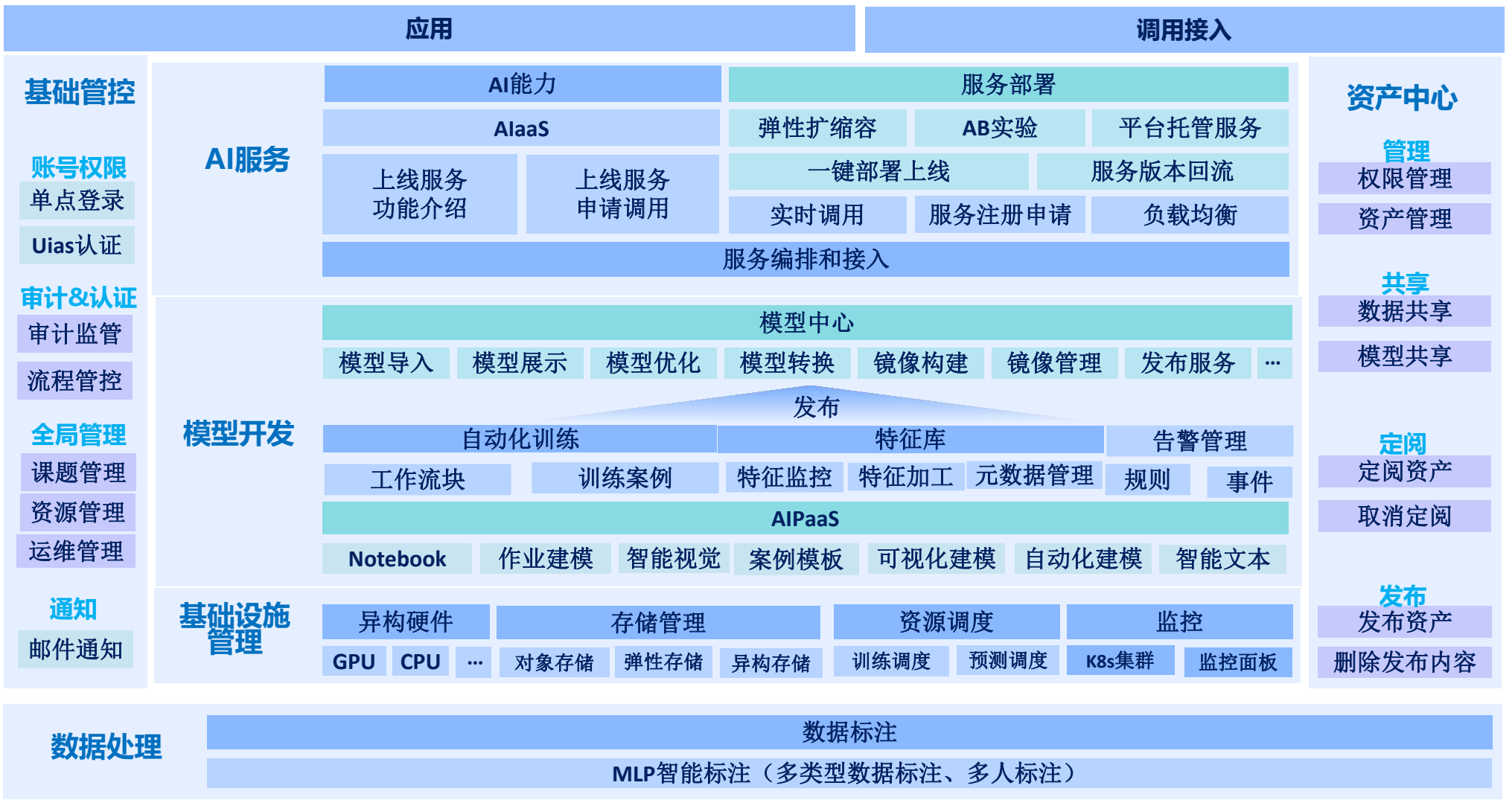


- 从数据标注、模型训练、服务部署到数据回流，提供平台化闭环的开发服务
- AI应用:经闭环开发后，发布的服务

深度学习平台MLOps流程



MLOps底座-深度学习平台系统框架



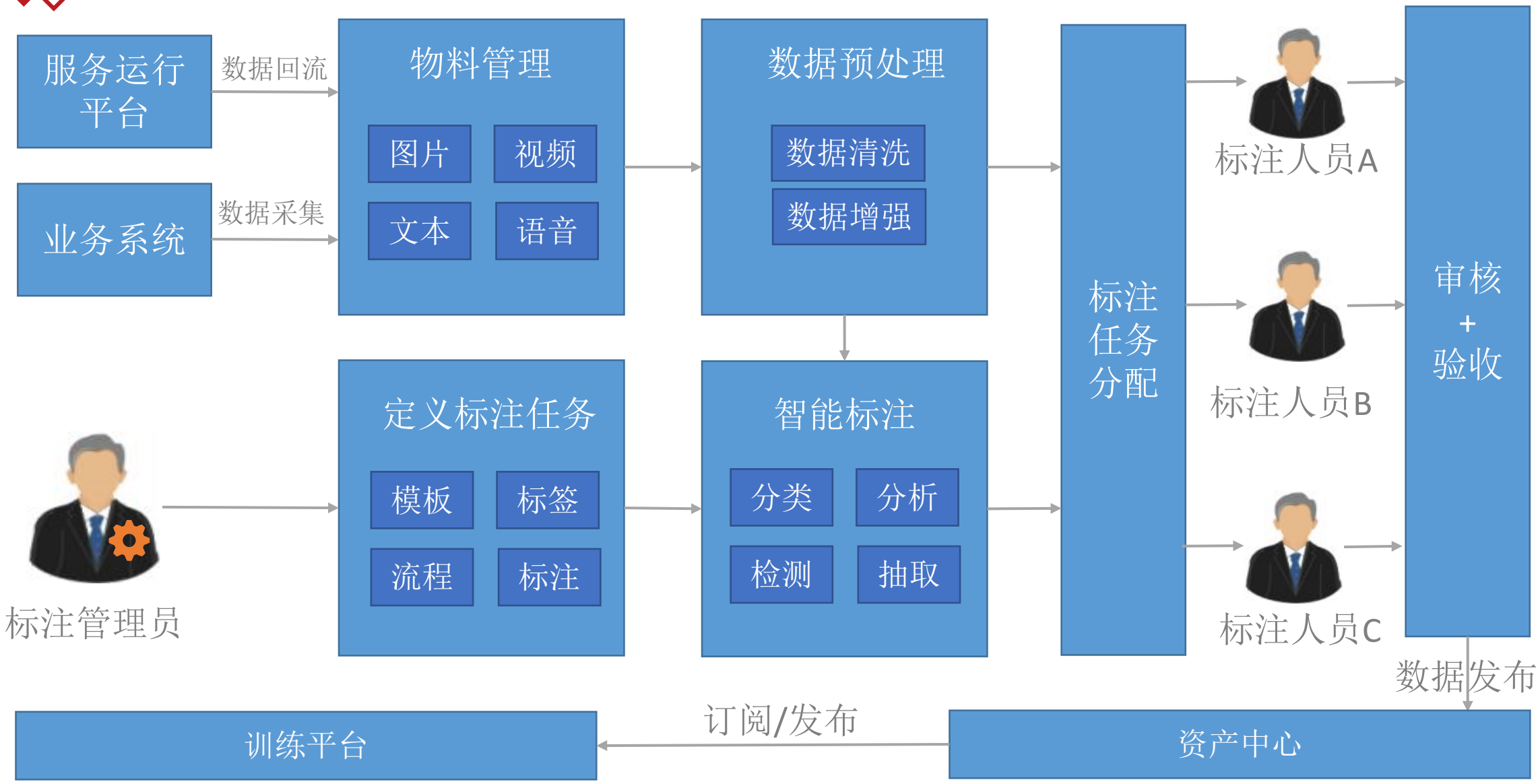
全行统一的人工智能平台

实现对人工智能模型从训练、测试、部署、运行、迭代的全生命周期的研发管理，引入多种机器学习、深度学习先进算法和模型，加速人工智能应用在全行业务场景的落地

