

MLOps发展趋势与 工作介绍

秦思思 中国信通院

目录

Contents

① 发展趋势

② 标准介绍

③ 工作介绍



MLOps行业背景

机器学习模型研发面临的四大难题：难管理、难部署、难监控、难协作
MLOps解决技术债的同时体现其优势：多、快、好、省

数据及模型缺乏统一管理

ML面向的对象是数据、算法和模型的有机整体

模型开发部署迭代周期长

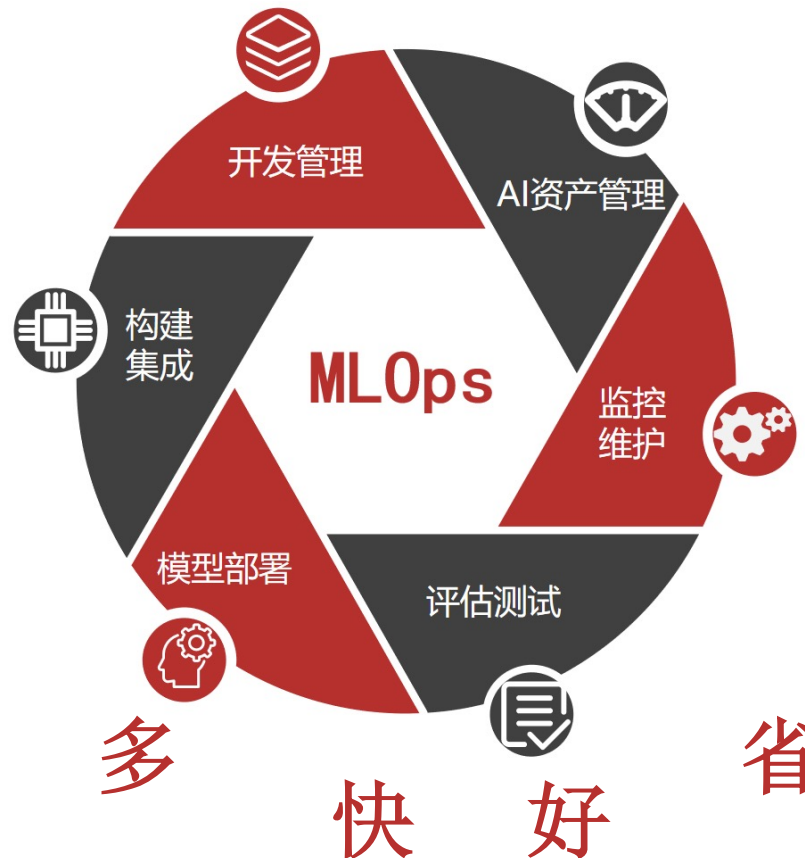
Algorithmia2020：很大一部分公司需要31-90天上线一个模型，其中18%的公司需要90天以上上线

模型监控体系不够完善

在部署的那一刻开始模型存在降级退化的风险：内容漂移和数据漂移

人员孤立无援

业务团队、运维团队和AI团队间的沟通鸿沟无法逾越



AI工程化发展：



工具层面



流程层面



MLOps



MLOps行业落地逐步深入，业务赋能不断提升

MLOps模型开发过程落地较为普遍，模型交付落地有待深入，模型运营作为高效闭环的核心落地有待推进和引导，AI资产沉淀及其风险治理成为**新关注点**



AI项目研发和交付效率大幅提升，某企业提高**40%**及以上



MLOps为**科技部门直接赋能**：提高IT部门科技服务能力，降本增效



MLOps为**业务部门直接赋能**：降低AI门槛，提高应用灵活度，直接赋能业务

2019、2020 连续2年进入Gartner数据科学与机器学习技术成熟度曲线，并被视为AI工程化的重要内容

2015 "Hidden Technical Debt in Machine Learning Systems" 首次提出机器学习生产化逐渐积累的技术债问题

3-5年内：

- IDC：到2024年，**60%**的中国企业将通过MLOps来运作其ML工作流程
- Cognilytica：MLOps市场规模将从2019年的3.5亿美元快速增长到2025年的**40亿美元**

2021

- 包括MLOps在内的**XOps**被Gartner列为2021年十大数据和分析技术趋势之一
- 吴恩达：MLOps将帮助每个人完成机器学习项目的整个生命周期
- 信通院牵头MLOps系列标准编制

