

大数据及人工智能培训云平台

校企数据科学人才培养新模式

黄德演

上海云支柱信息科技有限公司



大数据及人工智能培训云平台

校企数据科学人才培养新模式

演讲者介绍

- 黄德演：云支柱创始人兼总经理
- 2016年回国创业
- 之前是加拿大TERADICI在大中华区的业务总监及合作伙伴的技术负责人
- 过去10年先后参与Vmware私有云桌面的开发和技术支持，亚马逊AWS Workspace的开发和技术支持，所以对私有云和公有云这两个领域都有比较深入的研究
- 云支柱专注于开发企业IT上云服务的管理平台，即混合云工作空间解决方案

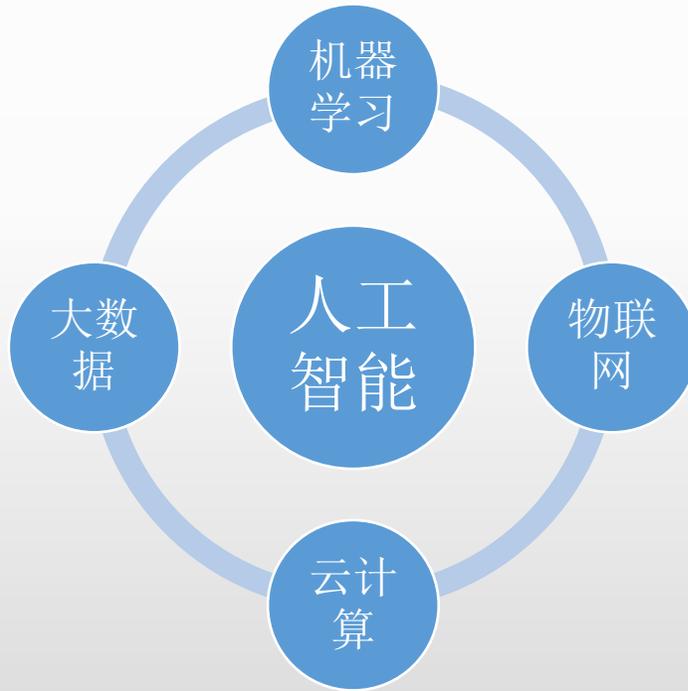
内容概要

- 技术浪潮：基于云计算，物联网大数据，深度学习等技术而兴起的人工智能AI浪潮，
- 社会变化：技术浪潮带来的企业转型和生活模式的变化
- 人才需求：社会变化对新型人才的需求
- 教育培训：人才需求对学校教育培训带来的机会和挑战
- 新型模式：机会和挑战倒逼新型的教育培训模式
- 案例分享：大数据及人工智能实训云平台



