

下一代互联网宽带业务应用国家工程实验室

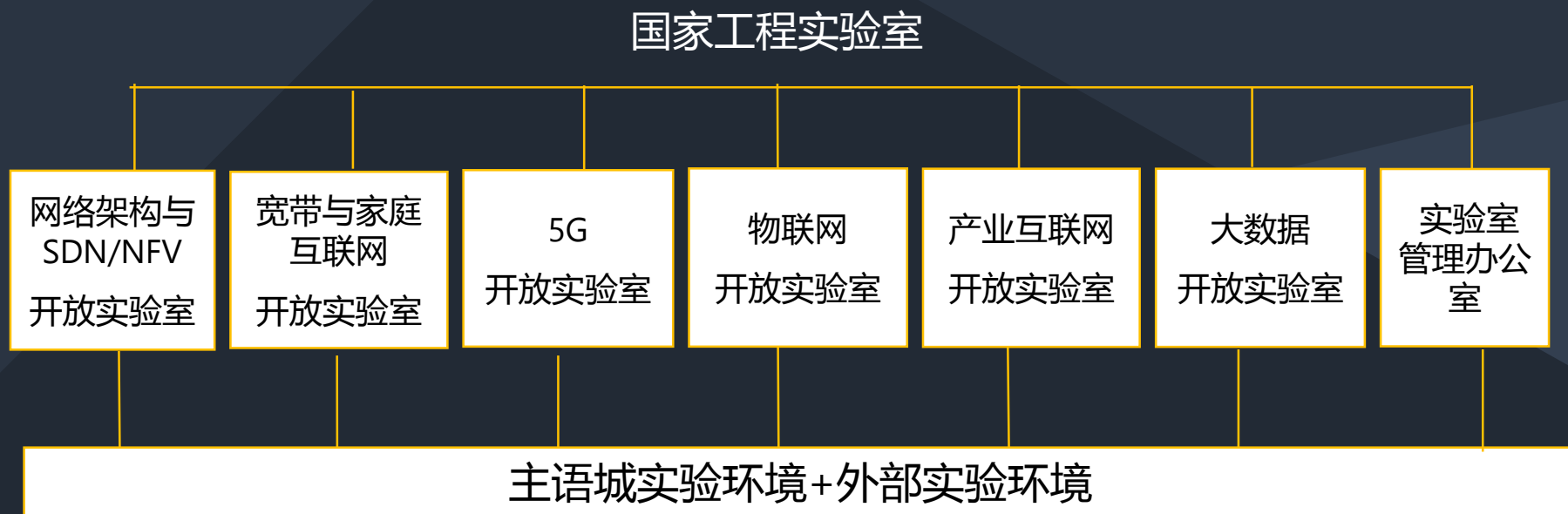
大数据开放实验室介绍

2017年5月

1

下一代互联网宽带业务应用国家工程实验室简介

2006年国家批复中国联通（原中国网通）组建“下一代互联网宽带业务应用国家工程实验室”。由刘韵洁院士领衔，组成了一支具有电信专业实践背景、具备高新技术创新能力的技术专家队伍。实验室贯彻“自主创新”的原则，依托来自国家多个部门70余项国家科技项目，项目总规模近10亿元，同时凭借联通集团的大量项目投入，在大数据、移动互联网、云计算、物联网、下一代互联网、新一代移动通信、三网融合等新一代信息技术领域产生了较丰富的技术与业务创新成果。



2

大数据开放实验室简介

依托于下一代互联网宽带业务应用国家工程实验室，大数据开放实验室以运营商大数据的采集、分析为基础，服务于中国联通网络建维优及内部运营支撑，并致力于运营商大数据在政务民生、交通、医疗、金融、汽车、旅游等行业的应用与合作。实验室目前已具备运营商BSS域和OSS域全量数据联合分析能力，构建开放的大数据分析应用平台，在多个服务行业已经形成完备的解决方案，致力于构建大数据行业应用生态链，实现合作、创新、共赢。



大数据平台、算法、应用

数据存储及分析能力：具备基于Hadoop、SQL的大数据分析能力，具备基于SAS、R、Modeler的数据挖掘能力，已经实现PB级的存储于处理。

数据处理能力：打通原始数据解析，清洗，入库，处理直至上层分析应用的全流程。



产学研合作

大数据产学研合作：北京邮电大学、同济大学数据科学中心、加拿大阿尔伯塔大学大数据研究中心、中科院信息工程研究所、深圳大数据研究院等；

承办年度大数据国际研讨会：ICSINC 2015 big data workshop、IEEE ISCIT2016 big data workshop、5GWN2O17 big data workshop。