2018

业内人士公开谈及工业生产中ML生命周期集成化管理的重要性，大公司开始尝试建立MLOps系统

MLOps发展趋势



2022年MLOps落地开花，核心内容逐渐清晰

MLOps：一个AI项目该有的模样

核心内容逐渐清晰：围绕**一个基础**、**两个关键**、**三个提升**打通从需求、开发、交付到模型运营的全生命周期

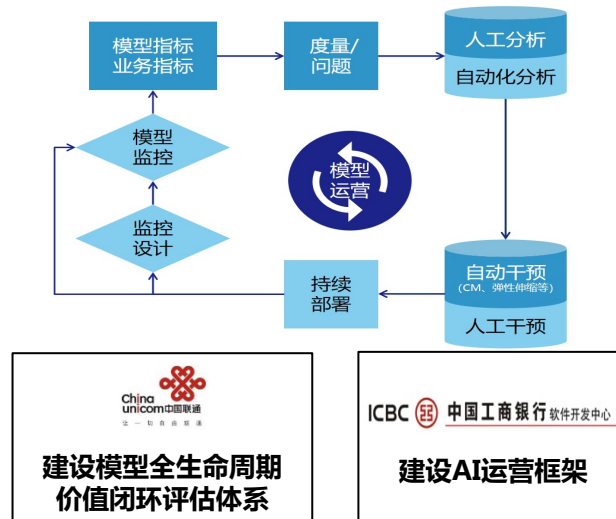
一个基础：搭建持续交付流水线，拓广度

以**持续交付（CI/CD）**为基础，搭建工厂流水线式的模型生产方式，提高规模化生产效率，拓宽MLOps落地广度



两个关键：加强持续运营和持续训练能力，推深度

以**持续训练CT**和**持续监控CM**为关键，搭建高效闭环的运营管理体系，提高ML可观察性，保证模型质量，增加赋能效果，推进MLOps落地深度



三个提升：数据、特征、模型等AI资产治理愈显重要，提高度

提升数据、特征、模型等**AI数字资产**管理能力，对其加以沉淀、安全管控和风险治理，提高企业级治理能力，提升MLOps落地高度





工具市场横向纵向同时发力，应用落地因地制宜

MLOps工具趋于**端到端**的同时，**专项工具**能力不断提升，工具市场热度持续攀升

MLOps 端到端工具

流水线编排工具

数据版本管理工具

交互式开发工具

拖拽式开发工具

元数据管理工具

实验管理工具

模型管理工具

特征存储工具

集成部署管理工具

模型监控工具

ML可观察性工具

...

MLOps 专项工具



端到端的工具平台具有**大而全**的特点



专项工具**纵向深入**，发展迅速，功能强大

应用行业落地**因地制宜**，选择合适的平台或工具链是落地实践主要方式，投入成本 **↑**

类别	平台式	工具链式
描述	以平台化方式采购或自研MLOps体系平台	采购各类专项工具组合能力
方法	方式一：一站式采购 端到端一体化平台 方式二：在已有AI开发平台上持续建设MLOps能力，打造 全能力AI中台	充分利用企业已有工具的基础上，增加MLOps工具，打通工具间的鸿沟，形成 一体式工具链
优点	一蹴而就，体系完善，功能强大	成本相比较低，无需改变过多用户习惯，局部落地较快
缺点	应用落地需要全员下决心，改变各角色行为习惯， 落地需要时间和坚持	工具间的衔接可能存在困难， 集成难度大 ，琐碎的优化改造需要大量人力成本