大数据及人工智能培训云平台 智能革命：人工智能时代

人工智能的关联技术



三大核心构件





大数据及人工智能培训云平台

智能革命：对企业和社会的影响

企业

数字化转型

服务化转型

- 数字化转型：应用-》数据
- 服务化转型：生产者-》消费者

社会

共享经济模式

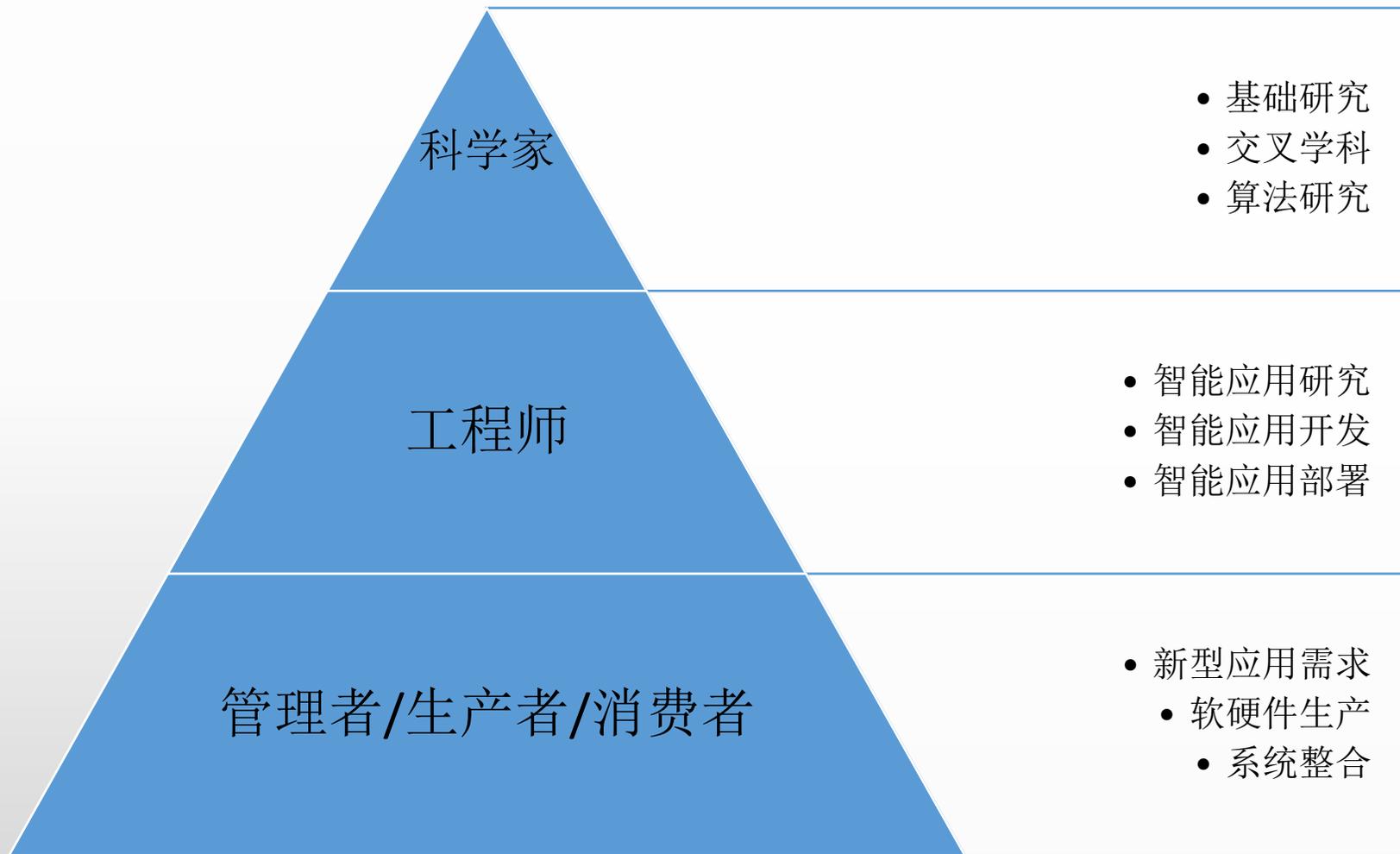
智能化社会

- 共享经济：所有权/使用权分离
- 智能化社会：人/人，人/物，物/物的连接和交互



大数据及人工智能培训云平台

智能革命：新型人才的需求层次





大数据及人工智能培训云平台

智能革命：市场对人才的需求和挑战

企业转型

- 数字化转型
- 智能化转型
- 服务化转型

- 数据应用已经渗透到各行各业
- 传统互联网公司解决信息交互的问题；新型的互联网公司是以数据驱动的应用
- 传统企业在向数字化转型及服务化转型过程中，需要打通企业内部的信息孤岛，连接产品和最终的用户，实现企业到数字化，智能化和服务化的转型

人才类型

- 数据和计算融合
- 软件和硬件结合
- 业务和算法结合

- 企业急需跨学科复合型人才
- 数据科学和计算科学的模糊及融合
- 人工智能产品端设备，边缘计算，网络以及云端的大脑，需要软件/硬件的结合
- 新型应用要结合业务和算法
- 使用软件工具先高级语言编程的转变
- 传统IT向云计算的转变（包括使用，部署，新型的云应用开发等）

市场的挑战

- 人才匮乏
- 自训艰难
- 产业链长

- 大数据及人工智能的人才缺口
- 市场需求巨大；巨头对稀缺人才的垄断
- 企业内部培养缺少师资及环境
- 交叉学科及软硬件的结合的要求延长了企业产业链



大数据及人工智能培训云平台

智能革命：教育培训市场的机会和挑战

市场机会

- 海量企业需求
- 新型专业设置
- 技术迭代更新

- 2016年2月：教育部公布新增“数据科学与大数据技术”本科专业。2016-2017年共批了35家高校
- 2016年9月：教育部公布新增“大数据技术与应用”专科专业，共批了62家高校
- 明年还会增加许多

