

基于AI+大模型的运维智能体探索与实践

江苏移动 数智化支撑部

2025.04





江苏移动 资深运维管理专家 胡炜

负责PaaS层架构规划和数智化运维能力提升，以及日常维护和系统优化工作，确保IT云安全稳定运行。

相关工作经验

- 具有多年的运维管理经验、运维系统规划经验、业务连续性管理经验、AIOPS运维能力构建经验。
- 集团IT运维高级专家，高级通信工程师，取得第一发明人国家专利12项；多次荣获省级管理创新一、二、三等奖及多项在岗技术革新二类成果奖等。



目录/ CONTENTS

1

AI+支撑思路

2

运维大模型应用实践

3

创新成果

4

运维经验推广



背景：AI与大模型技术融合，推动IT运维向智能化高效转型

近年来，大模型技术引发新范式、新时代、新机会，其中IT系统结构复杂度进一步提高，生成式AI大模型与AIOps结合成为必然趋势。大语言模型+Agent能够自主理解、规划决策、执行复杂任务，可以代替运维专家使用工具完成运维任务，让系统运维从+AI转向AI+成为可能。

集团公司战略引领

推进“AI+新兴技术”，促进新一代信息技术创新突破。发挥AI赋能其它技术创新的“元技术”作用，**依托其在知识整合、逻辑推理等方面的能力，洞察在传统技术研发中难以发现的规律，加速算力网络、大数据、区块链等技术的演进升级**，不断激发新一代信息技术的创新潜能。

——中国移动董事长杨杰

新技术范式带来革新

人工智能技术与运维服务相结合领域，当前发展：

- ①生成式AI，已经从技术萌芽期进入到期望膨胀期的顶点；
- ②ModelOps处于期望膨胀期和泡沫破裂谷底期边界处；
- ③自治系统在2023年进入技术萌芽期。

江苏公司领会落实

江苏公司紧扣“一二二五”战略实施思路，**深化数据治理**，持续**夯实数智技术底座**，加速“+AI”向“AI+”转变，加强转型人才培养，对内赋能公司**全域数智化转型**，对外拓展**能力价值变现**，为公司高质量发展贡献江苏力量。

依托中移聚智和智算两大基座，推进组织、技术、产品三大体系建设，发力对内注智和对外赋能两大方向，实现“全场景、多行业、一体化” AI+赋能支撑体系。

“1232”
工作思路



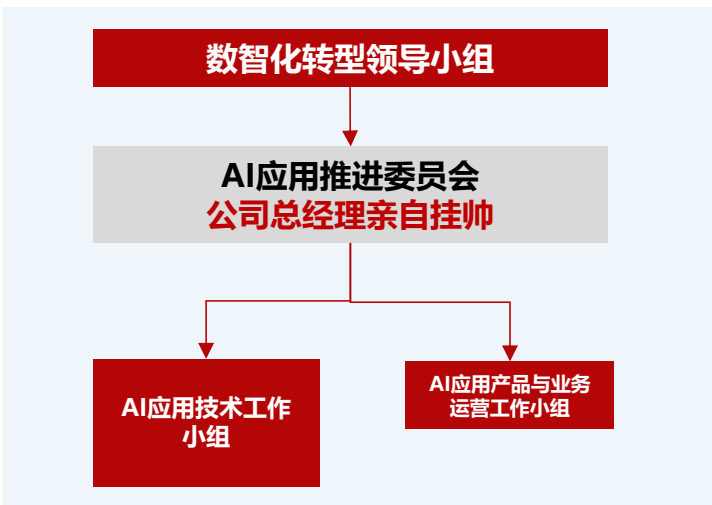
3 大体系 | 组织、技术、产品

- 产品方面：基于IT现有平台类组织方面，成立公司领导亲自挂帅的AI应用推进委员会，负责研究制定公司AI相关技术路线、产品业务规划及应用等整体方案及运营管理。
- 技术方面：做好多模态数据归集和湖仓一体存储，推进场景/行业大模型的建设和应用。
- 产品实施三化改造，建设大数据类、AI类、IT技术类平台产品。