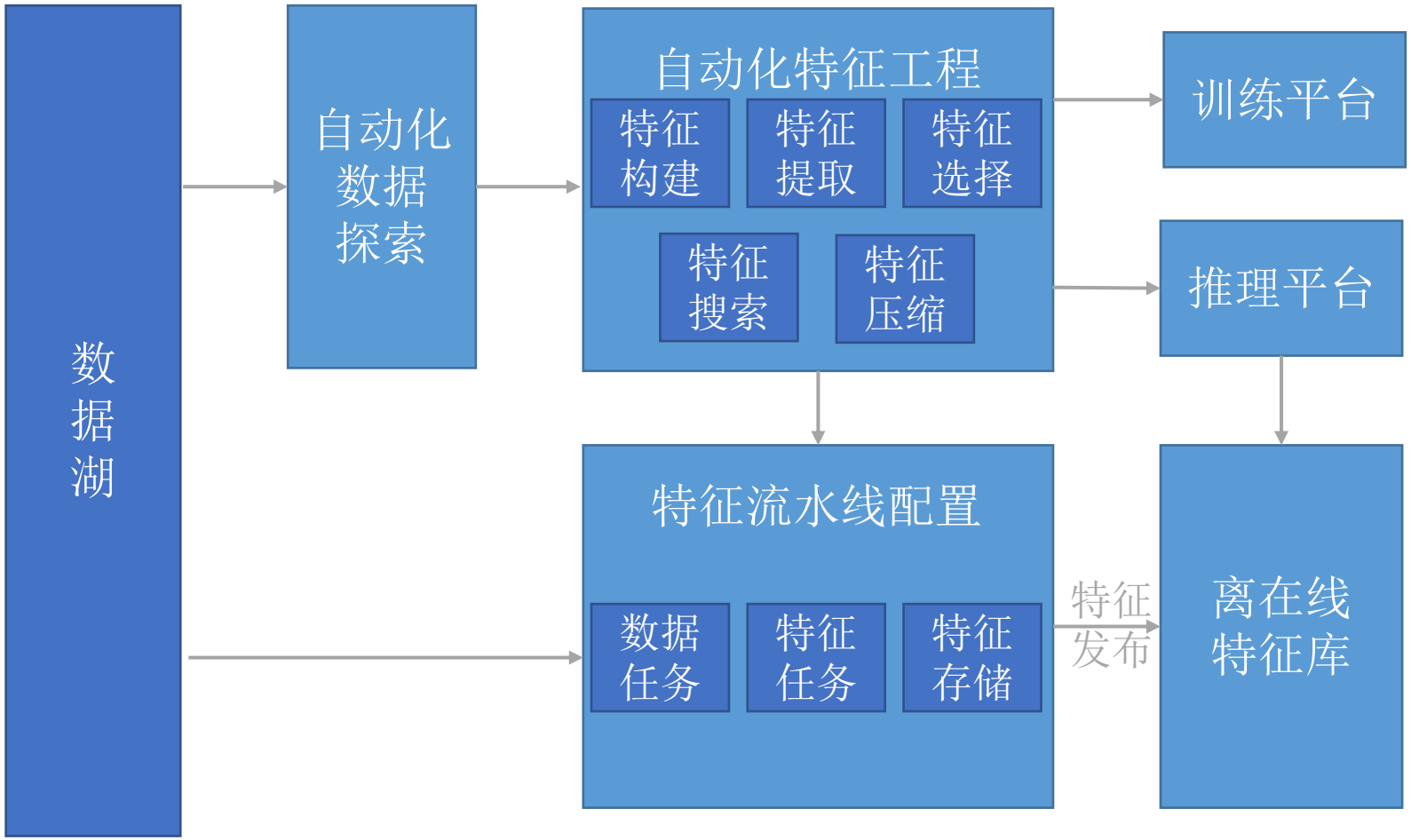
完善的安全机制

构建统一的风险模型实验室，提供集中的大数据及统计分析环境，实现模型开发与模型投产无缝衔接，提高模型迭代效率

深度学习平台组件-数据标注



深度学习平台组件-特征工程



自动化数据探索

支持多种自动探索及 丰富的可视化能力

自动化特征工程

可自动进行特征构建、提取、选择等任务

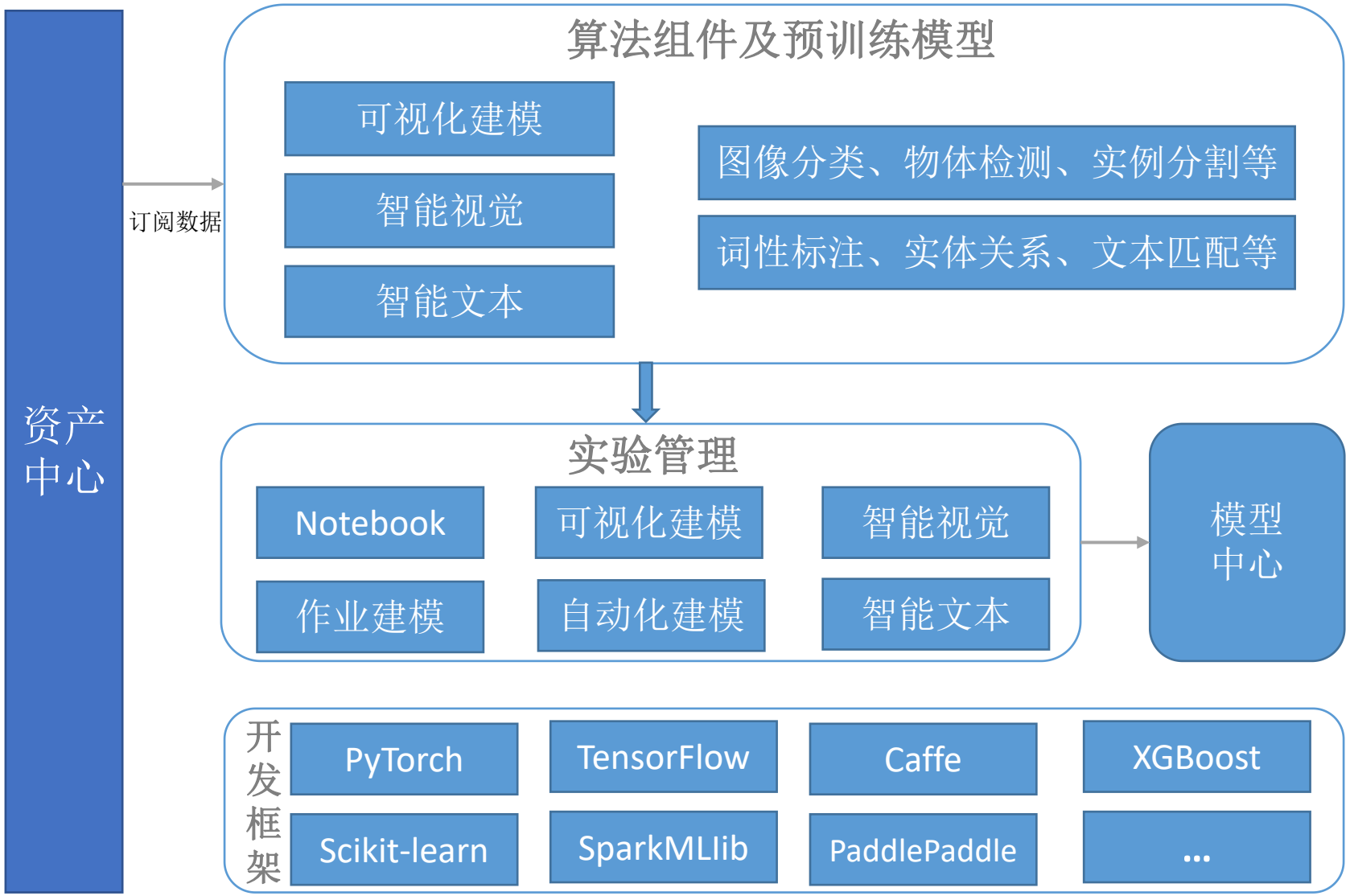
