

# 线上故障管理规范V1.0

## 线上故障管理规范

### 第一章 总则

为规范外研在线业务系统生产环境故障管理，针对线上故障定级，规范处理流程等，特制订此规范。

本规范适用于外研在线技术共享中心与各事业部技术部。

### 第二章 故障发现

业务系统发布生产环境并验收通过后，公司内、外部发现的问题都算线上故障。故障发现途径：

- 主动发现。研发或者运维查看生产环境日志，或者例行检查监控项时，看到了异常，进而发现了故障；
- 系统监控告警。监控项指标异常，触发告警；
- 关联系统故障追溯。上游或者下游系统的故障处理追溯，定位到可能和本系统有关系；
- 客服问题上报。用户使用时发现出了问题，反馈给公司客服人员，客服上报问题。

### 第三章 故障等级

故障等级	严重程度	描述	处理要求			业务恢复时限
			小组级处理	项目组/业务线级处理	公司级处理	
P0立刻	灾难问题，需要立即处理	1个主业务或多个非主业务产品线系统崩溃、主流程阻塞等问题；影响多校用户	/	问题发生后，项目组/业务线负责人负责问题定位和处理，如果2小时内无法定位问题，需要上报到公司级进行处理	指定合适的人员参与直至问题解决	<1小时
P1紧急	严重问题，尽快处理	1个主业务产品线1个次要流程阻塞；1个非主业务产品线主流程阻塞；影响多校用户	问题发生后，小组负责人负责问题定位和处理，如果2小时内无法定位问题，需要上报到项目组/业务线负责人	项目组/业务线组织人员进行问题定位和修复处理，如果4小时内依然无法定位问题，需要上报到公司级进行处理	指定合适的人员参与直至问题解决	<2小时
P2高优先级	重要问题	1个主业务产品线1个主要功能异常；  1个非主业务产品线1个次要流程阻塞；影响1校用户	问题发生后，小组负责人负责问题定位和处理，如果4小时内无法定位问题，需要上报到项目组/业务线负责人	项目组/业务线组织人员参与直至问题解决	/	<4小时
P3一般	一般问题	1个主业务产品线1个次要功能异常；  1个非主业务产品线1个主要功能异常；影响较多用户	问题发生后，小组负责人负责问题定位和处理，直至问题解决	/	/	<8小时

### 第四章 故障处理

- 故障处理原则：一切以恢复业务为最高优先级。
- 故障处理流程：发生故障>收到&确认故障 >上报至技术群>定位/处理故障>群同步处理进展>故障恢复>故障复盘>改进举措
- 处理过程中的反馈机制：以团队为单位，每隔 10~15 分钟在技术群或临时组建的故障处理群做一次反馈，反馈当前处理进展以及下一步行动。中途有需要立即执行的操作，须事先通报操作对业务系统的影响是什么，由负责人决策后再执行。
- 故障复盘：故障解决后，Confluence填写《线上服务故障统计》表；一周内组织相关人员对故障进行复盘（P0和P1故障必须复盘），建议采用COE分析方法，形成文档存入公司知识库。同时对故障等级进行认定，以及团队成员责任的归属。
- 改进措施：针对当前故障要做哪些改进措施，应对类似问题如何预防等；需要给出可实施的方案以及时间计划。
- 完善系统运维文档：对系统运维文档查漏补缺进行完善。相关人都可以依据运维文档，针对常见故障能紧急处理恢复业务。运维文档常规内容如下：
  1. 管理平台使用文档，包括服务启动/停止/重启等操作；
  2. 服务程序部署、回滚版本操作流程等；
  3. 服务程序部署目录、日志目录列表；
  4. 系统资源扩容操作；
  5. 常见故障列表及恢复手册。

### 第五章 附则

本规范的解释权归外研在线技术共享中心。

### 附件

#### 附件1 线上服务故障统计

序号	影响的产品	故障发现来源	故障开始时间	响应时间	恢复时间	故障等级	影响业务和用户描述	故障原因分析	解决方案	后续预防/改进措施
----	-------	--------	--------	------	------	------	-----------	--------	------	-----------


附件2 故障COE分析方法

项目	描述
标识	＜业务标识＞-＜等级P0/P1/P2＞-＜日期＞-＜顺序号＞
状态	草稿/执行中/结项
问题描述	
影响业务和客户范围	
原因分析（鱼骨图/因果分析图）4-6层原因	
经验总结：规则，经验，知识	
改进措施	
处理过程（指标，图表，代码）	