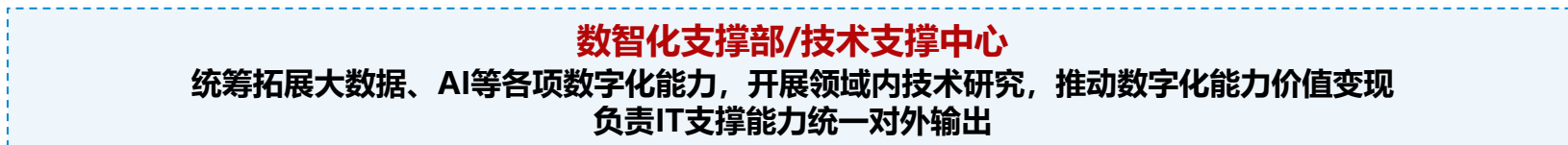
AI应用推进委员会

关键技术研究

IT能力三化改造



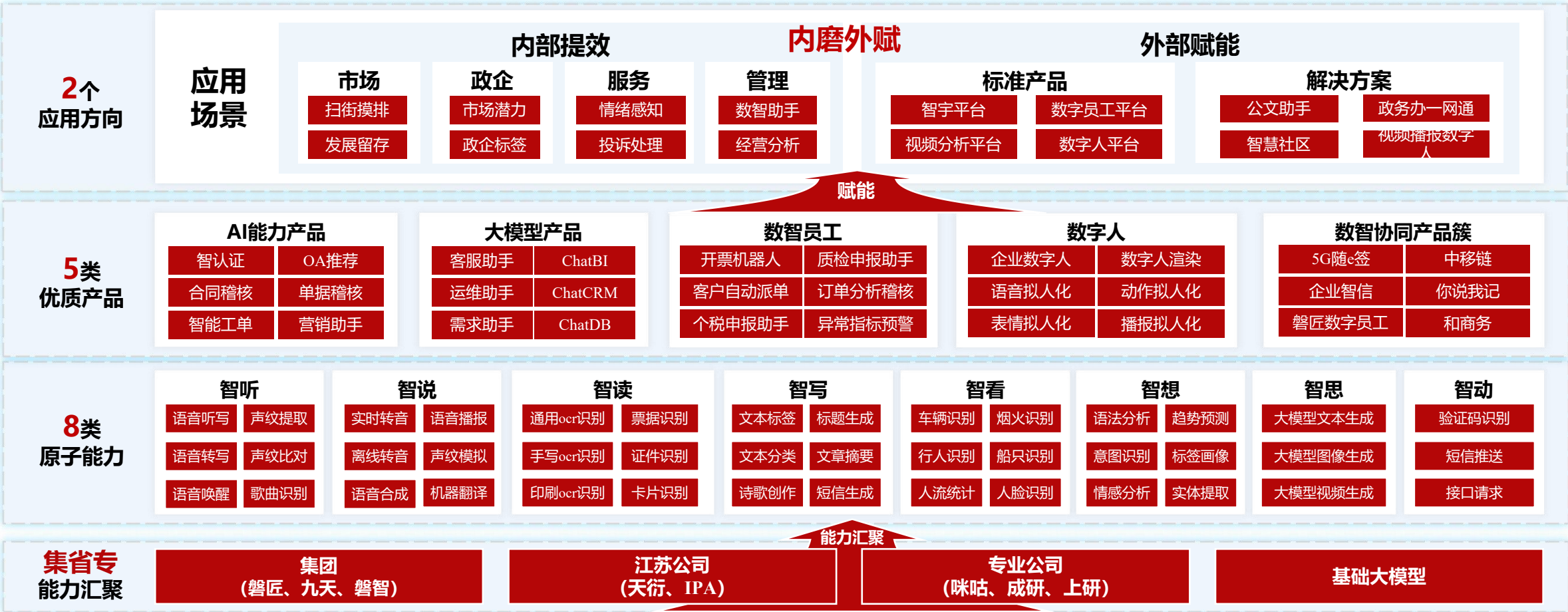
对内支撑数智化转型，对外输出IT能力和产品



发力2 个方向 | 对内注智、对外赋能

按照“资源整合、能力协同、体验优先、效率导向”的原则，“复用总部、汇聚省内、引入第三方”优质能力，持续丰富“听说读写看想思动”八大类AI能力图谱，精心打造五类优质产品，构建“AI+”数智能力支撑体系。

