无线城市与大数据

@ 中国移动 2012 移动互联网大会
13:50 08/21/2012 周二 北京国际会议中心 5A

By 邓侃 Ph.D
SmartClouder.com

- Red Lake 55,000 英亩矿区，
  GoldCorp 公司勘探几年，
  始终无法确定金矿矿点。

- 1999 年 GoldCorp CEO，
  在 MIT 得知 Linux 开源项目的传奇，
  决定把 Red Lake 矿区地质资料公开。

- GoldCorp 内部的反对意见，
  1. 地质资料是矿业公司的核心机密，
  2. 工程师们不配合，
     担心被外界揭露自己的无能。

- CEO 力排众议，
  公开自 1948 年始，整个矿区的地质资料。
  拨款 575,000 美元作为奖金。
  全世界各行各业、个人和机构都可以参与。

- 短短几周，来自全球的反馈滚滚而来，
  参与者包括，地质专家、数学家、物理学家、
  计算机专家，甚至还有退役军官。

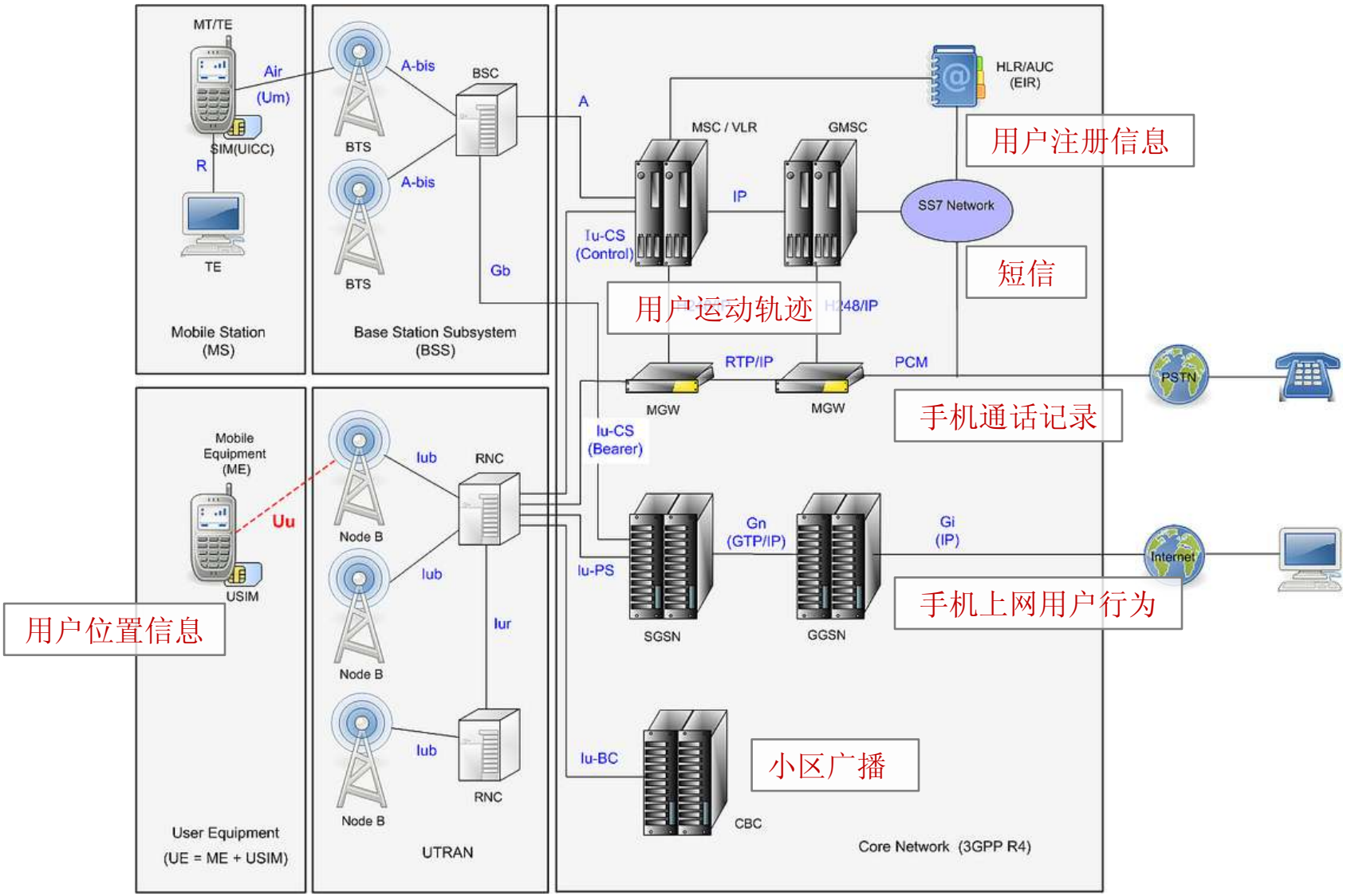
- 来自各行各业的建议，千变万化，例如，
  3D 地质结构视图，让勘探选址更简便。

