

# 2022 腾讯大出行犀牛鸟专项研究计划

## 研究课题

### 目录

1. 定位技术.....	3
1.1. 基于泛在信号的智能感知与定位技术研究.....	3
1.2. 基于图神经网络的定位算法研究与应用.....	3
1.3. 城市动态场景下基于 PPP-RTK 模式的高精度导航定位研究.....	4
1.4. 智能手机 RTK 解算.....	5
1.5. 全遮挡场景下高精度卫星导航定位.....	5
1.6. 基于卫星原始信号和 IMU、Camera 等信号的多源融合算法.....	6
1.7. 基于视觉 SLAM 技术的定位和建图研究.....	6
2. 计算机视觉.....	7
2.1. 激光点云的检测技术研究.....	7
2.2. 激光点云的线要素目标识别.....	8
3. POI 感知.....	8
3.1. 基于时空大数据的 POI 生命周期预测技术和应用.....	8
3.2. 基于运单数据的地址知识图谱构建与应用.....	9
4. 位置服务.....	10
4.1. 基于职住出行的城市地块图谱的推演预测.....	10
4.2. 城市素材库建设.....	10
5. 自然语言处理与语音生成.....	11
5.1. 导航引导语音拟人化合成技术.....	11
6. 未来网络.....	12
6.1. 面向未来交通的确定性网络机制研究.....	12
6.2. 面向未来交通的通信感知融合机制研究.....	12
7. 大数据分析.....	13
7.1. 大数据波动归因算法研究.....	13
7.2. 交通出行领域个性化资讯的自然语言生成模型.....	14
8. 智慧空间.....	14

8.1. 城市空间三维语义分析与计算模型研究.....	14
8.2. 数据驱动的城市生成与评估指标体系.....	15
8.3. 精确建筑能耗与用地属性关联模型.....	16

## 1. 定位技术

涉及研究领域包括：智能感知、深度学习、高精度卫星定位、计算机视觉等。

### 1.1. 基于泛在信号的智能感知与定位技术研究

**研究领域：**智能感知，泛在信号。

**关键词：**泛在信号，智能感知，定位技术。

**课题介绍：**

近年来，随着移动与普适计算领域的发展、各种移动和边缘智能设备持续增加，感知应用的广度和深度得到了不断拓展，涌现了基于泛在信号的智能感知新模式。该技术通常借助于日常使用的移动终端设备（手机、手表、音箱等），利用各类无线信号如 WiFi 信号、手机 UWB 信号、物联网 LoRa 信号、智能音箱声波信号等，提供对个体行为活动和位置的感知能力，服务于多种数字化、智能化、便捷化的感知任务。本项目旨在借助申请团队现有智能感知的研究成果，结合腾讯定位中心对普适场景下智能感知和目标位置估计的需求，实现定位精度的优化，探索多样化感知应用。

**科研目标：**

面向基于泛在信号的智能感知方向探索新技术和新应用。开展对已有 WiFi、LoRa、UWB、蓝牙等无线信号感知和定位效果的对比，研究定位性能提升方法。研究新兴消费类电子设备中 UWB 信号在各类感知应用如定位追踪、呼吸检测、手势识别等方面的应用潜力。产出高水平学术论文 1 篇，专利 1-2 项。

**可提供资源：**

腾讯定位中心部分脱敏数据和服务器资源；

实验验证所需场地。

### 1.2. 基于图神经网络的定位算法研究与应用

**研究领域：**深度学习，图神经网络。

**关键词：**空间定位，GCN，无线信号。

**课题介绍：**

随着移动互联网的发展，基于无线信号(CELL, WI-FI, 蓝牙等)的网络定位方式获得了较大的研究关注和应用。传统方法依赖于 GPS 数据挖掘无线信号的位置和特征，部分场景由于 GPS 信号缺失定位精度较差。基于信号共现关系和上下文信号切换关系，可以将无线信号构建为图

结构数据，描述无线信号关系，由于这部分数据量大，无需 GPS 辅助等优势，可以用于解决 GPS 信号缺失区域的精度问题。图神经网络具有直接对图进行特征提取和表示的优势，可以用于处理上述无线信号网络，提升定位精度。本项目旨在借助申请团队在 GCN 和定位领域的研究成果，将图神经网络、定位信号处理和深度学习结合产出端到端的定位系统，解决实际场景中的定位问题，并提出创新性的解决算法。

#### 科研目标：

结合定位实际场景和数据，构建创新性的定位信号图，并提出新型 GCN 和定位结合方案，在楼宇、地铁等场景上落地，提升定位精度 5%，产出至少 1 篇顶会论文和 2 项专利。