- 1 大小模型融合，构建全模型矩阵
- 2 汇聚省专能力，绘制全能力图谱
- 3 完善场景支撑体系，实现内外赋能





目录/ CONTENTS

1

AI+支撑思路

2

运维大模型应用实践

3

创新成果



运维大模型场景规划

FinOps云成本管理

云成本管控

成本观测量

成本趋势分析

云方案管理

上云方案生成

上云方案审核

AIOps运维自动驾驶

故障治理

故障预测感知

智能诊断评审

自愈方案推荐

故障协同调度

ChatOps

Chat巡检

Chat变更

Chat部署

Chat发布

DevOps敏捷交付

流程提效

变更方案智能评审

工程割接智能评审

架构调整智能评审

...

运营分析

上线质量监理

灰度运行分析

版本质量分析

故障回溯

专业领域助手

容器云平台智能体

多模态机房巡检

存储智能体

网络智能体

OB智能体

GodelDB智能体

Pulsar智能体

...

正在推进

后续实施

背景及必要性：云原生运维面临的问题和挑战

集、省、专公司核心系统都在迁移磐基，上云部署及日常运维工作非常复杂及工作量巨大。其核心价值：一是解决磐基迁移过程中的上云部署配置提效、二是规模化上云后尽早发现磐基平台异常和隐患、三是基于AI+手段赋能故障诊断和恢复实现端到端闭环修复。提升磐基交付人员、磐基一、二线运维人员及上云租户日常效率。

磐基背景：

磐基PaaS平台基于支撑超大规模集群的业务需求驱动，沉淀优秀云原生能力，总结一线运营运维生产实践，以应用为中心，灵活容器化快速部署，统一微服务治理，提供一站式上云，保障应用平稳运行。