- 外部参与者建议了 110 个矿点，
  50% 以上先前没有被 GoldCorp 想到过，
  其中 80% 的矿点，开采出 800 万盎司黄金。

- GoldCorp 股价从 1999 年初的 2.50 美元，
  飙升到现在的 36.84 美元，
  2011 年 4 月一度突破 55.00 美元。

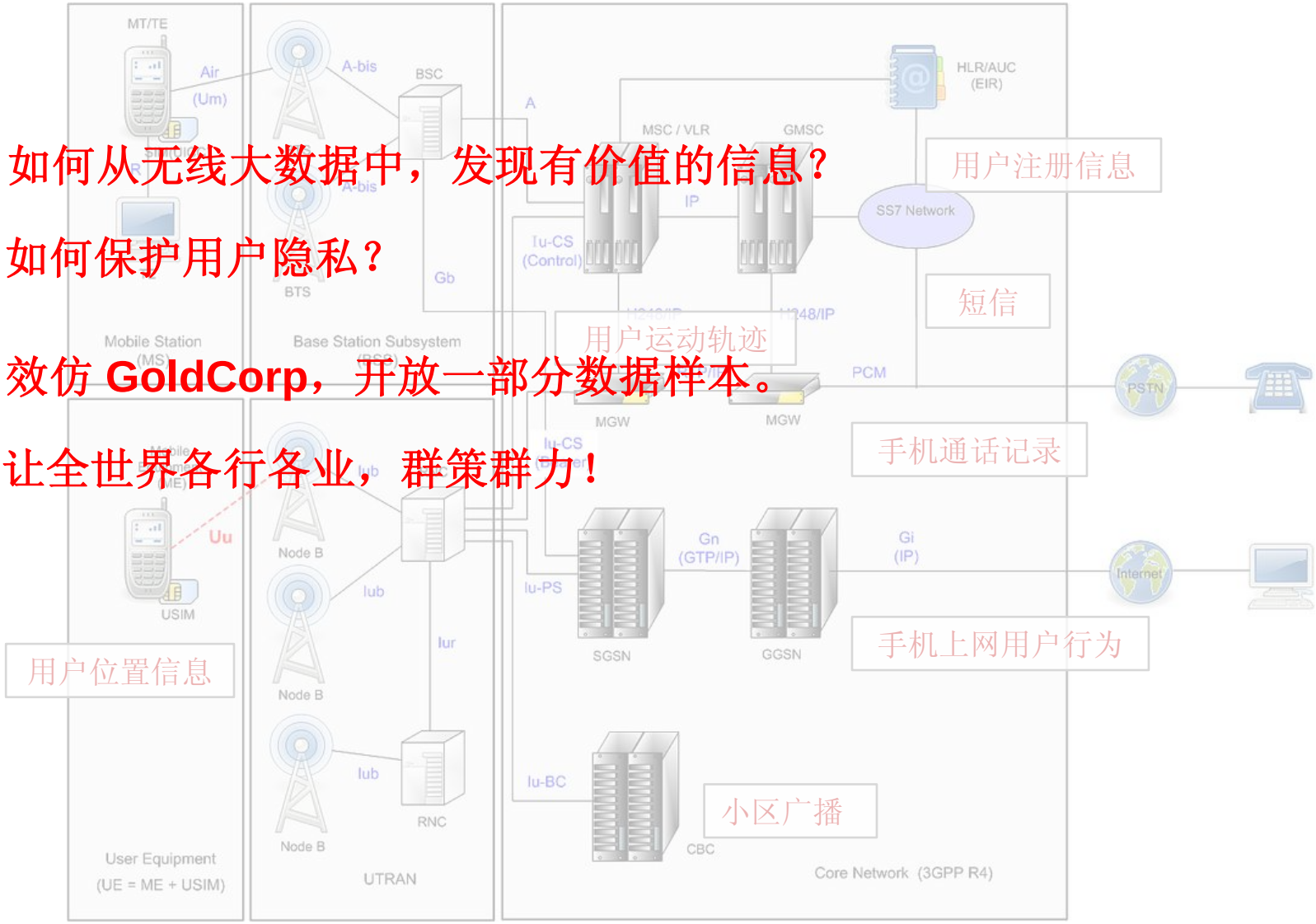
- GoldCorp 启示录:
  所谓公司核心机密信息，
  阻力主要来自于公司内部员工，回避外部竞争。