人才配置

团队拥有大数据研究人员35名，硕士31名，博士4名。其中大数据研究及分析挖掘人员20名，大数据开发人员15名，团队人才积累丰富的数据分析挖掘和开发经验。



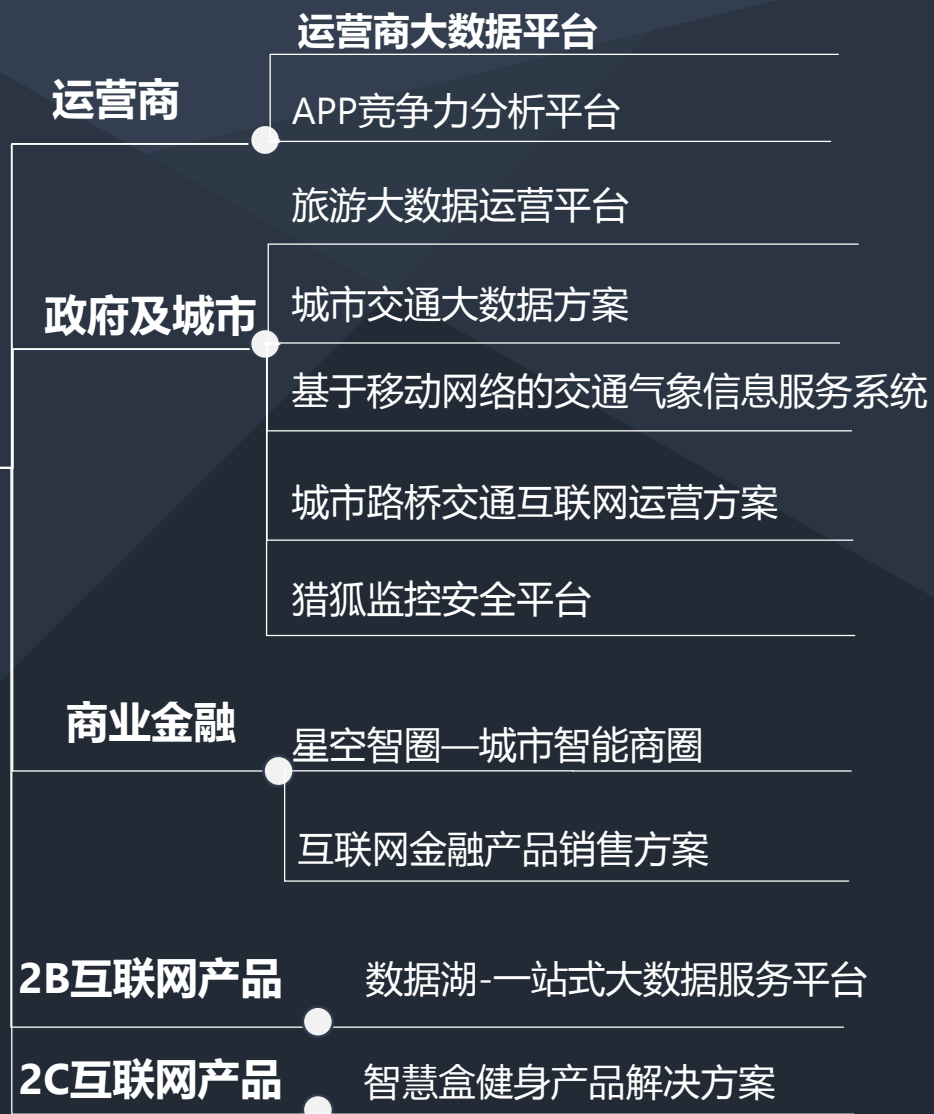
与行业标杆战略合作关系

与南京未来网络签订战略合作协议，在大数据融合、大数据分析 & 挖掘、行业应用等方面进行交流与合作；

与龙信数据签订战略合作协议，研发面向政府市场监管、城市发展规划、金融风险防控、数据创客等领域的大数据产品和解决方案等。

3

大数据开放实验室能力产品体系

大数据中心能力
产品体系

平台能力：大数据开放实验室基于Hadoop集群、MPP集群，实现每日亿级数据采集、解析、预统能力，PB级存储及处理能力。

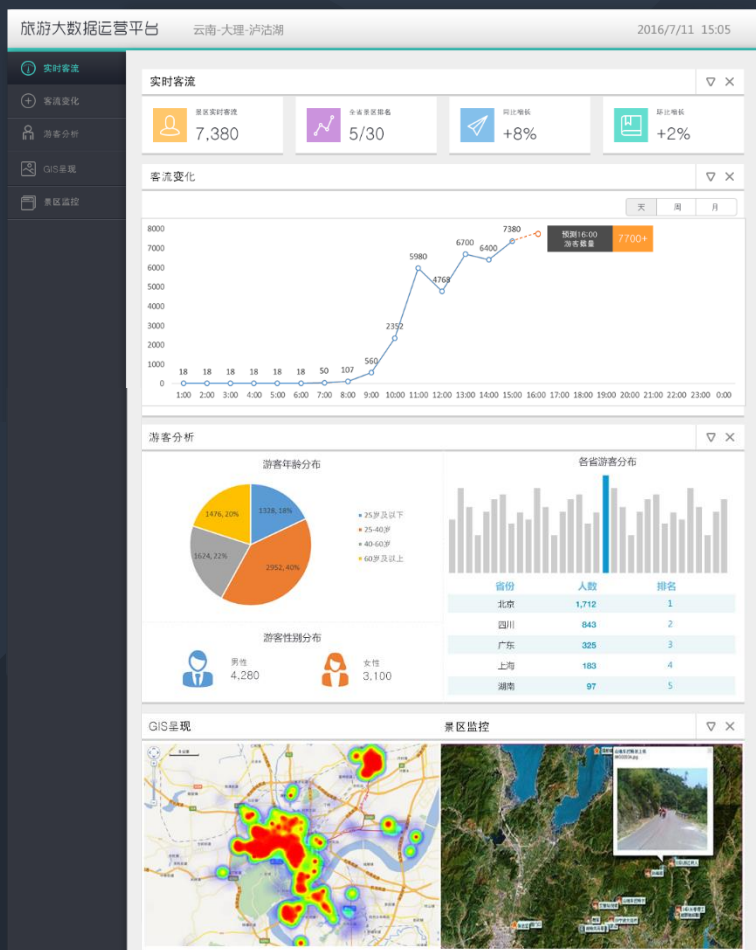
产品体系：深耕旅游、城市交通、商业金融、城市安全等行业，进行运营商数据关联分析。提出了包括旅游大数据运营平台、城市交通大数据方案、星空智圈等11款对应的行业需求解决方案或产品原型。

4

大数据产品案例1：智慧旅游（大数据+旅游）

旅游大数据运营平台

打造**景区监控运营平台**。利用运营商遍布全国的基站，7*24小时准确获取用户的实时位置信息，监控景区游客实时数量，并通过大数据算法对一段时间内**游客数量进行预测**。配合景区监控和人口密度GIS图，为景区的安全工作提供决策支撑。另外，结合运营商掌握的用户标签，对游客进行**旅游画像**，支撑景区的精细化运营。



一站式旅游

结合游客信息和景区线下资源，为游客提供**一站式的游览体验**。



订制服务

打造导游线上资源平台，为游客提供**线上选择导游服务**。



4

大数据产品案例2：星空智圈（大数据+商业）

星空智圈针对商圈、商户和消费者，分别设计了沃·领地、沃·商通、沃·友趣三个产品，为线下商圈提供智能商业服务。星空智圈整合了运营商用户数据（历史位置信息、用户标签、兴趣标签）、线下商圈数据、互联网爬虫数据，对上海市33个商圈进行深入运营分析。通过供需侧的对比，指导商圈供给侧的资源配置；指导商户精准投资；并为消费者提供个性化的消费建议。