建议：应用企业应综合考虑AI项目需要、组织结构、企业AI战略规划、已有技术资产等要素，选择落地实践方式。

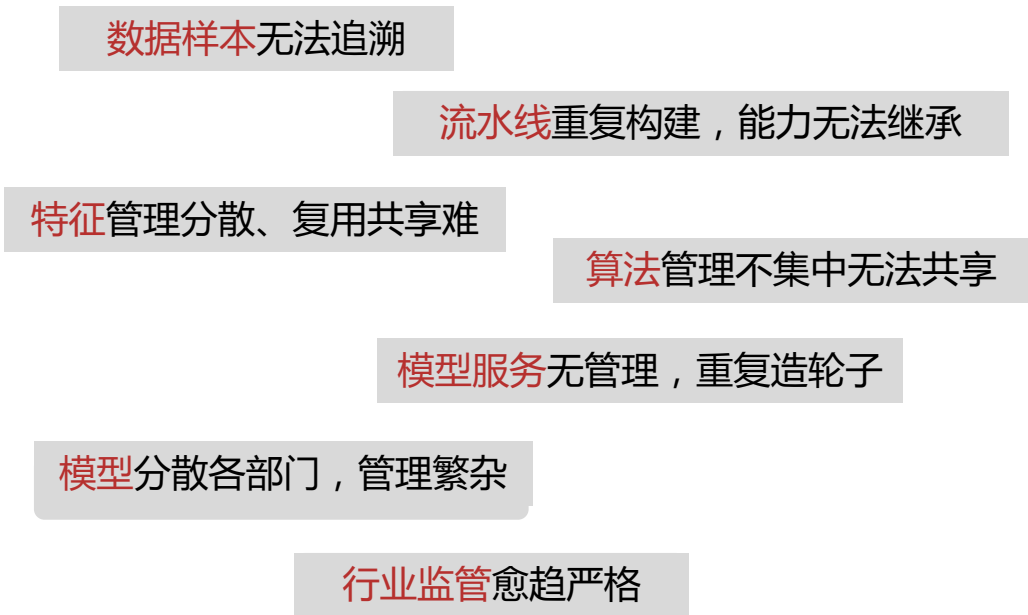
现实：没有一个平台能够一统江湖，没有一个工具满足所有行业需求。



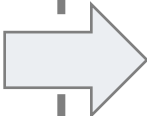
AI资产管理愈发重要，合规、安全、审计需求呼之欲出

模型类型及数量的增加、各类AI资产的增加和管理复杂，大规模AI应用的迫切，迫使AI资产管理提上日程

AI资产的沉淀和共享直接引发安全风险的担忧，模型的安全、合规、审计等需求突出，加强企业AI治理顺势而为



发文标题	商业银行互联网贷款暂行办法
发文标题	中国银保监会2020年第9号文
发文部门	中国银保监会
发文标时间	2020年7月





模型运营成为MLOps成败的关键，始末交加

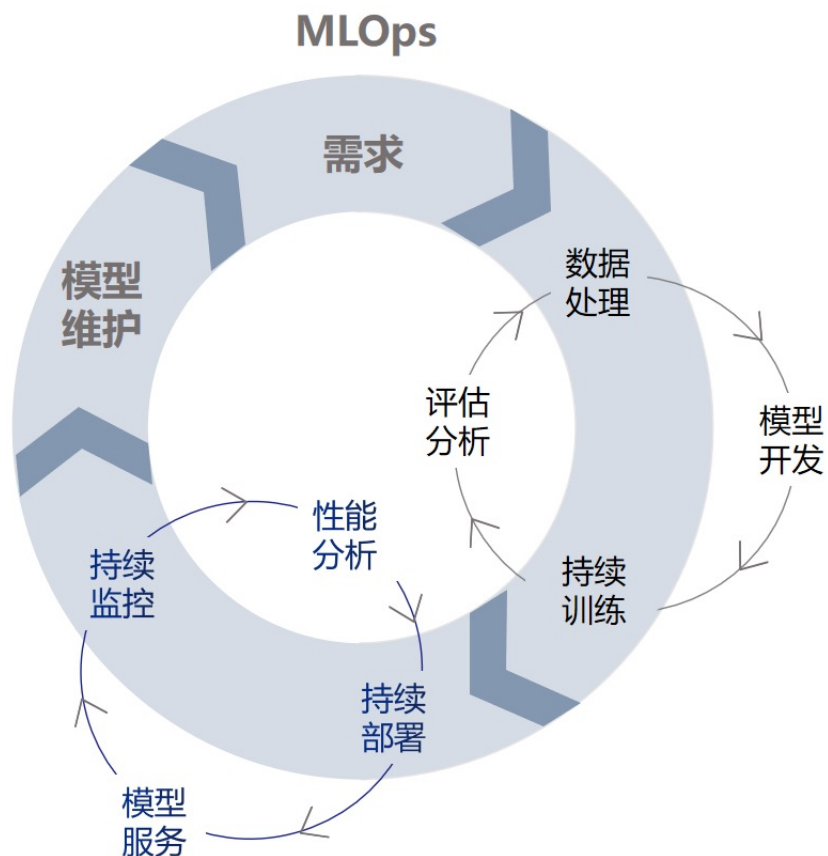
高效运营是保证模型服务质量、业务效果的重中之重，始末交加驱动运营闭环
建设模型**全生命周期高效运营体系**，提高ML**可观察性**，加速MLOps持续运转