#### 可提供资源：

1. 脱敏定位数据；
2. 测试机器和场地；
3. 计算资源和图计算平台。

## 1.3. 城市动态场景下基于 PPP-RTK 模式的高精度导航定位研究

**研究领域：**高精度卫星定位。

**关键词：**PPP-RTK，网络 RTK，高精度定位。

#### 课题介绍：

PPP-RTK 技术融合了 PPP 技术和 RTK 技术的优点，相比于传统的网络 RTK 技术，PPP-RTK 技术在基准站数量相同的情况下覆盖范围更广，且其播发的数据为具有物理意义的误差改正数，与之相对，传统的网络 RTK 技术播发虚拟观测值，其内含改正数为多项误差之和。因此，与传统网络 RTK 相比，PPP-RTK 技术播发的改正数具有物理意义，且与测站无关，稳定性和外推性都更好。但是 PPP-RTK 技术也有其缺点，由于模型精度问题，其终端算法收敛速度相比于传统 RTK 算法存在一定的劣势，且在硬件支持层面 PPP 改正数不如 RTK 广。另外，由于服务端需要分离各项误差，因此 PPP-RTK 技术其需要较长时间收敛。

#### 科研目标：

提出一种算法，在服务端并行进行网络 RTK 和 PPP-RTK 解算，二者互相校验，互为备份，在正常情况下，播发网络 RTK 结果，在原始物理站数据出现中断时，基于 PPP-RTK 服务计算得到的结果生成虚拟观测数据，利用其改正数外推性好的特点提升系统容灾能力。产出高水平学术论文 1 篇，专利 1-2 项。

#### 可提供资源：

部分实验测站数据、服务器资源、路测设备和车辆。

## 1.4. 智能手机 RTK 解算

**研究领域：**高精度卫星定位。

**关键词：**智能手机，大众应用，RTK，车道级导航，多源融合。

**课题介绍：**

智能手机是卫星导航定位技术大众化应用的重要载体，其定位精度是绝大多数大众位置服务应用的基础。提升智能手机的定位精度，对推动高精度定位服务的大众化具有重要意义，有利于低成本车道级导航、精密农业、智能穿戴等技术的实现。出于对成本、体积和功耗的考虑，智能手机的卫星定位模块在观测质量、稳定性等方面远不及专业设备，极大地限制了定位模块自身的定位精度。目前手机通过安卓标准接口可以透出更多卫星原始观测量信息，以及手机上的 IMU 等传感器信息，通过 RTK 解算和多源融合的方式，提升智能手机上的定位精度，提升手机定位精度能到米级。

**科研目标：**

提出一种手机端输出原始观测量信号，包括载波、伪距等其他安卓标准接口和差分数据的智能手机 RTK 算法，通过状态预测、解算、跟手机内 IMU 融合、滤波策略等算法优化，达到提升定位结果精度和稳定性的目的。产出高水平学术论文 1 篇，专利 1-2 项。

**可提供资源：**

智能手机，差分服务，路测车辆资源。

## 1.5. 全遮挡场景下高精度卫星导航定位

**研究领域：**高精度卫星定位。

**关键词：**伪卫星，RTK，隧道。

**课题介绍：**

在室内、地下、隧道、水下、高山或城市峡谷等环境，微弱的卫星导航信号不足以穿透各种物理遮蔽物，导航终端用户无法使用卫星导航服务，成为“空白区域”。在交通领域中，交通产业快速发展带来的安全问题日益凸显，紧急救援、应急指挥等空间信息服务需求迫切。隧道类环境既不能实时导航、也不能监控定位，还属于救援盲区，容易造成救援延迟，错失救援的最佳时机，造成更大的经济损失，用户因无法提前预知隧道内通行是否通畅而不能提前做出避开隧道绕行的合理线路，造成盲目堵塞隧道交通甚至发生二次事故等问题

**科研目标：**

给隧道内的车辆或手机提供相对连续的位置信号，提供一种室内环境下卫星导航定位增强系统方案：通过即时接收室外区域的卫星导航信号或互联网获取基准信号，并在室内实时生成和播发与室外区域真实卫星导航系统信号相同步的导航信号，解决在不对导航终端进行任何改动下，实现室内卫星导航定位问题。产出高水平学术论文 1 篇，专利 1-2 项。