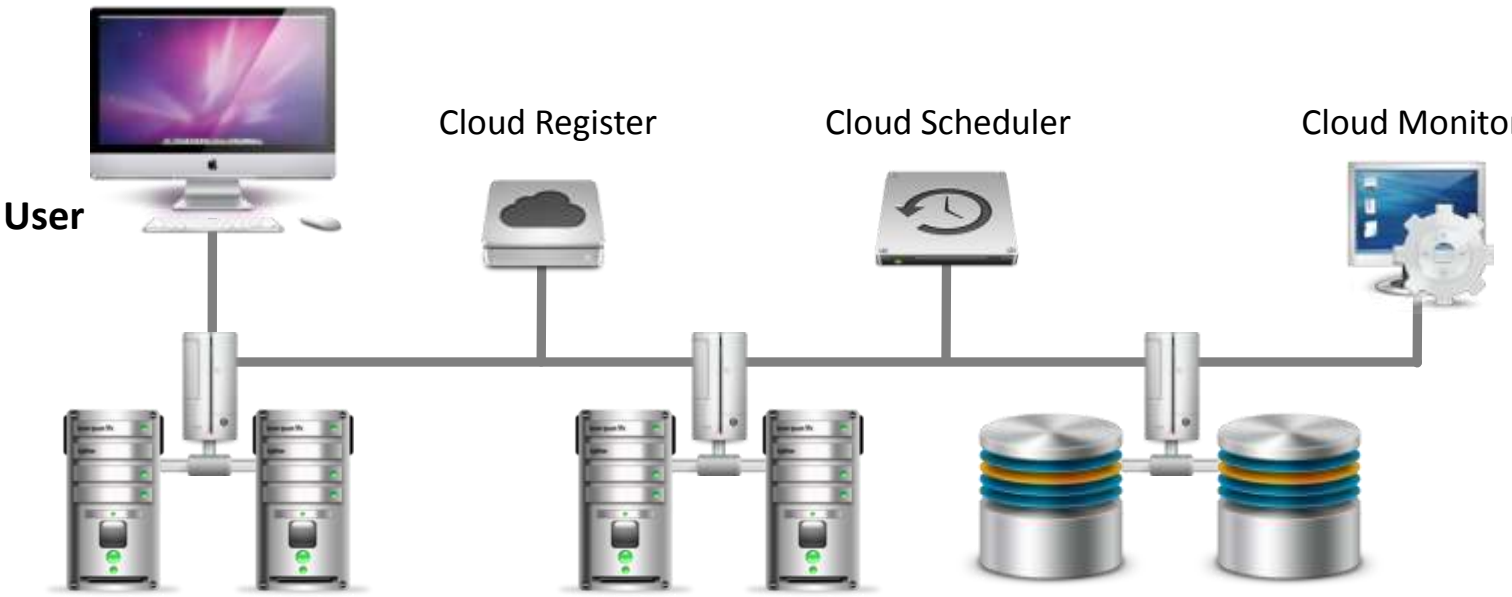
移动网络，蕴含大数据

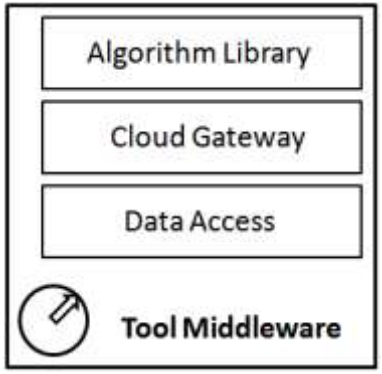

MT/TE
Air
(Um)
SIM(UICC)
R
TE
Mobile Station
(MS)
A-bis
BSC
BTS
A-bis
BTS
Gb
Base Station Subsystem
(BSS)
A
HLR/AUC
(EIR)
MSC / VLR
GMSC
IP
SS7 Network
Iu-CS
(Control)
H.248/IP
MGW
RTP/IP
MGW
PCM
PSTN
Iu-CS
(Bearer)
Mobile
Equipment
(ME)
Uu
USIM
User Equipment
(UE = ME + USIM)
Iub
RNC
Node B
Iub
Node B
Iur
Iub
RNC
Node B
UTRAN
Iu-PS
Gn
(GTP/IP)
SGSN
GGSN
Gi
(IP)
Internet
Iu-BC
CBC
Core Network (3GPP R4)
用户注册信息
短信
用户运动轨迹
手机通话记录
用户位置信息
手机上网用户行为
小区广播