离在线一体特征库

1.解耦特征生产和特征消费环节，降低数据端到端的依赖，提升模型迭代效率

2.实现不同团队间的特征共享和复用，支持特征溯源和追踪

3.同时支持离线特征高吞吐使用场景和在线特征低延迟使用场景

深度学习平台组件-AIpaaS



多种建模方式

1.即开即用的代码式开发环境
2.零代码可视化工作流
3.自动化建模AutoML
4.实验管理
5.生产线建模：智能视觉、文本等

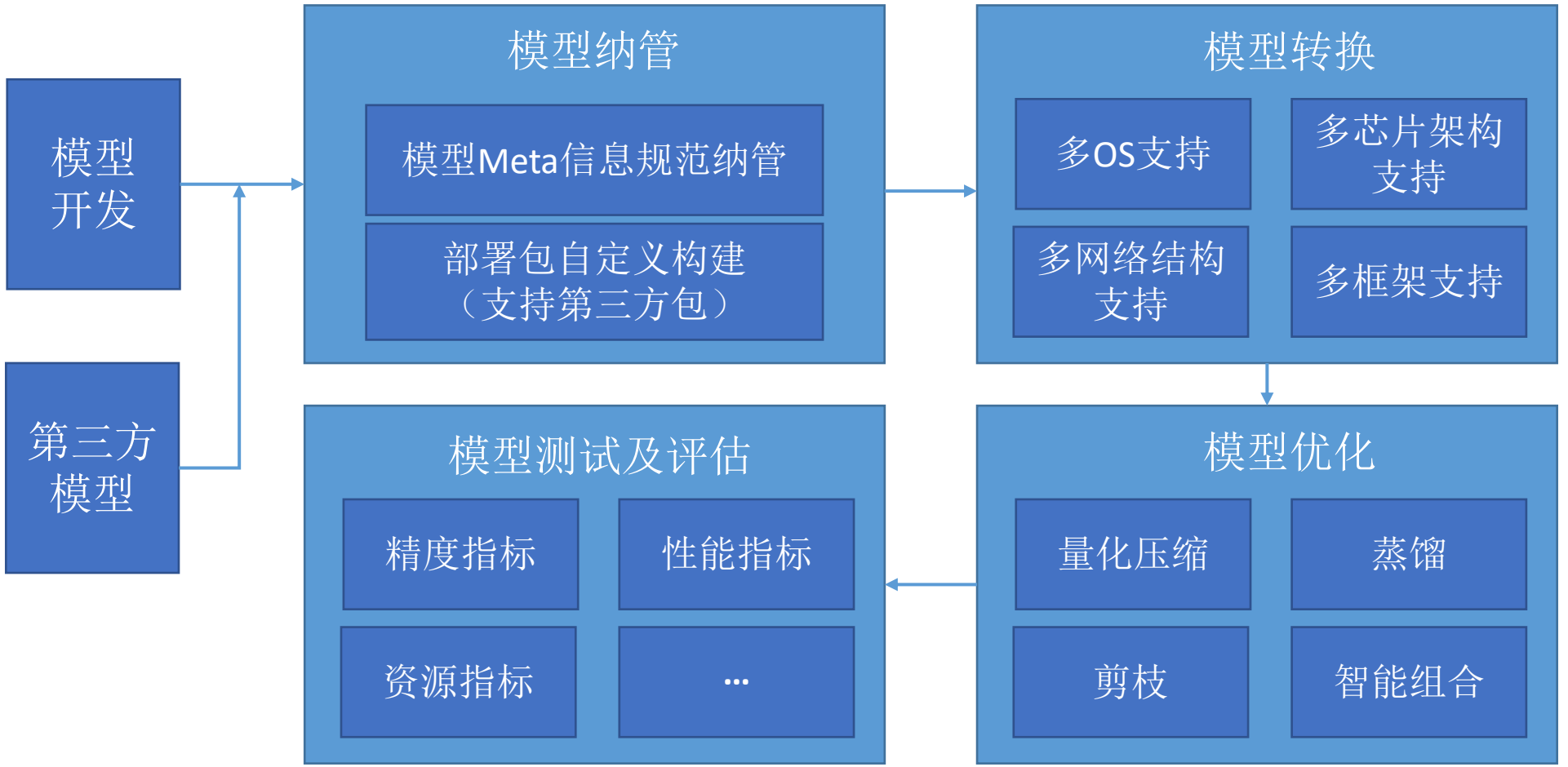
丰富的开发框架