学校的挑战

- 师资缺乏
- 课程陈旧
- 培训环境

- 师资自身的培养及新型学科的开发需要一定的周期
- 企业和高校在竞争高层次的人才
- 传统课程的设置及基础研究的滞后
- 新技术的快速迭代发展
- 交叉型的学科对教育培训的支撑环境的高要求如计算资源，数据资源，新型的培训体系等



大数据及人工智能培训云平台

智能革命：挑战倒逼传统教育模式变革

挑战 =》 机会

校企合作：数据科学人才培养新模式

学校 -》 企业： 学校基础理论研究及前沿技术研究和企业市场新兴需求的结合

•企业 -》 学校： 企业前沿技术的应用和学校研究方向的结合

•学校 + 企业： 资源共享； 供需对接； 平台共建；



大数据及人工智能培训云平台

案例分享：校企合作打造大数据实训云平台

- 目标：提供具有完整的一站式大数据及人工智能实训环境
- 方法：整合云计算资源+海量数据资源+算法框架环境+培训内容+实训管理平台+师资

合作伙伴

亚马逊创梦实训中心

上海云支柱信息科技

长江光启创新学院

上海数萃大数据学院

实训平台

云计算资源整合

实训云平台管理和接入

数据和培训方法整合

课件 + 算法实训 + 师资



大数据及人工智能培训云平台

案例分享：共建实训云平台的优势

提供校企大数据及人工智能人才培养新模式

- 技术：结合前沿的云服务技术
- 体验：高效及卓越的性能体验
- 覆盖：任意时间地点培训服务

云端
培训

完整
方案

- 整合：按需求整合方案
- 统一：标准化模板化
- 扩展：模块化可扩展

- 响应：极速响应业务需求
- 弹性：即需即用节省资源
- 应用：满足各类高端实训

弹性
资源

专业
合作

- 专业：发挥校企各自专长
- 分工：专业领域快速迭代
- 供需：针对性培训合适人才



大数据及人工智能培训云平台

案例分享：合作伙伴简介

亚马逊创梦实训中心：亚马逊中国唯一授权的实训基地；提供亚马逊原厂的技术支持

上海云支柱信息科技：亚马逊技术合作伙伴；开发并运营混合云工作空间管理平台及接入服务

长江光启创新学院：提供海量数据，各类培训方法及联合开发培训课件

上海数萃大数据学院：开发大数据培训课件和算法研究；提供完整的大数据及深度学习的培训

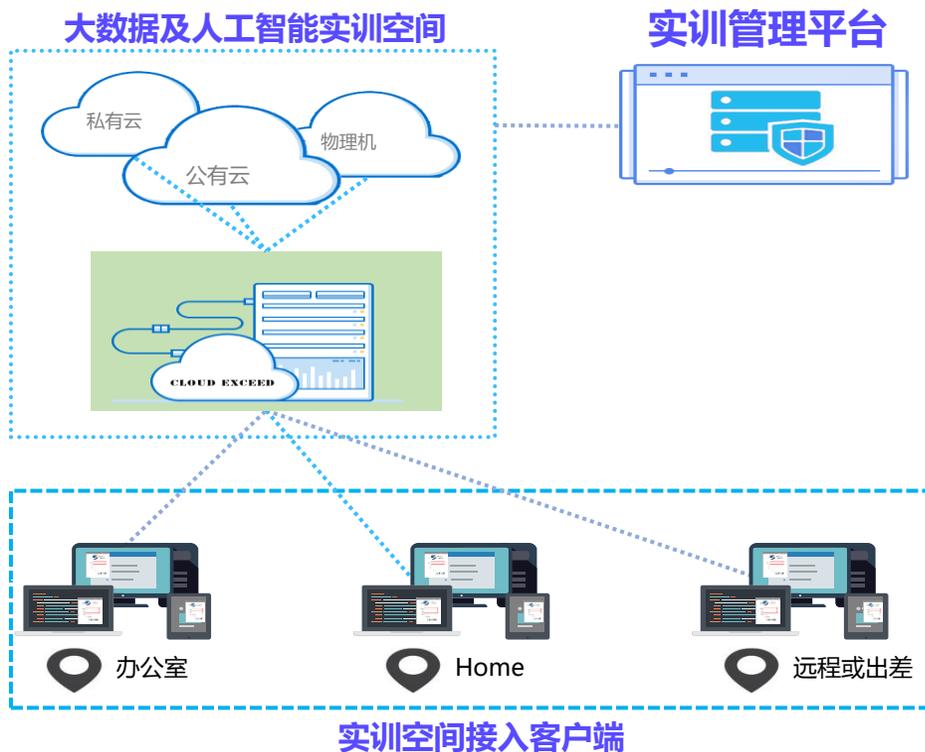


案例分享：诺传实训云平台的架构设计

混合云训空间 + 实训管理平台 + 多样化客户端接入

混合云实训空间

- 高效整合公有云和私有云资源
- 完整的计算 / 数据 / 算法 / 培训课件云端集成
- 不断整合跟踪全球领先技术
- 即需即用高效便捷
- 按使用付费成本低廉