- 如何从无线大数据中，发现有价值的信息？
- 如何保护用户隐私？
- 效仿 **GoldCorp**，开放一部分数据样本。

  让全世界各行各业，群策群力！



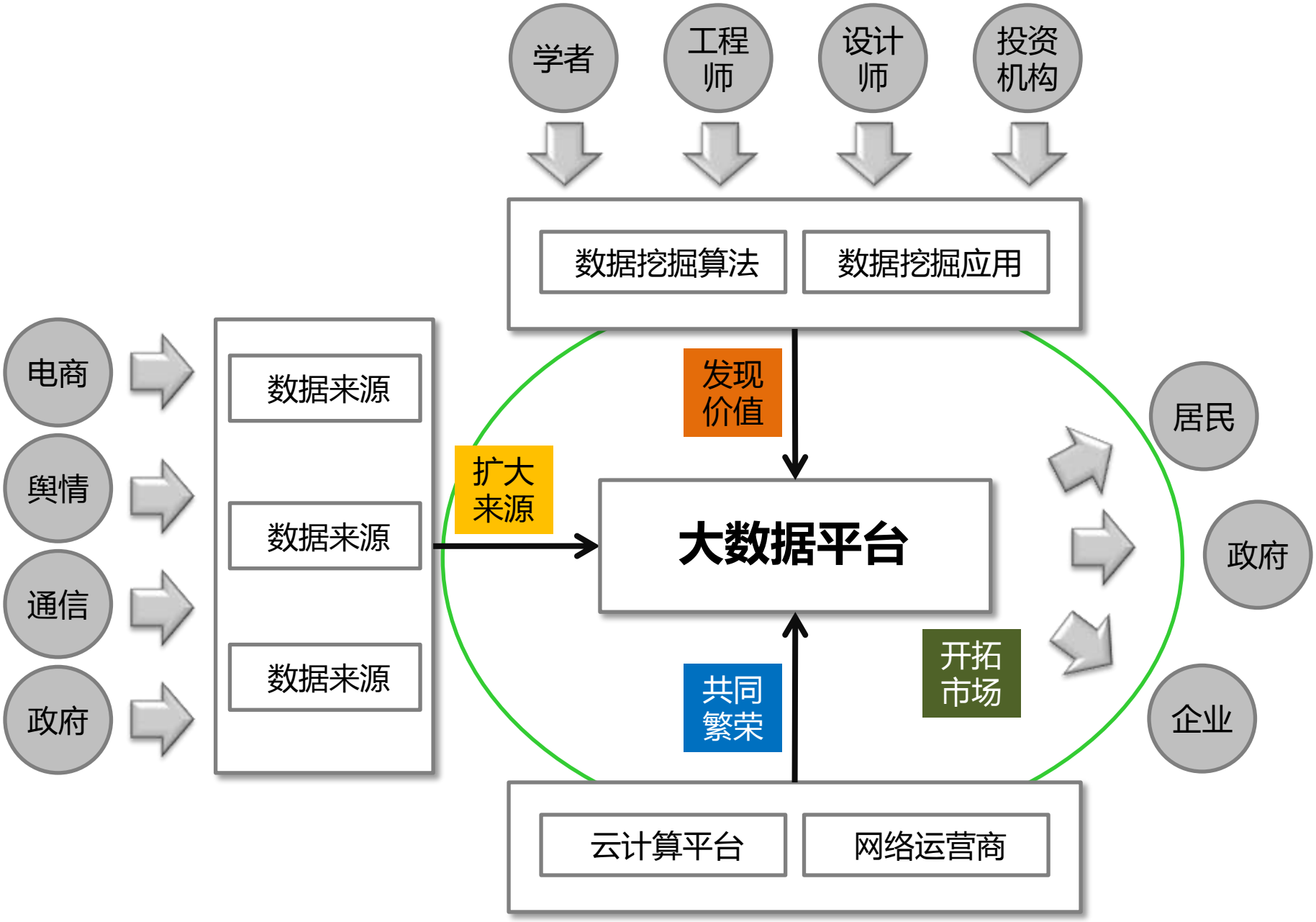


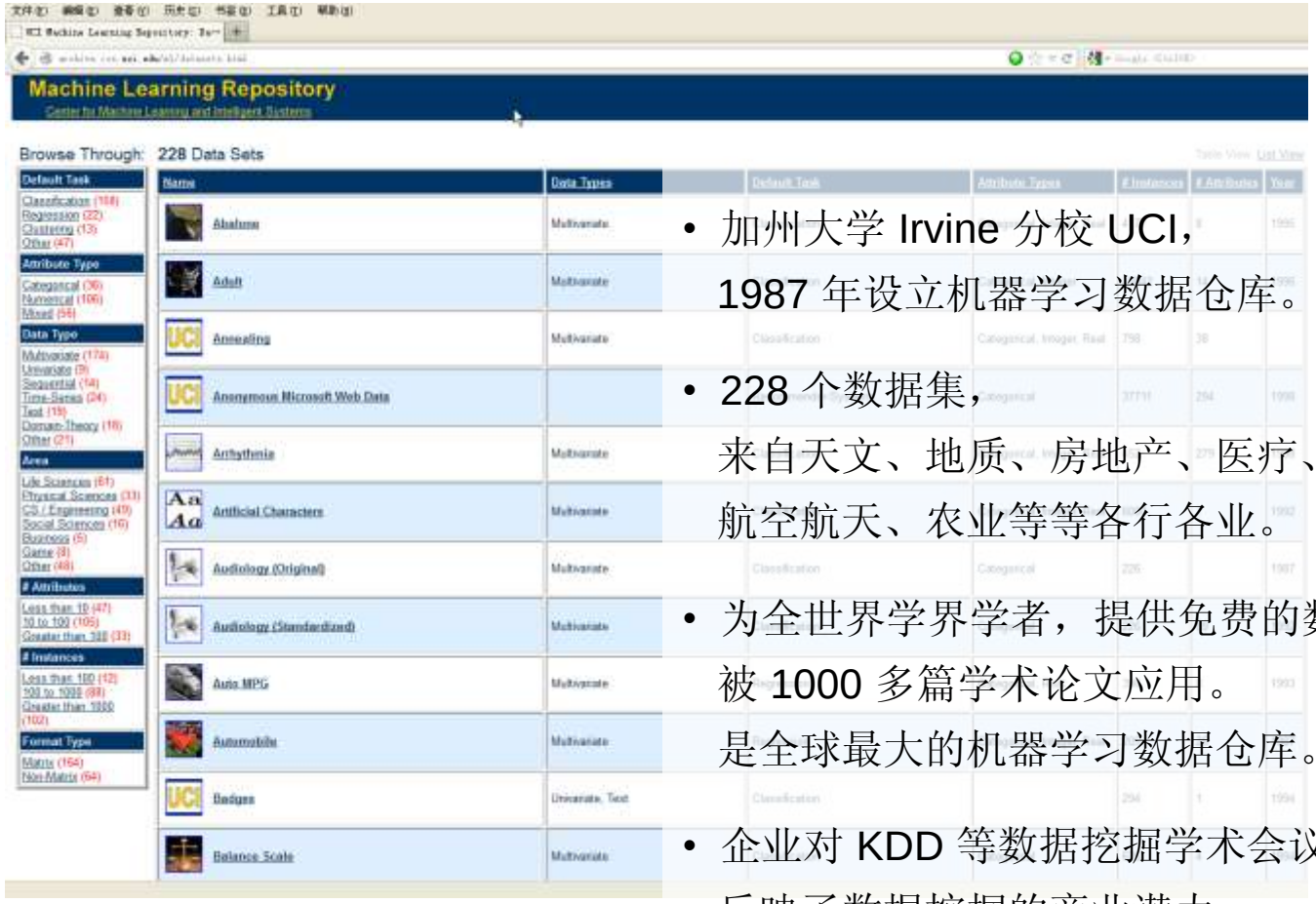


假如中国移动等等大数据来源
能够开放数据样本

假如有一个开放的
大数据存储与处理云平台

假如能够吸引全球各行各业
共同参与共同发展共同繁荣


学者
工程师
设计师
投资机构
数据挖掘算法
数据挖掘应用
电商
舆情
通信
政府
数据来源
数据来源
数据来源
扩大来源
发现价值
大数据平台
共同繁荣
开拓市场
居民
政府
企业
云计算平台
网络运营商



- 加州大学 Irvine 分校 UCI，1987 年设立机器学习数据仓库。
- 228 个数据集，来自天文、地质、房地产、医疗、汽车、体育、航空航天、农业等等各行各业。
- 为全世界学界学者，提供免费的数据下载。被 1000 多篇学术论文应用。是全球最大的机器学习数据仓库。
- 企业对 KDD 等数据挖掘学术会议的踊跃赞助，反映了数据挖掘的商业潜力。