平台支持主流机器学习和深度学习开发框架，支持以自制Docker镜像的形式支持其他框架和第三方软件库

内置多种预训练模型

包含多种经优调后的预训练模型

深度学习平台组件-模型中心

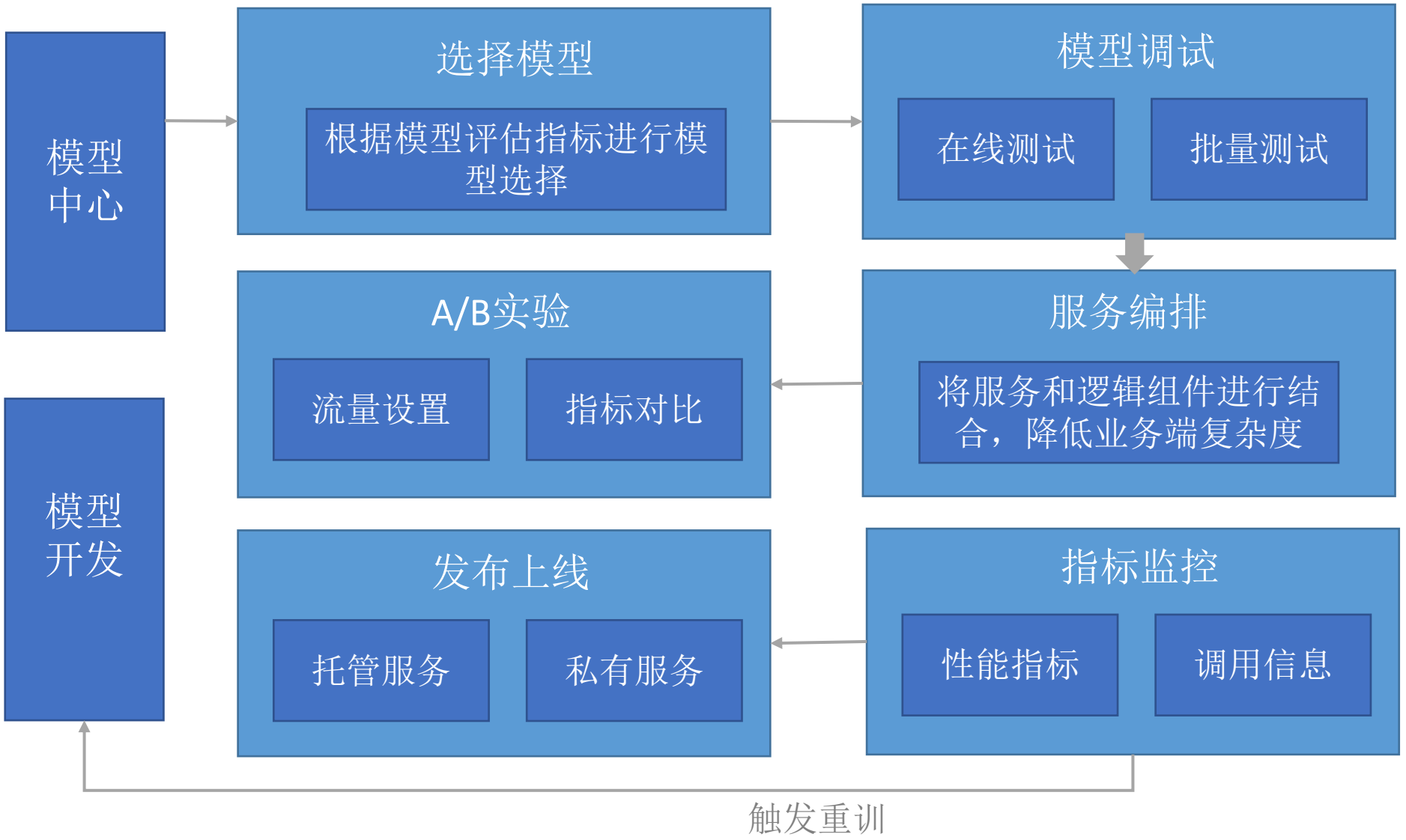


1.强大的模型纳管能力，构建AI模型统一管理规范

2.领先的模型优化技术，极致优化模型推理性能

3.专业的模型工程化支持，满足云、边、端各类硬件的部署要求，边缘部署面向10+种硬件芯片

深度学习平台组件-服务部署



服务发布