AI+大模型背景：

随着AI技术的突破性发展，通过构建运维智能体，实现运维智能问答，运维任务执行、分析解读和决策处置，能有效提升隐患发觉和问题修复的效率。

面对的问题



实际的挑战



1、如何提升核心系统迁移容器云效率？

核心系统迁移PaaS容器云平台过程中因涉及较多的技术栈，**运维和交付人员**需要大量的时间处理迁移过程的上云及部署配置等问题，从而影响交付的效率。



2、如何尽早的发现容器云平台异常和隐患？

随着规模化系统上云，日常运维人员面临大量的运维保障工作，**一线需要及**对发现问题隐患进行分析定位、二线专家制定处置方案及修复。

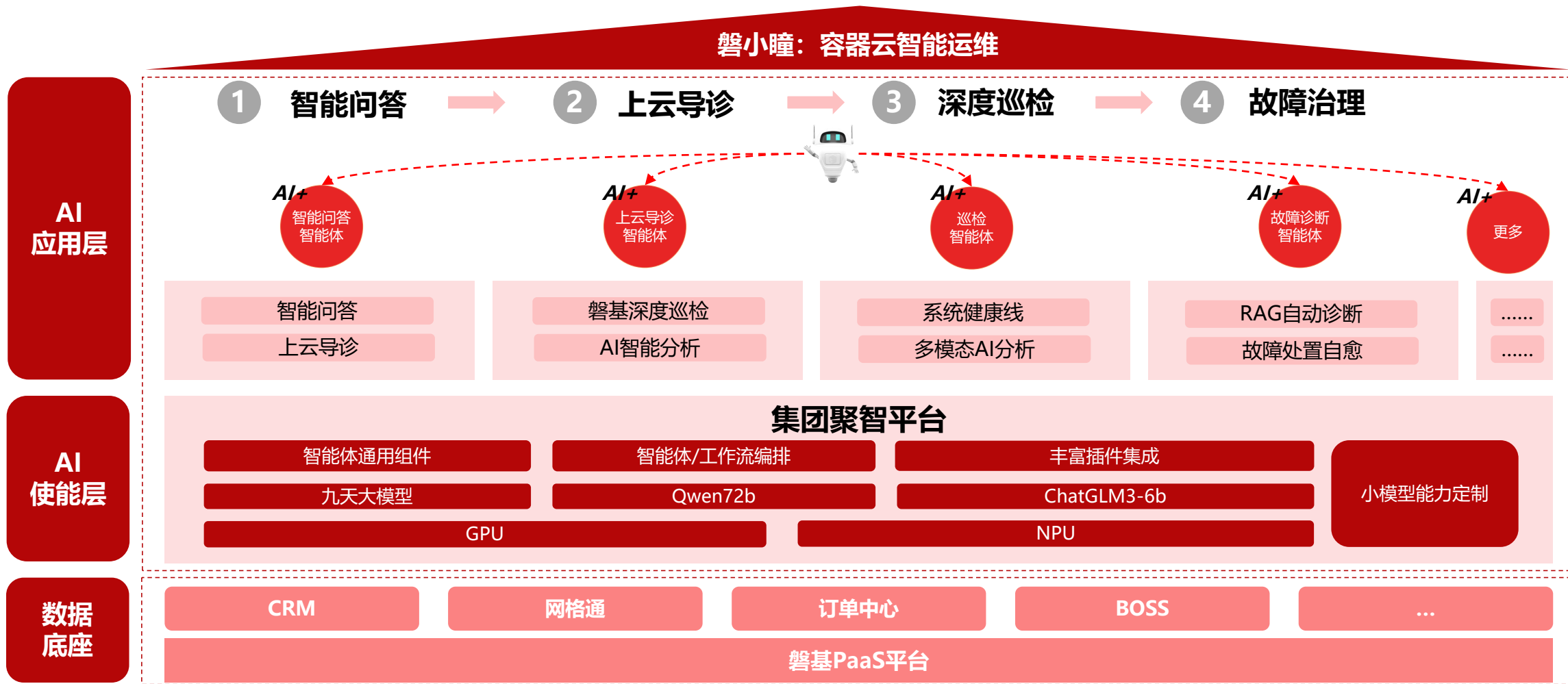


3、基于AI+大模型如何赋能故障诊断和恢复？

利用**大模型智能匹配**对应的自愈工具，进行节点隔离、应用重启、集群扩容等自愈操作，**实现故障端到端闭环修复**。

总体方案：利用大模型技术共筑容器云防线

引入集团公司基于九天和主流大模型的聚智智能体开发平台，自研孵化容器云运维智能体“标品”，实现容器云运维高价值能力，助力全面提升全网磐基运维效能。





依托智能体平台敏捷构建智能体

基于中移聚智智能体平台作为AI使能层，**五步敏捷构建**智能问答、上云导诊、深度巡检、故障治理等多个智能体，解决集、省、专公司容器云迁移及日常运维共性问题。