实训管理平台

- 企业级分布式SaaS管理平台
- 一站式混合云资源管理
- 资源 / 数据 / 课件模板化管理
- 实训空间的自动化运维

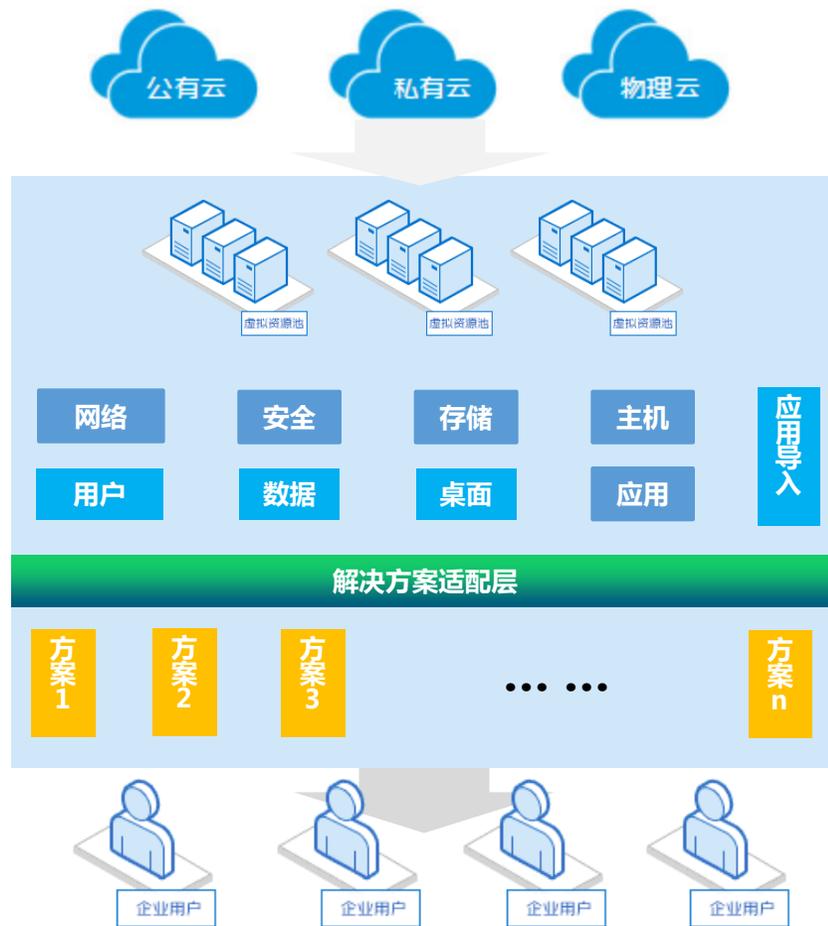
多样化客户端接入

- 根据需要支持各类客户端接入
- 支持专业级应用性能要求
- 随时随地的网络接入提供培训



大数据及人工智能培训云平台

案例分享：诺传实训云平台的技术实现



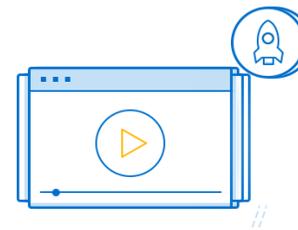
云资源管理平台

整合管理云资源及云服务
提供便捷的自动化管理平台



数据闭环管理

数据的传输存储处理
自主可控安全低成本



超融合接入客户端

随时接入专业工作桌面
保障云服务的安全可控



自助方案定制

用户自助方案建模
统一应用导入接口



大数据及人工智能培训云平台

案例分享：诺传实训云平台的技术特点



云资源的任意搭配

支持公有云，私有云，物理云的任意混搭



培训方案驱动建模

根据实训课件的不同需求配置实训空间，模板化，标准化



高性能远程接入

根据需要提供任意客户端接入如浏览器，专业级远程桌面等



计算资源的灵活调度

海量云计算资源满足各类实训环境的弹性需求



数据的集中管理和共享

收集，清洗和集中归档满足实训平台数据的共享和管理



培训课件和环境的整合

完整的课件和培训环境的整合提供培训的效率和质量



自动化一键部署

一键式部署实训云平台，高效便捷



策略驱动的资源消费

策略设置实现云资源按需消费的自动化控制



单一界面全局监控

单一管理界面和统一接入降低使用门槛和培训效率



大数据及人工智能培训云平台

案例分享：诺传实训云平台的使用流程





大数据及人工智能培训云平台

案例分享：诺传云实训平台的管理：云支柱+亚马逊创梦

- SaaS管理平台
- 自助式服务
- 自动化运维





大数据及人工智能培训云平台

案例分享：平台的管理模块：合作伙伴：云支柱信息科技 + 亚马逊创梦

实训平台管理模块

资源管理

方案管理

应用管理

模板管理

策略管理

工作空间

网络管理

计算资源

用户管理

诺传云

+ 创建主机 + 添加硬盘 + 发送命令 + 发送定制命令 操作 刷新

子网名称	主机别名	主机标识	描述	主机类型	分配用户
subnet-254	rdp-gateway	i-03ac4e89f...	RDP Gateway	t2.medium	
subnet-254	nat	i-06969ca5...	NAT	t2.medium	
subnet-1	WIN-DS3TA...	i-0906a9dd...	dl-training	p2.xlarge	-RDSH-
subnet-1	WIN-5MBU...	i-06a67263...	Template	t2.medium	
subnet-1	WIN-1I41S...	i-0d46f0895...	Test	t2.medium	user1
subnet-253	WIN-I4ND3...	i-0dfaf0654...	AD Server	t2.medium	
subnet-253	WIN-R2QP3...	i-0e00c9bb...	pc-w7-2	t2.large	-RDSH-
<input checked="" type="checkbox"/> subnet-253	WIN-3M80L...	i-0616212d...	pc-w7-1	c3.xlarge	-RDSH-
subnet-253	WIN-EQTO...	i-003d0129...	pc-w7-3	t2.large	-RDSH-