支持在线预测服务、离线预测服务，满足服务发布过程中的在线调试、异步批量预测、模型转换、压缩和加密等出处理

服务编排

通过服务的编排组件，对一系列的单点能力进行编排

指标监控

自动检测部署的模型，并在检测出现问题时发出邮件警告

深度学习平台组件-资产中心

浦发银行 SPD BANK

GOPS 2022 shenzhen

资产中心

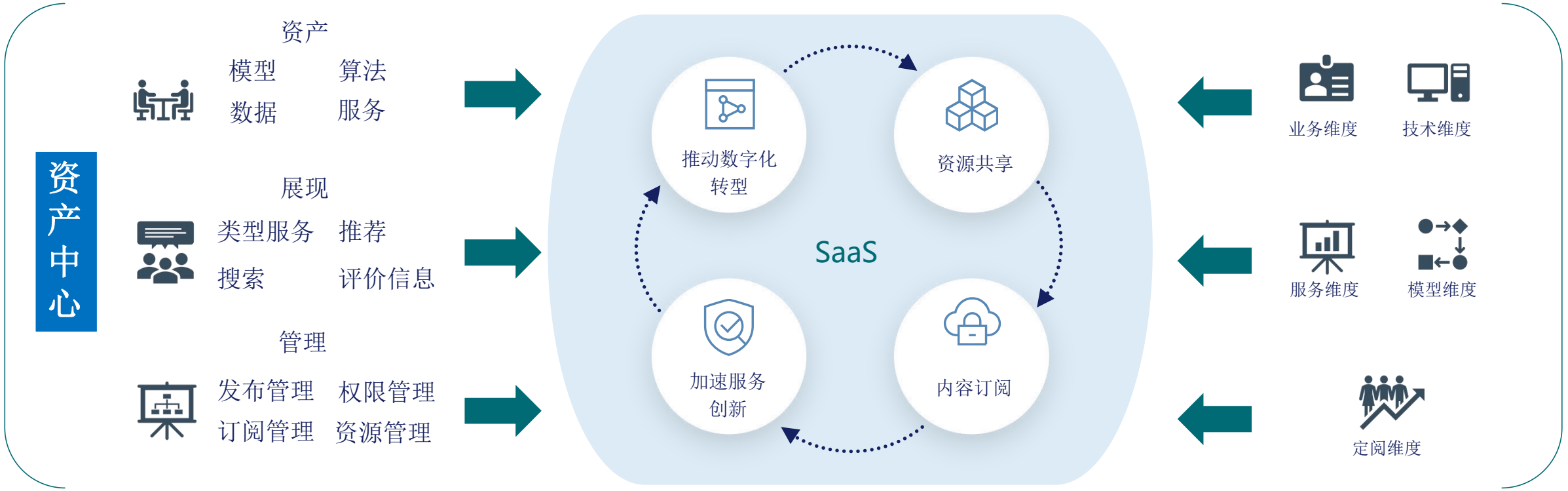
AI资源（包括模型、数据集、特征工程、工作流、产线、算法等）是一个即插即用的仓储型共享平台