商圈产品—沃领地

面向对象：商圈规划部门

产品简介：构造商圈全息画像，分析商圈运营情况。并指导商圈供给侧改革，挖掘商圈潜力

产品模块：商圈规划运营分析、商圈潜力挖掘



配图为商圈运营分析模块界面

商户产品—沃商通

面向对象：单个商户

产品简介：利用大数据分析，帮助线下商户提升投资精准度，提高投资回报率。

产品模块：店铺选址和广告投放



配图为店铺选址模块界面。

消费者产品—沃友趣

面向对象：消费者

产品功能：提供个性化消费推荐和同趣朋友推荐的APP产品

产品模块：自画像、趣周边、趣部落



配图为APP部分截图

4

大数据产品案例3：智能交通（大数据+交通+气象）

结合运营商数据、城市规划数据、气象数据，为城市提供交通出行服务。

利用运营商掌握的用户历史时间-位置信息，结合城市规划数据，得到城市周期性的静态空间分布和时空轨迹分布。再结合高精度、全要素气象预测数据，可以预期一段时间内城市动态全景图，为交通部门提供风险预警管控建议，为政府决策部门提供高精度实时预报预警服务，为个人提供定制化的出行服务。



人口分析

分析目标城市的常驻/流动人口数量及比例、流动人口来源地分布情况。分析目标区域居住人口、工作人口、休闲活动人口密度分布。



人流轨迹分析

分析目标区域高峰时段人口流动轨迹，实时构建城市人群聚集及迁徙地图。



跨城通勤分析

分析目标城市与周边区域的跨城通勤情况。并分析目标城市与其他城市间的通话联系强度分布。



气象交通服务

利用运营商信令数据，叠加气象信息，预测城市拥堵等其他交通情况，为交管部门提供决策依据。



政府气象预警

利用气象实况/预报信息，提供高精度实时预报预警服务，预测气象风险，规避灾害。



个人定制服务

结合个人历史轨迹，匹配小颗粒度的气象信息，为用户提供精细的出行指导。

4

大数据产品案例4：数据湖（一站式大数据服务平台）

“实时数据湖智能分析系统”提供一站式大数据服务，能够汇聚企业内部各类数据，通过实现对异构数据的数据虚拟化，让用户在使用过程中更加得心应手，覆盖企业大数据应用的各个方面，协助企业建立全链条的大数据生态体系。整个产品体系中包含9个子产品，各个子产品之间实现松耦合设计，企业可根据自身需求进行灵活定制。



基于ESB架构的数据采集服务

提供各种异构数据的采集服务，为数据湖各子产品间提供可靠的数据和信令传输通道



云存储服务

为用户提供一种按需分配的“黑盒”云存储服务，配置简单，方便易用



数据解析预处理服务

对数据湖中的结构化、半结构化、非结构化数据进行解析，生成易于分析、存储的结构化数据文件



分布式计算服务

够提供快速、多功能、PB级分布式计算服务，提供高效的数据访问性能及实时分析性能



实时计算服务

能够提供针对流式数据的实时化分析工具，具有较强的扩展性，保证较低的数据延迟



分布式数据库服务

集成OldSQL/NoSQL/NewSQL等多种类型数据库系统，提供多种数据处理服务



元数据管理服务

提供数据资产管理的统一视图，实现数据从数据采集、ETL处理、数据分析和结果展现的全面管控



可视化任务管理及调度服务

提供人机交互友好的用户界面，供用户进行各类任务的添加、配置、删除，轻松管理整个系统中各类任务的管理与调度



商业智能挖掘服务

以多数据源为基础，结合机器学习的方法帮助用户做出商业决策，实现商业价值，包括BI、OLAP、数据挖掘3个功能

5

大数据未来发展方向

以数据开放和数据融合为核心，打造以数据市场为中心的大数据生态，整合多源数据，打造能力平台，创新行业应用。深耕大数据行业应用，特别是以创新治理能力为核心的政府大数据应用，和以产业升级、供给侧结构性改革为核心的行业大数据。

创新行业应用



整合多源数据，
打造大数据能力平台



建设大数据IT
基础架构

行业

智慧旅游 城市规划 城市交通 大数据金融 电信行业 智能医疗
其他政府大数据应用 其他以产业升级、供给侧结构性改革为核心的行业应用

行业定制大数据平台

数据采集 数据解析预处理 云存储 分布式计算 元数据管理
数据统计及挖掘 数据安全及备份 商业智能挖掘 可视化任务管理及调度

基础设施

MapReduce：分布式计算框架 HDFS：分布式文件系统 Hbase：实时分布式数据库
Hive：数据仓库 Pig：数据流处理 Mahout：数据挖掘库 Flume：日志收集工具
Ambari：安装、部署、配置和管理工具 Sqoop：数据库ETL工具

谢谢！

Thank you