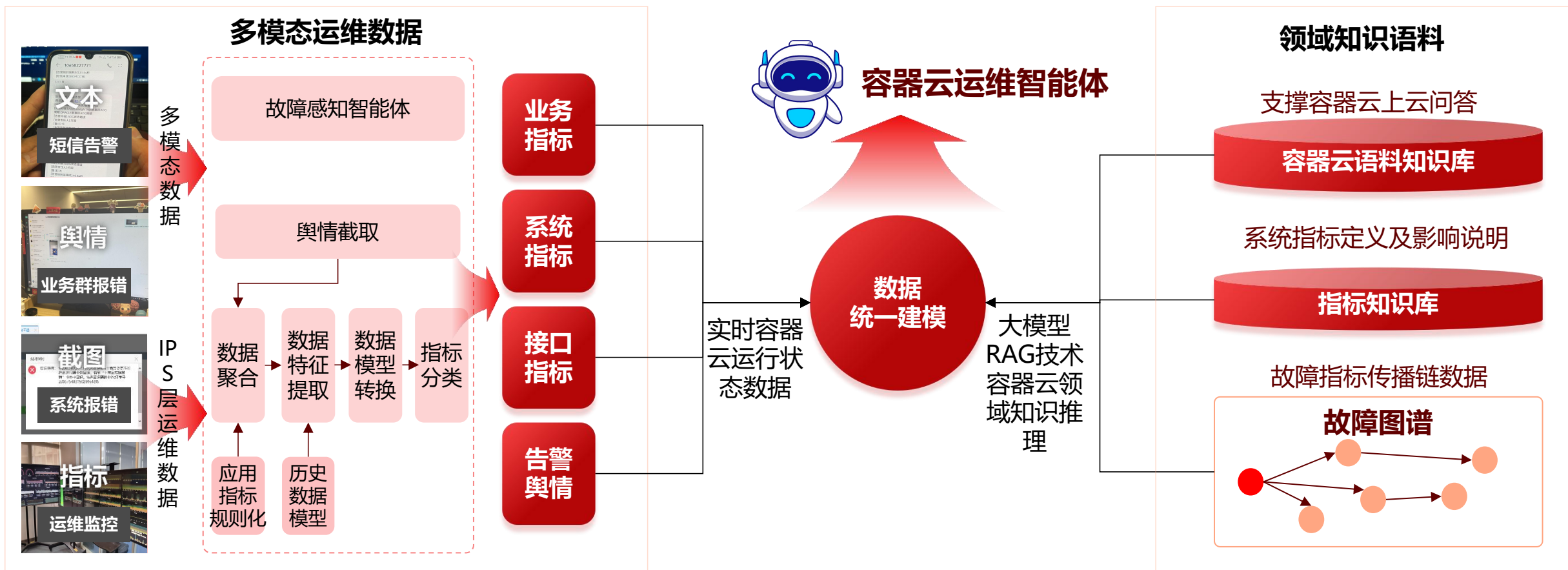
多模态数据融合赋能容器云运维智能体

将全层级的多模态运维系统触点数据和领域大模型知识语料，利用AI算法对全层级多模态指标进行汇聚，**统一建模、转换、智能分类**，供**磐小瞳**进行**训练、推理**。

《《《 监控多渠道信息感知 》》》

《《《 数据统一建模 》》》

《《《 大模型语料标准化管理 》》》



AI智能服务机器人提升上云效率

CoPilot智能服务机器人嵌入容器云平台门户，引导租户自助式上云，**代替人工提供流程解答、上云导诊、租户管理等智能服务**，人机交互助力上云提效80%。

智能服务

沉淀**200+**k8s技术栈语料及典型上云案例

上云流程

使用问题

部署配置

...

上云导诊

封装**40+**通用运维工具集

快速自检

节点隔离

扩容重启

...

租户管理

支撑**10000+**node、pod等管理对象

节点检查

应用管理

纳管确认

...



聚智
底座

API管理平台（在省侧部署，通过应用封装调用容器云接口服务）

聚智平台大模型/智能计算资源

人机交互助力上云提效**80%**

AI深度巡检全面检测容器云异常隐患

针对容器云平台统一技术底座，实现对**基础资源、K8S集群、中间件、数据库和应用**等五类资源进行分析，代替专家进行异常隐患挖掘及诊断分析，解决传统“快照式巡检”无法提前发现系统隐患痛点。