安全托管

安全私密的方式对资源进行管理

资产沉淀及共享

资产SaaS化，加速资产落地，提高资源重复使用率，减少重复建设



深度学习平台-MLOps标准建设情况

浦发银行 SPD BANK

GOPS 2022 shenzhen

2021年 | 2022年迄今 | 2022年底

MLOps 能力项

进度管理　资源管理　数据标注

数据质量评估　开发环境管理

已完成

数据接入　流量管理

模型部署

数据组织　模型选择

算法组件管理　模型管理　集成发布

模型配置　集成过程管理

升级

风险管理　评估指标管理　模型监控

集成版本管理　评估任务管理

实验管理　元数据管理

已完成

超参数优化　流水线构建

开发环境管理　特征管理　数据预处理

数据版本管理　特征选择和处理

评估指标管理　构建制品管理

更新策略管理　场景构建

升级

探索性分析

建设中

测试报告管理

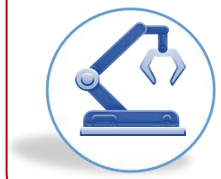
构建计划管理

测试任务管理

MLOps实践案例-基础大模型训练

浦发银行 SPD BANK

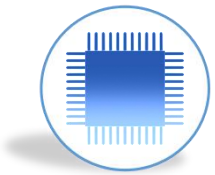
GOPS 2022 shenzhen



- 上百亿金融语料
- 多类亿级业务数据
- 数据预处理

难点：数据量多、特征管理难

自动化数据分析、自动化数据标注、数据集管理、特征管理



- 上百块GPU加速卡
- 大规模分布式训练
- 百亿参数模型构建
- 超参调优、训练迭代

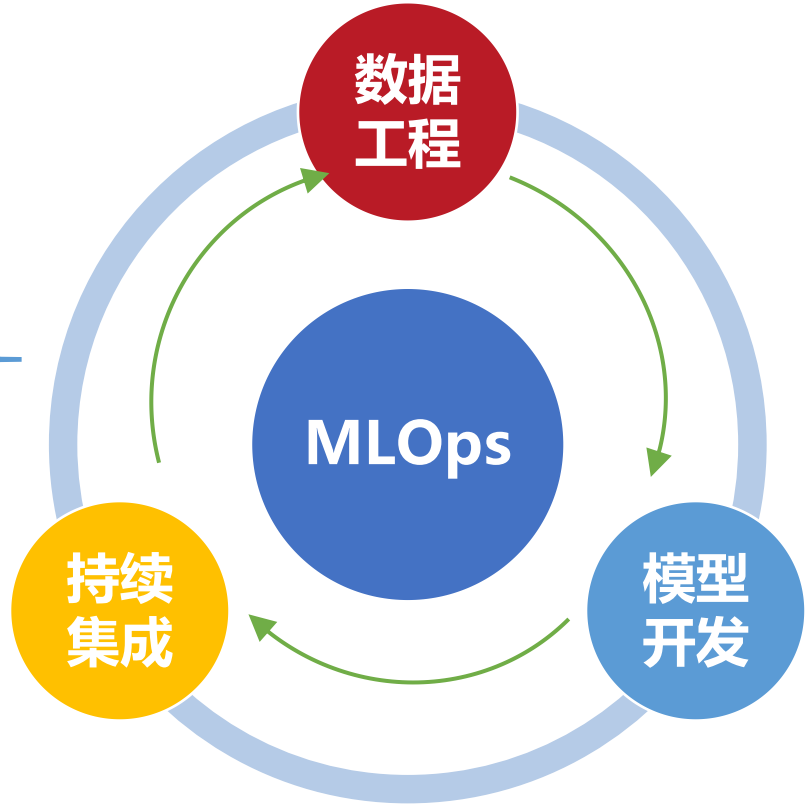