详细信息 性能数据 存储信息 费用信息 命令历史 实例安全组信息 实例会话信息

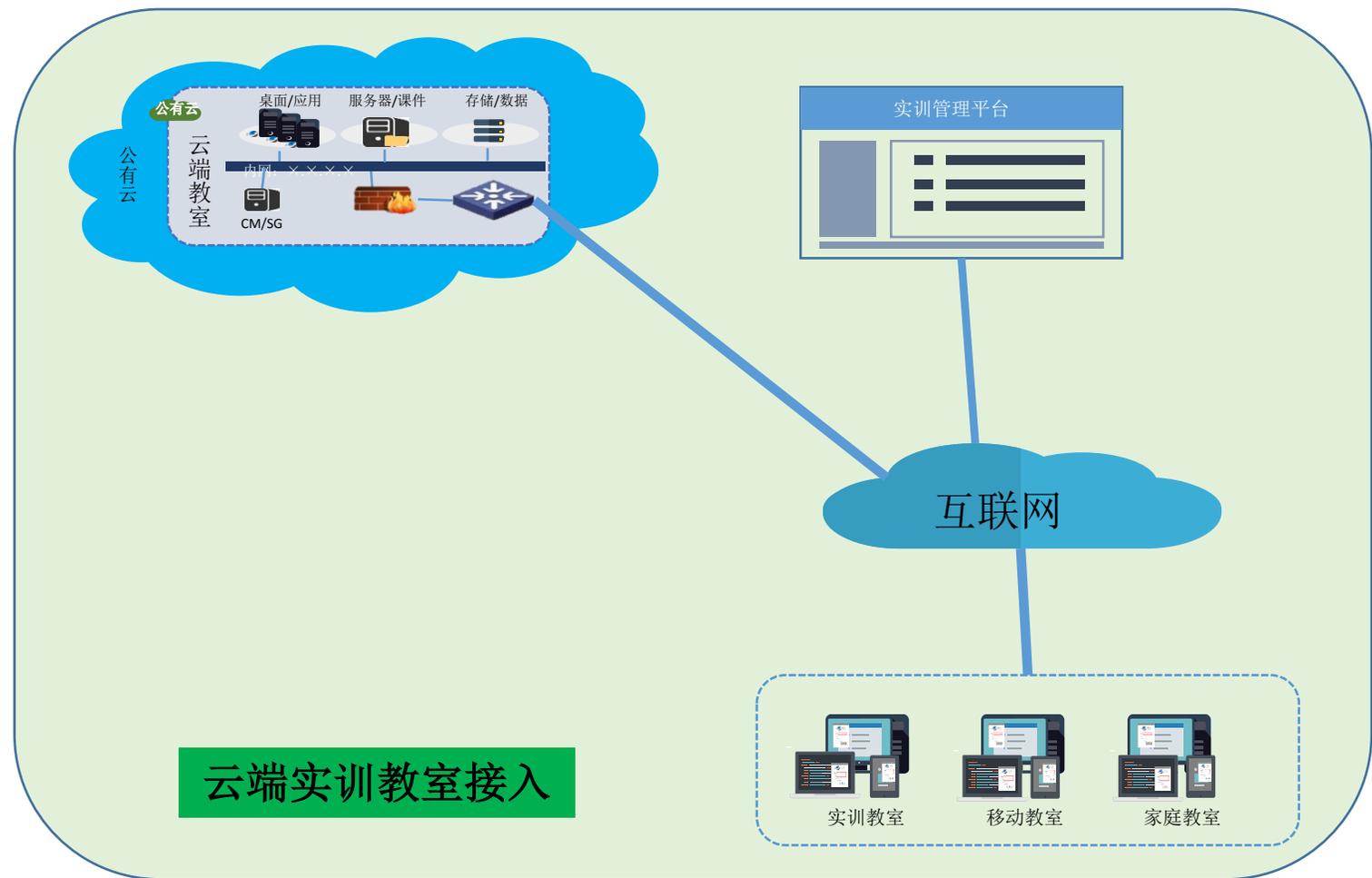
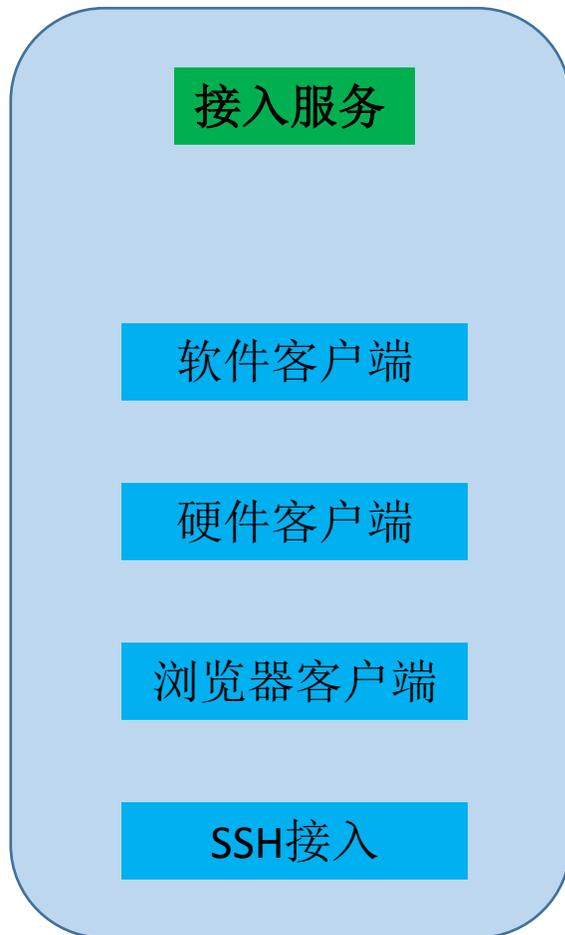
```

SystemState
  PrivateIp 10.1.253.163
  PublicIp
  Description
  NonRootDeviceVolumeIds vol-0d13319bea72fa53c
  PlatformValue Windows
  IsSubnetAdmin false
  IamRoleArn arn:aws-cn:iam::686892099902:instance-profile/CEInstanceProfile
    
```



