

采石矶数据质量基础系统

华为云使用说明

文档版本: 1.0

深圳计算科学研究院

版权所有 不得复制

版权所有 © 深圳计算科学研究院 2022。 保留一切权利。

除非深圳计算科学研究院另行声明或授权,否则本文件及本文件的相关内容 所包含或涉及的文字、图像、图片、照片、音频、视频、图表、色彩、版面设计 等的所有知识产权(包括但不限于版权、商标权、专利权、商业秘密等)及相关 权利,均归深圳计算科学研究院所有。未经深圳计算科学研究院书面许可,任何 人不得擅自对本文件及其内容进行使用(包括但不限于复制、转载、摘编、修改、 或以其他方式展示、传播等)。

注意

您购买的产品、服务或特性等应受深圳计算科学研究院商业合同和条款的约 束,本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围 之内。除非合同另有约定,深圳计算科学研究院对本文档内容不做任何明示或默 示的声明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约 定,本文档仅作为使用指导,本文档中的所有陈述、信息和建议不构成任何明示 或暗示的担保。

E	录		ii	ii
1	产品特	点		2
2	产品介	·绍		3
3	使用流	程		4
	3.1	登录华	为云	4
	3.2	选择商	品	4
	3.3	购买商	品	4
	3.4	业务访	问 IP 和端口开通	4
		3.4.1	公网访问业务系统	4
		3.4.2	虚拟私有云内部访问业务系统	5
	3.5	镜像使	用指导	5
		3.5.1	镜像使用	5
		3.5.2	虚拟机登录	5
	3.6	使用系	统	5

1 产品特点

在全球全面利用大数据进行现代化治理的背景下,数据日益融入到政府决策、 社会治理、公共服务、生产制造、消费流通等环节,低质量无序数据存在的危害 风险越来越突出。根据美国数据仓库研究所统计数据显示:数据质量问题每年造 成美国工业界大约 6110 亿美元的经济损失,约占美国 GDP 的 6%,同时 80%的公 司都能从低质量的数据中发现重大的成本改进,因此解决数据质量的需求越来越 大。但目前在全球市场上销售的数据质量主力产品 80%以上是以 ETL 系统为主, 其数据质量规则依赖人工经验配置,表达能力弱,只能支持单表、单行规则,对 数据质量的提升非常有限。同时 AI 在数据质量方面还处于探索阶段,当前成熟 的机器学习模型不具备解决数据质量的所有问题的能力,主流模型比如 Bert, GPT 等含有非常多参数,如何处理海量数据一直是一个难点。为此我们研发了采 石矶系统解决上述问题,该产品以自动管理为特征,融合逻辑规则与人工智能, 支持数据规则的自动挖掘、分析和推理,提供数据错误的自动监测和纠错等功能。

2 产品介绍

本系统在数据质量奠基性理论指导下,通过统一逻辑框架下规则和 AI 的结 合,实现数据内部潜在规则自动发现。支持单表单行、单表多行、跨表规则的挖 掘,并通过规则执行完成数据潜在错误的发现,提供确定性的修复建议,解决数 据一致性、时效性、精确性、完整性和实体的同一性问题。面向集中式/分布式 数据、关系型数据,打造具有可信数据采集,规则发现、数据查错、实体增强、 数据剖析、模型管理、规则管理等功能的一站式数据治理(数据质量)解决方案。

3 使用流程

3.1 登录华为云

使用华为云账号及密码,登录到华为云云商店主页。

3.2 选择商品

1. 在华为云云商店中,选择商品。

2. 在页面搜索栏中搜索:采石矶数据质量基础系统。

3. 点击商品名称,进行商品详情页面。

3.3 购买商品

- 1. 查看商品信息、设置服务选型后,单击"立即购买"。
- 若购买页面无可用的虚拟私有云,请单击"创建虚拟私有云",按照界面提示创建 虚拟私有云后再进行购买。
- 应用包含镜像费用和云服务器费用,具体收费可以通过商品详情页面提供的电话或 邮箱咨询。

可参考华为云云商店的商品购买流程:

华为云文档路径:文档首页> 云商店> 用户指南> 购买流程> 通用商品 购买流程

参考链接:

https://support.huaweicloud.com/ugmarketplace/buyer_topic_0000004.html

3.4 业务访问 IP 和端口开通

3.4.1 公网访问业务系统

当需要在公网访问采石矶数据质量基础系统时,需要购买:

1. 一个公网 IP, 并开放系统业务访问端口 8186。

- 2. 端口 8186 为业务系统的登录端口。
- 3. 业务首页登录的 URL 为: http://业务系统的公网 IP:8186

3.4.2 虚拟私有云内部访问业务系统

当只需要在虚拟私有云内部访问采石矶数据质量基础系统时:

1. 需要开放端口 8186,为业务系统的登录端口。

2. 业务首页的 URL 为: http://业务系统的私有云 IP:8186。

3.5 镜像使用指导

3.5.1 镜像使用

按照华为云指导进行操作:

华为云文档路径:文档首页> 云商店> 用户指南> 商品使用指导> 镜像 类商品使用指导

参考链接:

https://support.huaweicloud.com/ug-marketplace/zh-cn_topic_0203254655.html

3.5.2 虚拟机登录

按照华为云镜像类商品的安全策略要求,禁止 root 用户通过 SSH 登录虚拟 机系统,首次登录可以使用:管理控制台远程登录(VNC 方式)。

按照华为云指导进行操作:

文档首页> 弹性云服务器 ECS> 用户指南> 实例> 登录 Linux 弹性云服务器> Linux 弹性云服务器远程登录(VNC 方式)

参考链接:

https://support.huaweicloud.com/usermanual-ecs/ecs_03_0136.html

3.6 使用系统

基于镜像完成虚拟机初始化之后,可以按照申请的 IP 及端口访问系统。

5