难点：计算量大、资源有限

资源调度、实验管理，模型评估

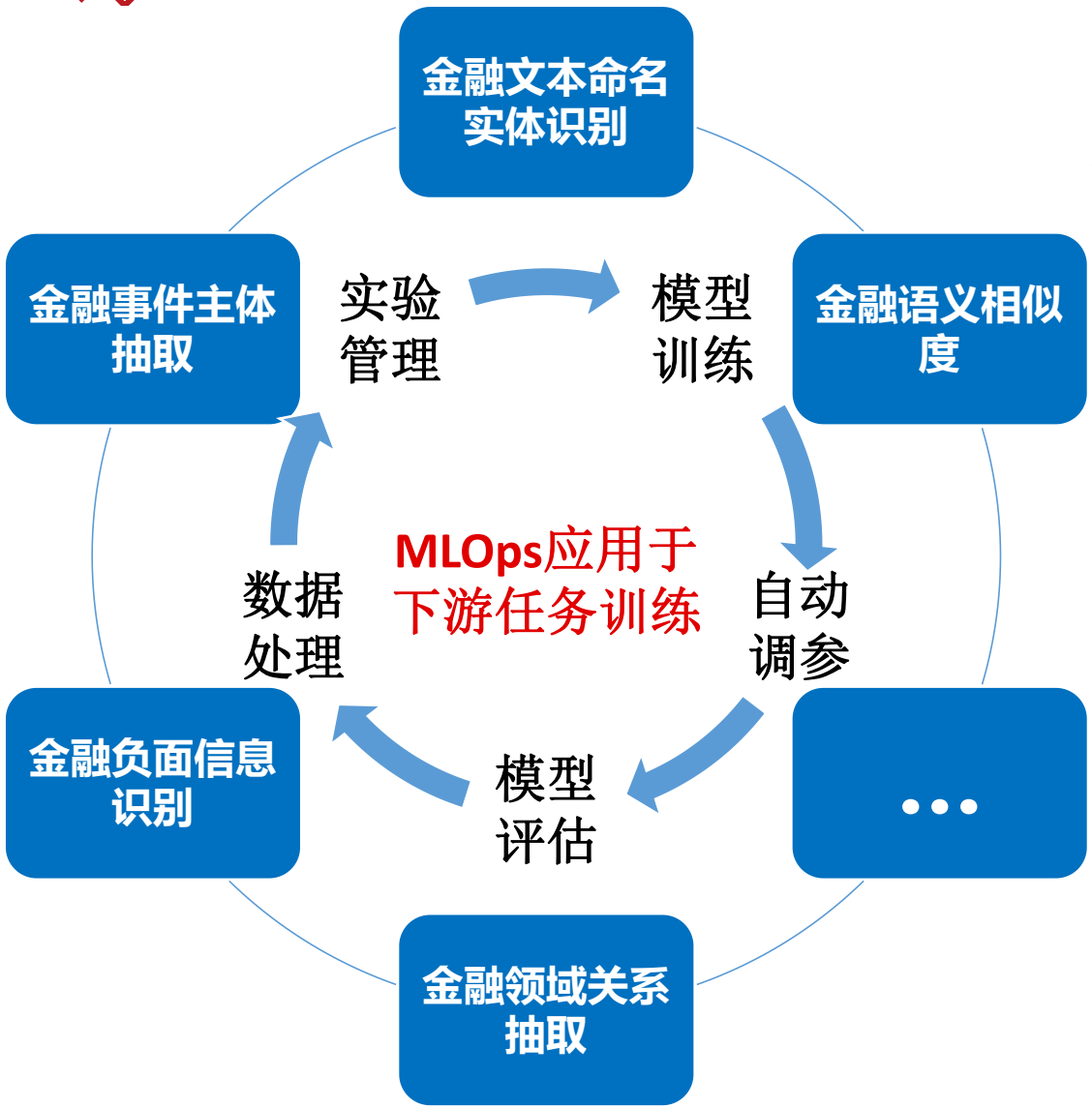
- 大模型推理
- 效果跟踪
- 模型迭代

难点：模型验证慢、推理速度慢

模型加速、自动重训、模型监控



MLOps实践案例-大模型下游任务批量调优



大幅降低下游任务落地时间！

效率提升

- 同步验证20多个下游验证任务
- 每个任务并发5个实验流水线
- 3天内完成所有下游任务

MLOps实践案例-风险传导

浦发银行 SPD BANK

GOPS 2022 shenzhen

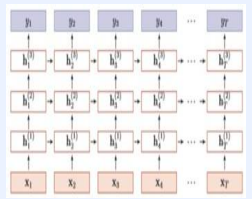
基于时序图神经网络的风险传导模型

原始数据集

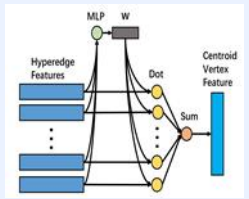
- 客户、集团基础数据
- 担保、投资、股权、上下游
- 工商、舆情、司法、征信等

表征数据集

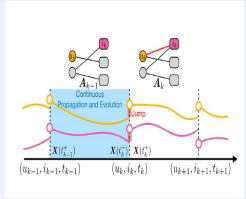
- 关系数据的多维向量
- 图节点表征邻接矩阵
- 时序关系的编码序列



深度循环神经网络



图神经网络



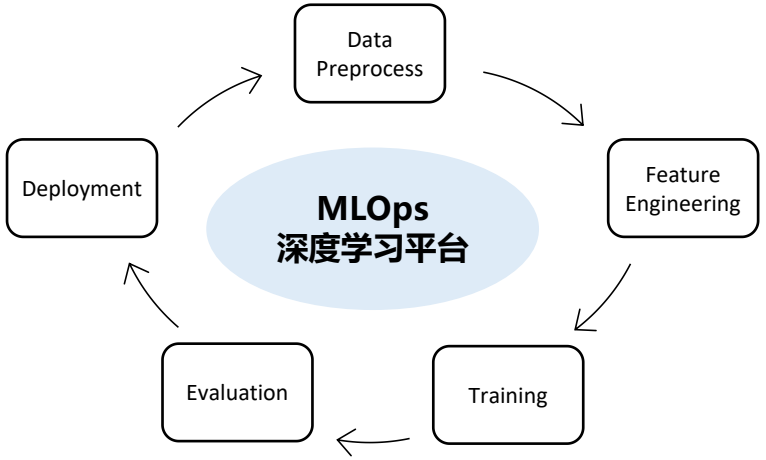
交互传播演化模型

难点问题

算力问题	• 企业间拓扑关系复杂，**需要大量算力资源支持**
人力问题	• 多维数据处理及特征工程，**需要耗费大量人力资源进行处理**
时间问题	• 风险传导场景使用的时序模型，**需要耗费大量时间进行迭代训练**
调优问题	• 动态时序图网络结构的模型，**需要进行大量超参调优工作**