大数据及人工智能培训云平台

案例分享：实训平台的接入服务





大数据及人工智能培训云平台

案例分享：实训平台的云资源服务：合作伙伴：亚马逊创梦实训中心

IaaS服务



AWS计算服务

- 通用型计算实例（高达64核CPU，256G内存）
- CPU优化计算实例（高达32核CPU，60G内存）
- 内存优化计算实例（高达128核CPU，2T内存）
- GPU计算实例（每个高达16个GPU，64核CPU，488G内存）



AWS存储服务

- 块存储EBS
- 对象存储S3
- 文件服务EFS



AWS网络服务

- 弹性公网IP
- 虚拟私有云
- VPN连接

分析服务



Amazon Athena

数据查询



Amazon EMR

MapReduce



Amazon Kinesis

流数据分析



Amazon CloudSearch

搜索服务

AI服务



Amazon Machine Learning

机器学习



Amazon Rekognition

图像识别



Amazon Lex

语音文本



Amazon Polly

语音识别



大数据及人工智能培训云平台

案例分享：实训平台的实训教室：合作伙伴：亚马逊创梦实训基地

亚马逊创梦实训基地



实训基地云教室



云教室可任意扩展复制



大数据及人工智能培训云平台

案例分享：实训平台的数据服务：合作伙伴：长江光启创新学院

数据搜索引擎



数据库

数据库搜索

- 已经整理的公开信息的数据库（如：统计局、行业、协会、调研机构等）
- 企业自己开放的数据库（如：农业企业、物联网、健康行业等）
- 用户上传的数据库以及国外最新公开数据库（如：kaggle, 知乎用户数据库等）



数据集

数据集文件搜索

- 企业贡献数据集（如：仓储、生产、销售等）
- 用户爬虫数据（如：摩拜单车、知乎文章、大众点评评论等）
- 比赛开放数据集（如：房地产数据，个人征信数据）



程序代码

程序代码

- 公众号推送文章中涉及到的代码
- 培训课程和参考实例的代码
- 企业开源和开放的代码

分析工具和技能

Python和R的在线开发环境

代码学习平台

可视化分析软件
在线学习

数据分析实践

参与面向产业研究的数据分析活动社群

参与面向泛娱乐行业数据分析社群

参与到组建行业分析方向的分析社群内



大数据及人工智能培训云平台

案例分享：实训平台的培训课程体系：合作伙伴：上海数萃大数据学院

完整的“课程体系 + 云平台实操环境”：标准化配置，模板化管理

大数据分析课程体系

- 资深讲师团队
- 高校博士与行业专家
- 完整课程体系
- 面向企业和高校
- 贴心培训平台(亚马逊云平台)
- 支持 R, python, spark, scala

长江龙盾创新学院
数萃大数据学院
2017年11月4日

推荐方案

- 基础培训课程(常年开放)
 - A-1. 大数据导论(陈耀斌、林耀典)
 - A-2. 统计分析基础(刘洪、魏立群)
 - A-3. 数据科学与数据治理(魏立群、李洪)
 - A-4. 大数据基础(魏立群、魏立群)
 - A-5. python 语言基础(魏立群、李洪)
- 中级培训课程(常年开放)
 - B-1. 数据可视化(魏立群、李洪)
 - B-2. 高级统计分析(刘洪、魏立群)
 - B-3. 贝叶斯分析的应用(魏立群、李洪、魏立群)
 - B-4. 统计机器学习(魏立群、李洪)
 - B-5. 网络爬虫与文本挖掘(魏立群、李洪)
- 拓展培训课程(不定期开放)
 - C-1. 商业大数据分析(魏立群、李洪)
 - C-2. 深度学习(魏立群、李洪)
 - C-3. 大数据编程与开发(魏立群、李洪)
 - C-4. 大数据平台技术及应用(魏立群、李洪)
 - C-5. Python 机器学习(魏立群、李洪)
- 大数据实训训练营
 - D-1. 大数据分析实训训练营(Python)(魏立群、李洪、魏立群)
 - D-2. 大数据分析实训训练营(Hadoop)(魏立群、李洪)
 - D-3. 大数据分析实训训练营(Python)(魏立群、李洪)
 - D-4. Web 全栈工程大数据可视化实训训练营(魏立群、李洪)
 - D-5. 量化金融大数据分析训练营(魏立群、李洪)