**可提供资源：**

伪卫星硬件设备、差分服务、测试环境、测试车辆等。

## 1.6. 基于卫星原始信号和 IMU、Camera 等信号的多源融合算法

**研究领域：**高精度卫星定位。

**关键词：**GNSS，惯性导航系统，视觉定位，多源融合。

**课题介绍：**

传统的 GNSS 定位技术在作业时容易受到外界观测条件的限制，在高速运动的载体上或在高架桥下、城市峡谷中及树木遮挡严重的区域，卫星信号的接收能力将急剧减弱，导致无法满足导航定位的要求。近年来，随着移动测量和位置服务等行业的快速发展，单一导航系统已无法满足用户对定位结果的可靠性需求。因此，GNSS 技术与其他导航技术相融合的多源定位技术逐渐发展成为目前导航领域的研究热点之一。目前的导航定位方法除了基于全球导航卫星系统(GNSS)的绝对定位外，还包括无需外界信号输入的相对定位，如基于惯性导航系统(INS)的位置推算和基于图像处理的视觉导航。这类相对定位技术所使用的传感器如惯性测量单元(IMU)和摄像头不易受外界干扰，能与 GNSS 技术相互补充，通过多源融合定位提高定位结果的可用性和可靠性。

**科研目标：**

提出一种卫星信号与 IMU、摄像头等数据融合的多源定位算法，通过对不同传感器在高速载体或城市峡谷等环境下采集的数据进行融合解算，解决卫星信号中断或受干扰时定位的可用性 & 可靠性问题。产出高水平学术论文 1 篇，专利 1-2 项。

**可提供资源：**

GNSS 硬件设备数据，IMU 设备数据，摄像头设备数据，路测车辆资源。

## 1.7. 基于视觉 SLAM 技术的定位和建图研究

**研究领域：**计算机视觉，SLAM。

**关键词：**V-SLAM，特征提取和匹配，语义 VIO。

**课题介绍：**

随着自动驾驶的发展和计算机视觉技术的成熟，车辆对于自身定位的精度要求不断提升，视觉 SLAM 作为机器人自主导航的经典方案，获得了越来越多的关注，在研究深度和创新性方面不断提升。该技术可以辅助车辆在封闭场景如停车场进行同步定位和建图，解决停车场内部卫星信号无法定位的问题。此外，在室外大场景，视觉 SLAM 技术同样可以提升定位精度，并

在道路语义层面进行建模，构建道路语义地图。但目前视觉 SLAM 仍旧面临诸多挑战，如场景变化、纹理缺失、运行效率等问题，因此，本项目旨在借助申请团队现有的视觉 SLAM 研究成果，结合腾讯地图在各种场景高精度定位和建图的需求，探索和实现高精度定位和建图应用。

#### 科研目标：

研究和探索面向封闭场景和室外开阔场景的视觉 SLAM 新技术和新应用。开展对已有 SLAM 框架、语义 VIO 等技术方案的效果对比，研究性能提升方法，探索视觉 SLAM 在自动驾驶领域的应用潜力。产出高水平学术论文 1 篇，专利 1-2 项。

#### 可提供资源

1. 脱敏数据和服务器资源；
2. 实验测试车辆和验证所需场地。

---

## 2. 计算机视觉

涉及研究领域包括：计算机视觉、点云识别。

### 2.1. 激光点云的检测技术研究

**研究领域：**点云识别。

**关键词：**3D 关键点检测，彩色点云，变化发现。

#### 课题介绍：

对整个自动驾驶系统来说，高精度地图是核心技术之一，为自动驾驶提供超视距的感知能力，为车辆决策控制提供依据。高精地图生产需要消耗大量的人力财力，而生产效率的提升则非常依赖自动化算法的能力提升。在高精地图打底阶段，对于各要素的自动化算法的识别精度具有很高的要求，在更新阶段，则需要能够精确的找出点云上发生变化的要素。基于以上考虑，我们提出如下科研问题：

1. 如何利用激光点云数据，在局部点云上精确的回归标牌、红绿灯、杆状物精确的顶点位置；
2. 如何通过图像的数据提升在激光点云上的检测的效果；
3. 如何通过同一位置采集的多趟点云数据，准确的识别出发生变化的位置；

具体研究课题包括但不限于以下方面：

1. 精确的 3D 关键点检测；
2. 相机、激光雷达等传感器融合；
3. 基于相机生成彩色点云；

#### 4. 变化发现。

##### 可提供资源：

加偏的海量高精度点云数据及标注结果。

##### 科研目标：

基于学术界和业界的最新研究，发表点云相关领域的顶级会议论文 1 篇，并在部分场景实现业务落地。

## 2.2. 激光点云的线要素目标识别

**研究领域：**点云识别。

**关键词：**车道线分割与检测，彩色点云，点云与 RGB 图像融合。

##### 课题介绍：

车道线是高精地图业务中的最关键元素，精度是其最重要属性，在城市环境 2D 投影图噪音大，识别容易受车辆残影，护栏等干扰，直接在 3D 点云中检测车道线就非常必要。该课题主要研究如何基于点云深度学习，端到端地进行 3D 车道线检测，实现车道线的超高精度检测，简化算法流程，提高数据质量，加快数据更新速度，极大提高生产效率。基于此我们提出如何科研问题：