大数据样本捐赠

中国移动将开放部分网络日志。
仅广东移动日信息处理 45 亿条上网日志。

中国移动将开放部分处理过的话单。
话单系统日新增记录 1TB。

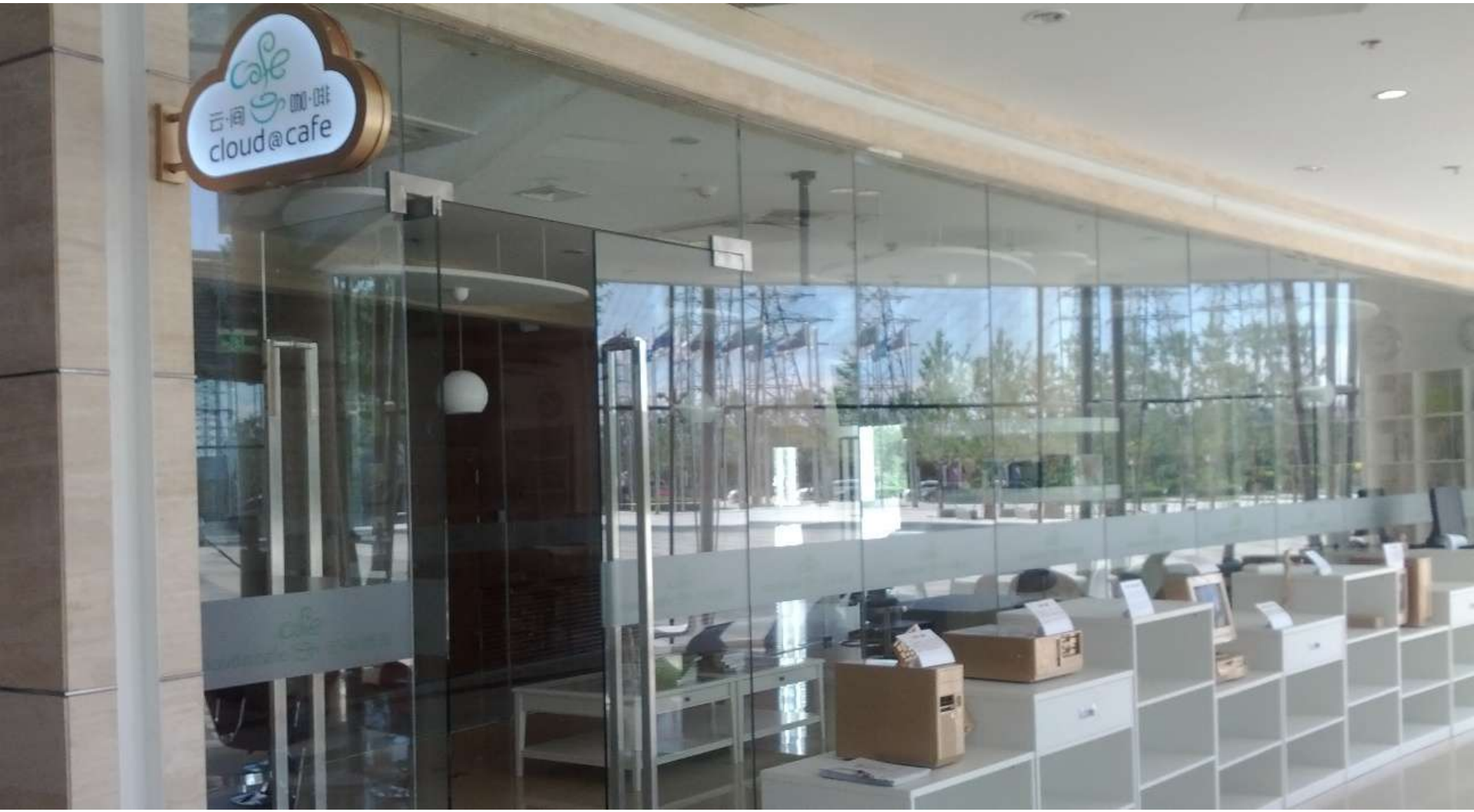
联通网络流量监控。
230 个计算节点的并行计算系统。

城市居民水电煤社保交费记录。
提供 2000 年至今的部分处理后账目。

门诊预约记录
北京各三甲医院的门诊及预约记录样本。

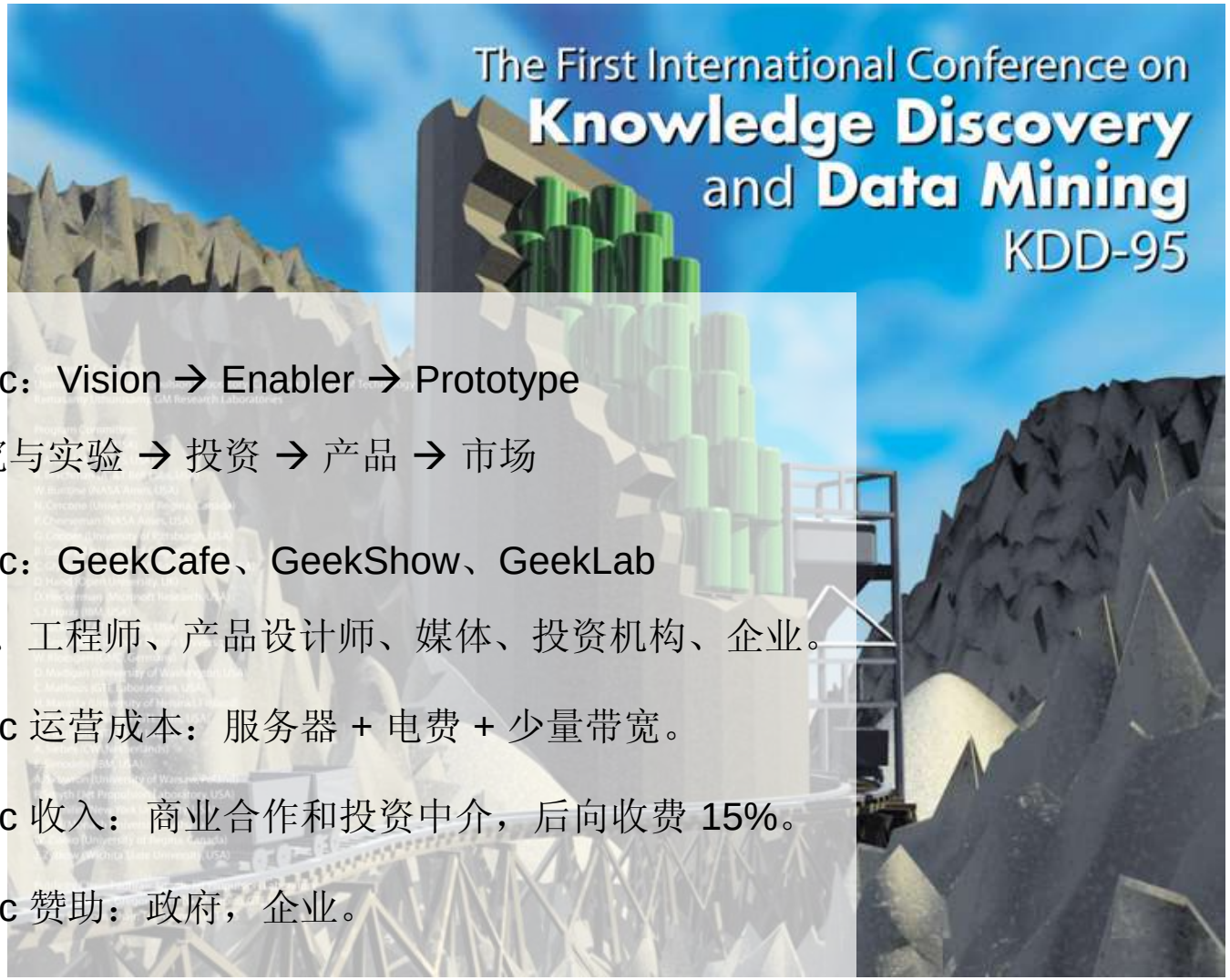
- 中关村软件园，云基地外观
- GeekRepublic = GeekCafe + GeekShow + GeekLab

- 中关村软件园，云基地，GeekCafe



- 中关村软件园，云基地， GeekShow

- 中关村软件园，云基地， GeekLab



- GeekRepublic：Vision → Enabler → Prototype

  需求 → 研究与实验 → 投资 → 产品 → 市场

- GeekRepublic：GeekCafe、GeekShow、GeekLab

  学者、专家、工程师、产品设计师、媒体、投资机构、企业。

- GeekRepublic 运营成本：服务器 + 电费 + 少量带宽。

- GeekRepublic 收入：商业合作和投资中介，后向收费 15%。

- GeekRepublic 赞助：政府，企业。

- 在职员工是否可以加入 GeekLab？
- 如何成为新会员？会员制。
- 如何保护会员隐私？
- 知识产权归谁所有？
- 是否会有法律纠纷？
- 。。。