始

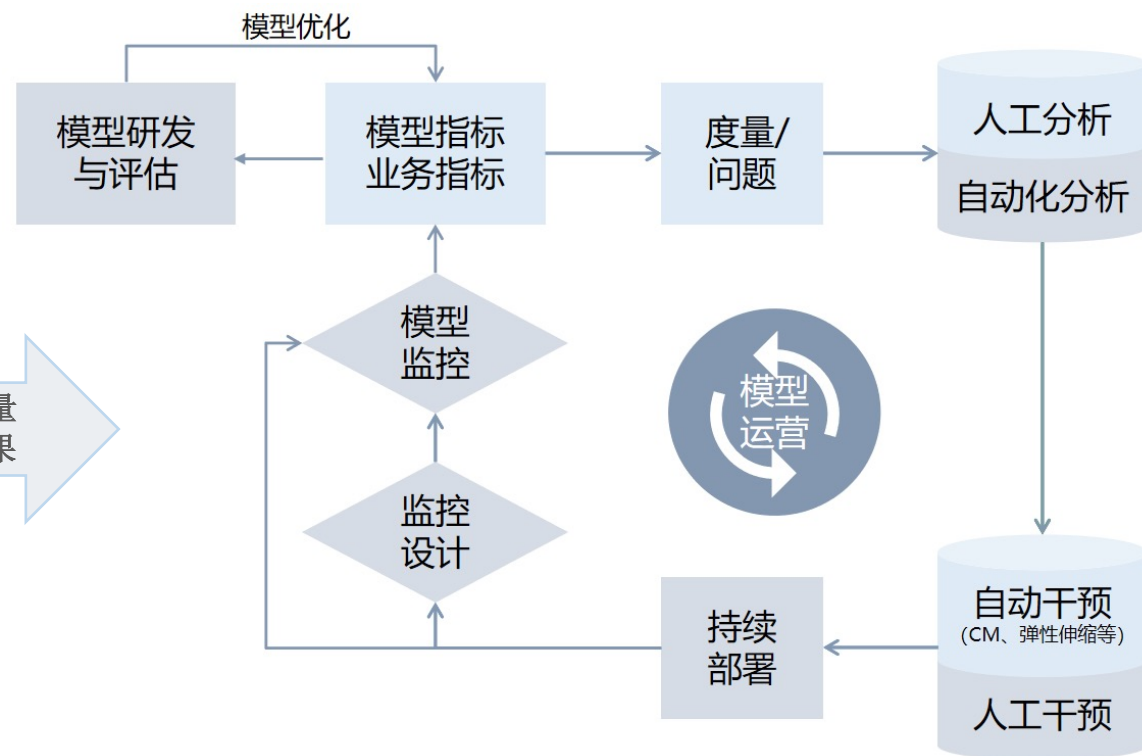
模型服务的开始环节

模型研发的最后环节

末



模型指标：保证模型质量
业务指标：评价模型效果



目录

Contents

① 发展趋势

② 标准介绍

③ 工作介绍



标准体系架构

本系列标准包括了MLOps的过程管理和ModelOps模型治理等内容，共分为7个标准

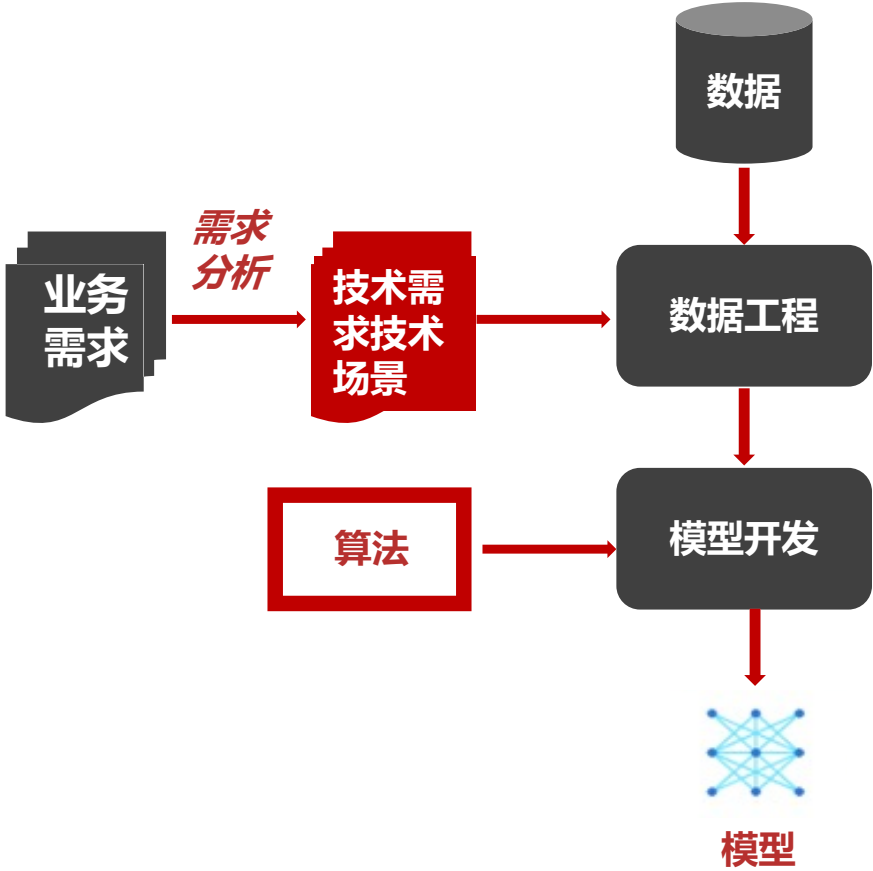




开发管理标准介绍

开发管理：将业务需求转化为技术需求，并根据已有数据，开展数据处理和模型开发，从而输出模型，作为集成和测试环节的输入

开发管理（标准一）																														
需求管理							数据工程							模型开发																
需求分析			测试用例设计		项目计划		数据收集		数据探索		数据处理		特征工程		模型构建		模型训练			评估与选择										
需求确认	可行性评估	场景设计	测试用例编写	测试用例管理	进度管理	资源管理	风险管理	数据接入	数据组织	探索性分析	数据质量评估	数据预处理	数据标注	数据版本管理	特征选择与处理	特征管理	流水线管理	开发环境管理	算法组件管理	流水线构建	实验管理	超参数优化	元数据管理	模型管理	评估指标管理	评估任务管理	模型选择			
3个能力域									10个能力子域									28个能力子项									200多个能力要求			





模型交付标准介绍（征求意见稿）

模型交付：将代码、模型和配置等作为输入，对其进行构建打包、集成测试等过程，形成可交付的模型及模型服务部署至相应环境，最终为业务系统提供服务

模型交付（标准二）																																	
持续集成						部署发布								模型测试						配置管理						度量反馈							
构建管理			持续集成			模型编排			持续部署			发布管理			测试执行			测试管理			版本管理			变更管理			环境管理			度量指标		质量驱动改进	
构建过程	构建配置	构建计划	模型转换与优化			集成过程	集成反馈	场景构建	模型配置	集成发布	模型管理	模型部署	更新策略管理	运行监控	流量管理	效果评估	模型下线	模型评估	测试分层与设计	自动化测试	测试数据管理	测试结果管理	版本控制	分支管理	依赖与模型管理	变更过程	变更追溯	环境类型与资源管理	环境依赖与配置	度量指标定义	度量数据管理	度量分析	反馈改进