※ **在有限时间和有限投入的情况下，采用MLOps有效的解决了风险传导模型建设过程中的难点问题**

MLOps实践案例-风险传导



训练、测试、评估、迭代全流程流水线管理

自动化缩短整体建模流程6个月

自动化数据处理、统计分析、特征工程方法

数据处理有效节省人力70+人天

支持多维度模型评估方法和重训策略

完成模型效果的自动化评估、对比和重训

通过实验管理同时发起一百多条训练流水线，每条流水线迭代优化上百轮

创新实验室-深度学习平台

上百块GPU加速卡

同时并行100多个模型的数据特征处理、训练和模型评估，减少人工干预，总计节省了100+天的训练时间

浦发银行
SPD BANK

03

MLOps在金融行业的展望和挑战

金融行业实施MLOps的差异

在金融行业，实施MLOps不仅意味着在机器学习系统构建流程的所有步骤中实现自动化和智能化，还需要顾及监管、安全、审计和可解释性等。

事项	MLOps	金融领域MLOps
持续训练/持续监控	✓	✓
数据来源	✓	✓
训练结果可复现	✓	✓
模型管理	✓	✓
数据存储安全/传输安全		✓
特征衍生合理性/可解释性		✓
训练过程可追溯		✓
模型可解释性		✓
服务攻击监控		✓
符合监管要求的规范、流程制度		✓

MLOps在金融行业的展望

浦发银行 SPD BANK

GOPS 2022 shenzhen

安全

- 完善的鉴权和授权体系：抽象出系统组件和功能，通过最小授权原则进行赋权
- 更好的数据保护：通过加密方式保护数据的存储和传输
- 完备的审计：确认职责，识别和阻止恶意活动

监管

- 可追溯：除数据、模型等静态文件可以追溯，操作、流程等动态对象也可以追溯
- 可解释：通过主要特征的获取、模型训练过程来辅助模型的可解释性
- 可监控：监控生产中的数据漂移，对预先定义的策略做出反应
- 可复现：根据保存的基础信息，恢复模型训练结果

运维

- 提高开发质量：在开发环境中，通过自动化或人为的review机制来审核各个流程
- 降低预生产工作量:通过预制CI/CD流水线来降低预投产前的工作量
- 提升部署和监控效率：将CD和自动化的模型监控相结合
- 完备跟踪和告警：自动化模型指标跟踪，以便在出现异常时向用户发出警报

金融行业实施MLOps的挑战-数据安全性

浦发银行 SPD BANK

GOPS 2022 shenzhen

数据源

- 数据验证：数据可信可靠验证
- 数据管理：数据源纳入元数据进行管理

数据可追溯

数据全生命周期管理，包括数据获取时间、数据来源、数据量、数据标签、采样方法及过程等内容具有可追溯性

数据使用

- 数据存储：防止数据存储时被窃取或篡改，明确的数据存储安全防护措，数据存储介质标识的记录和跟踪
- 数据传输：数据传输过程中防止被窃取或篡改，明确安全防护措施
- 数据使用：使用数据时应有完整签名或校验码验证环节

数据内容

- 噪声值或异常值：具备数据检测、过滤噪声值和异常值的能力
- 数据分布：数据分布合理，避免出现因数据分布不均而导致的模型训练不准确
- 数据标签：数据应具备唯一标签，便于记录和管理
- 隐私数据：保障数据隐私，避免用户敏感信息泄露，涉及隐私应进行加密处理

金融行业实施MLOps的挑战-模型安全性

浦发银行 SPD BANK

GOPS 2022 shenzhen

模型存储

- 防止模型存储时被窃取或篡改，该环节必须制定明确的安全防护措施
- 模型存储介质的标识进行记录

模型传输

- 防止模型传输过程中被窃取或篡改，该模型传输环节必须制定明确的安全管理要求和安全防护措施

开源模型

- 使用开源模型或第三方模型时，需对该模型是否安全进行评估
- 基于开源模型或第三方模型做出的模型，需进行安全性评估

模型标识

- 模型应有唯一标识，如添加水印，具有对恶意相似模型的识别度
- 对模型的标识进行记录

模型训练

- 模型训练过程中，应具备对应的安全防护配置，如审计日志，保障训练步骤安全，防范训练过程被窃取，避免训练信息泄露

模型升级

- 模型版本可追溯性
- 模型升级的详细步骤、修改内容保留审计日志

模型仓库

- 模型需导入模型仓库，并在模型仓库中进行统一管理；
- 方便从模型仓库中抽调每个模型的要素，包括建模脚本、软硬件环境、建模人员、迭代次数及参数迭代

金融行业实施MLOps的挑战-服务安全性

浦发银行 SPD BANK

GOPS 2022 shenzhen

预上线服务	服务部署	服务监测
1）服务测试报告：包括但不仅限于安全性测试、性能测试、功能测试，保证服务可以正确上线； 2）测试结果详细记录，并对测试过程及结果进行内部评审。	1)服务部署可追溯性，包括服务部署人员、部署时间、部署脚本及软硬件配置信息，脚本应详细说明; 2)服务升级：升级内容、时间、操作人详细记录在审计日志中。	1)监测上线服务的运行状态及性能，及时反馈存在的缺陷问题; 2)特征数据质量分周期监控报告：日常监控（每天监控）。定期监控（周、月、季度监控）； 3）故障恢复，包括告警、自动定位、提醒处理等功能。

浦发银行 SPD BANK

GOPS 2022 shenzhen