基础资源、K8S集群、中间件、数据库和应用等五类资源

核心服务异常、harbor仓库异常、磐维数据库异常、应用状态异常、中间件异常、pod类状态异常、k8s插件类异常

节点隔离、K8S服务重启、应用隔离、应用重启、应用扩容、harbor仓库恢复、磁盘清理、中间件恢复、配置修改

场景1：5类120+
关键指标保障

场景2：7类90+
异常诊断分析

场景3：9类40+
隐患修复工具

容器云智能体异常获取、诊断分析、自愈

1、任务发起

2、信息获取

3、诊断分析

4、生成处置
建议及报告

5、处置工具

设置定时任务或主动发起

利用 workflow 获取对应信息

大模型根据组合信息进行诊断分析

大模型给出建议并生成报告

大模型调用处置工具

全年15次重保活动0故障

深度巡检时间16小时缩短至5分钟

容器云稳定运行200+天



目录/ CONTENTS

1

AI+支撑思路

2

运维大模型应用实践

3

创新成果

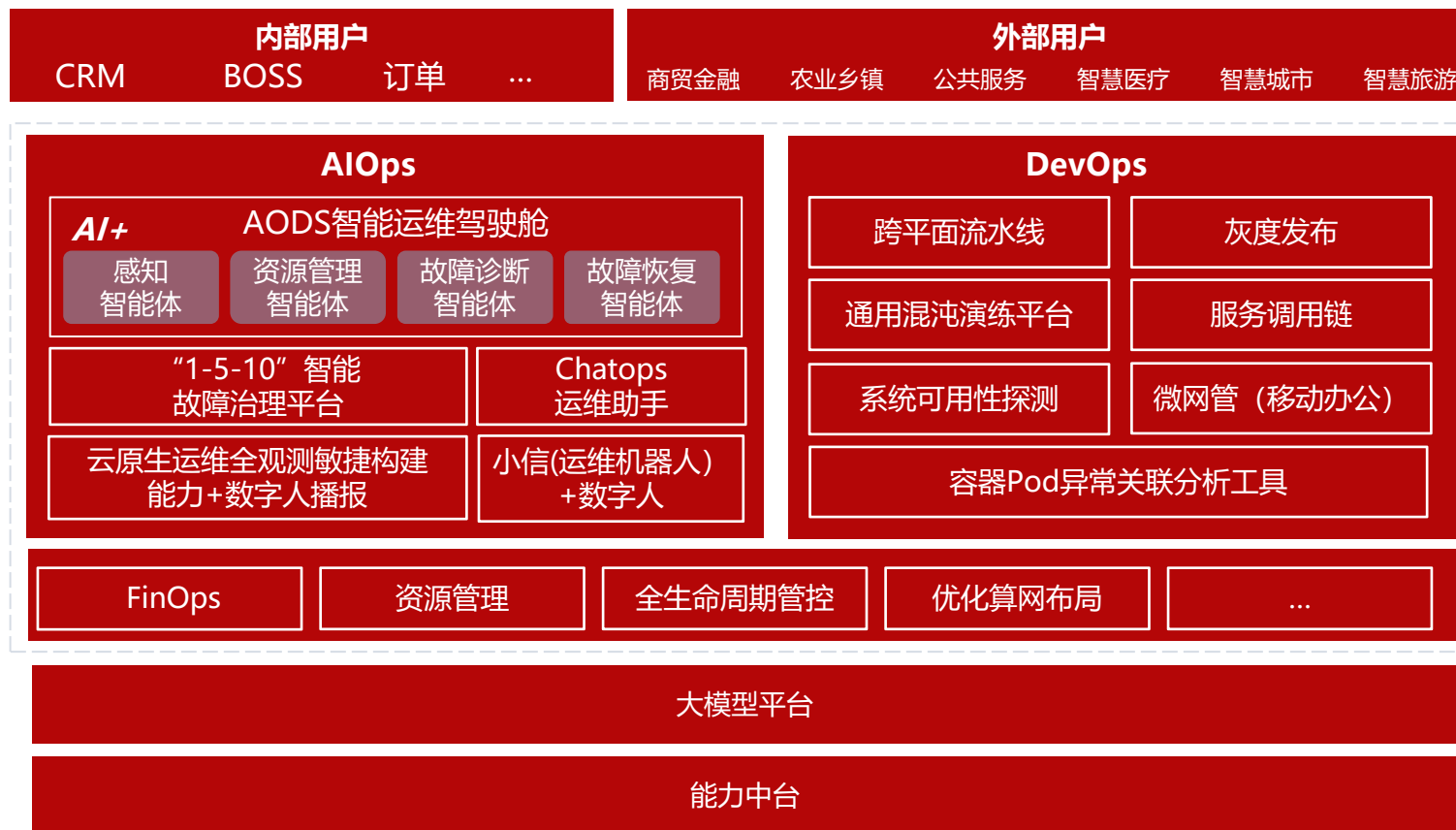


打造大模型赋能的智能运维产品族

1个目标

数智化转型 高质量发展

1套框架



1个底座

更高的客户满意度

更快的上线发布

更稳的系统运行

更优的资源布局

完善智能运维能力体系的构建，打造了一批数智化运维产品，加快数智化转型，促进高质量发展。

T H A N K S

感谢大家观看