参编单位



有40余家单位参与“开发管理”和“模型交付”标准编制
标准依托于中国通信标准化协会（CCSA）和中国人工智能产业联盟（AIIA）同步推进



ICS 35.240
L67

AIIA/P

中国人工智能产业发展联盟
AIIA/P 0006—2022



ICS 35.240
L67

AIIA

中国人工智能产业发展联盟
AIIA/P

人工智能研发运营一体化
(Model/MLOps) 能力成熟度模型
第二部分：模型交付

The Capability Maturity Model of Model/MLOps
Part 2: Model Delivery



人工智能研发运营一体化
(Model/MLOps) 能力成熟度模型
第一部分：开发管理

The Capability Maturity Model of Model/MLOps
Part 1: Development Management

2022-10 发布

2022-10 实施

中国人工智能产业发展联盟 发布



— 瀚亭科技 —



中国人工智能产业发展联盟 发布

目录

Contents

① 发展趋势

② 标准介绍

③ 工作介绍



工作范围

标准评测

- ✓ MLOps标准及评测
- ✓ ModelOps标准及评测
- ✓ 其他专项标准及评测



产业研究

- ✓ 落地实践指南
- ✓ 产业观察报告
- ✓ 技术工具图谱



生态建设

- ✓ 技术沙龙交流
- ✓ 产业活动峰会
- ✓ 最佳实践推广




咨询服务

- ✓ 项目咨询顾问
- ✓ 战略规划指导
- ✓ 从业人员培训
- ✓ 技术公开课程



首家”开发管理“参评企业—**中国农业银行**
三级 领先级

参评项目：人工智能服务体系——掌银生活页信息流推荐模型

能力成熟度	一级	二级	三级	四级	五级
需求管理			 中国农业银行 AGRICULTURAL BANK OF CHINA		
数据工程					
模型开发					

系统化开发管理能力

支持数据工程和模型开发流水线部分自动化执行，实现了数据和模型等资产的可追溯和共享能力，以及实验过程和结果的可跟踪

体系完
备性

自动化
程度

流水线
执行程
度

AI资产
管理能
力



Thanks

开放运维联盟

高效运维社区

DevOps 时代

荣誉出品



想第一时间看到高效运维社区
的新动态吗？