The screenshot shows a desktop environment with a Google Chrome browser window displaying a slide titled "解放双手万倍提效 文学化统计编程" by 汤银才, dated 2017年11月7日. To the right, the RStudio interface is open, showing the R console with the command `install.packages("kableExtra")` being executed. The RStudio environment pane is empty, and the file explorer shows the current directory structure.



大数据及人工智能培训云平台

案例分享：实训平台的培训课程推广：合作伙伴生态

合作伙伴共同推广培训项目





大数据及人工智能培训云平台

案例分享：实训平台的使用演示: GPU vs. no-GPU比较

如何充分利用云资源达到最好的教学体验

深度学习算法训练 – TensorFlow + No-GPU

437秒

```

C:\Users\stui\Downloads\d1_2d\python>pyth
Extracting ../data/mnist/train-images-idx
Extracting ../data/mnist/train-labels-idx
Extracting ../data/mnist/t10k-images-idx3
Extracting ../data/mnist/t10k-labels-idx1
2017-11-27 13:21:12.026780: W C:\tf_jenki
36\tensorflow\core\platform\cpu_feature_g
't compiled to use AVX instructions, but
could speed up CPU computations.
accuracy at step 0 : 0.0619
accuracy at step 10 : 0.3429
accuracy at step 20 : 0.5591
accuracy at step 30 : 0.6653
accuracy at step 40 : 0.7368
accuracy at step 50 : 0.7765
accuracy at step 60 : 0.8068
accuracy at step 70 : 0.8364
accuracy at step 80 : 0.8474
accuracy at step 90 : 0.8612
prediction is: 6
QWindowsWindow::setGeometry: Unable to se
w/'MainWindowClassWindow'. Resulting geom
, custom margin: 0, 0, 0, 0, minimum size
5>.
    
```

深度学习算法训练 – TensorFlow+GPU

30秒

```

2017-11-27 05:21:12.306156: I c:\tf_jenk
-gpu\py\36\tensorflow\core\common_runtin
2017-11-27 05:21:12.306156: I c:\tf_jenk
-gpu\py\36\tensorflow\core\common_runtin
2017-11-27 05:21:12.309156: I c:\tf_jenk
-gpu\py\36\tensorflow\core\common_runtin
rFlow device (<gpu:0>) -> <device: 0, nam
accuracy at step 0 : 0.1367
accuracy at step 10 : 0.3435
accuracy at step 20 : 0.5745
accuracy at step 30 : 0.6911
accuracy at step 40 : 0.7513
accuracy at step 50 : 0.7884
accuracy at step 60 : 0.816
accuracy at step 70 : 0.8367
accuracy at step 80 : 0.8546
accuracy at step 90 : 0.8627
prediction is: 6
QWindowsWindow::setGeometry: Unable to s
w/'MainWindowClassWindow'. Resulting geo
, custom margin: 0, 0, 0, 0, minimum siz
5>.
    
```



大数据及人工智能培训云平台

案例分享：实训平台的体验课程：深度学习在云端



讲师介绍



史春奇老师

史春奇，日本京都大学博士，美国Brandeis大学博士后。数据科学家，数萃大数据学院讲师，飞利浦亚洲研究院资深数据科学家。12年以上人工智能、机器学习经验。曾在网易有道、百度搜索从事海量文本、图片处理和大数据平台研发工作。

授课时间：2017年12月23-24日（两日）

授课地点：创梦云成实训创新中心（上海市杨浦区国定东路200号1号楼1410）

环境一览





大数据云平台

案例分享：Q&A

欢迎提问