1. 如何利用激光点云数据，实现精确的车道线等线要素检测；
2. 如何通过图像的数据提升在激光点云在线要素上的分割的效果；
3. 如何通过多趟采集，提升线要素的分割准确率以及噪声滤除；

##### 可提供资源：

加偏的海量高精度点云数据及标注结果。

##### 科研目标：

基于学术界和业界的最新研究，发表点云相关领域的顶级会议论文 1 篇，并在部分场景实现业务落地。

## 3. POI 感知

涉及研究领域包括：NLP、知识图谱、数据挖掘。

### 3.1. 基于时空大数据的 POI 生命周期预测技术和应用

**研究领域：**NLP、数据挖掘。



**关键词：**时空大数据挖掘、时序预测、深度图网络。

**课题介绍：**

利用多种来源、多种模态的时空大数据情报，去还原真实世界的变化，是提升 POI 基础品质的一种重要手段。其中涉及到深度学习、大数据挖掘、计算机视觉、自然语言处理等多领域前沿技术的综合应用。目前，相关技术的研发分散在多个项目，缺乏体系，应用单一，对情报价值的发掘还有较大提升空间。希望通过高校合作的方式，结合时序预测、深度图网络等领域的前沿技术，统一建模 POI 全生命周期的感知预测算法，包括但不限于 POI 新增、热度趋势、闭店等挖掘，并将成果应用在商业选址、消费推荐、广告投放等多个领域。

**可提供的资源：**

POI 数据集，脱敏后的支付、运单、WIFI 连接数据集，计算资源与平台，标注人力

**科研目标：**

调研学术界最新进展，研究时空大数据挖掘算法，合作发表高质量论文或专利 2 篇，相关技术成果用于新增、过期项目，贡献 POI 覆盖度提升 8%，过期率降低 4%

## 3.2. 基于运单数据的地址知识图谱构建与应用

**研究领域：**NLP、知识图谱。

**关键词：**自然语言处理、知识图谱、大数据挖掘。

**课题介绍：**

地址文本的处理需求广泛存在于地图业务的各个方面，如地理位置服务、POI 数据的质量提升、POI 数据聚合等。而多来源、多范式的地址文本数据往往混有大量错误的、缺失组件的地址数据，地址文本本身的信息不足以支持我们进行深度的分析和处理，所以构建一个地址知识图谱就显得越来越重要。知识图谱这一新兴技术，在遇到地址文本这一场景时，面临了数据巨大、层级关系复杂等挑战。另一方面地址数据也有其独特的特征和需求，需要针对性的对现有知识图谱技术进行调整和拓展。现在相关的研究还方兴未艾，因此希望通过高校合作的方式对地址知识图谱的构建和应用进行探索与研究。

**可提供的资源：**

地址库数据集，脱敏后的运单数据集，计算资源与平台，标注人力

**科研目标：**

调研学术界最新进展，研究图谱挖掘算法，合作发表高质量论文或专利 2 篇，构建地址知识图谱，省市区 100%覆盖，乡镇街道 90%覆盖，主要道路 80%覆盖，地标建筑/区域实体 80%覆盖，主要地名 70%覆盖，从属/附近关系覆盖率超过 80%

## 4. 位置服务

涉及研究领域包括：知识图谱、推演预测、自动化建模。

### 4.1. 基于职住出行的城市地块图谱的推演预测

**研究领域：**时空推演预测。

**关键词：**城市地块知识图谱，推演预测。

**课题介绍：**

城市人口按照地块的空间划分，呈现出地块的轮廓识别，地块的属性识别（包括但不限于：静态属性的 poi 类型、个数、活跃时间、辐射范围及动态属性人口变化、现势性变化等），地块的连接识别（人车物流传播），这些都是知识图谱刻画的数据基础，在这些刻画之上，我们通过学习实体、属性、连接的关系来进行推演和预测。能够有效的针对人口变化对地块的经济结构，人群类别，出访到访时间及需求产生预测能力，这些能力都是未来在智慧城市中对于人的变化所产生的社会效应的一种评估支撑。针对一下问题我们需要关注几个方面的技术构建：

1. 地块的连接，人的轨迹移动在不同地块场景的语义识别和出行链判断；
2. 出行链的完整性，关键节点的识别率，以及出行方式、时间的准确性都是该技术方向的重要指标；
3. 基于属性变化的推演预测技术，目前比较成熟的方案是以 GCN 网络来实现不规则地块的多任务学习预测，未来希望引入社会学/因果推断类模型；
4. 构建 3 是基于群体数据推演和预测，同时展开基于个体的预测和轨迹补全部分，从底层不断完善预测的技术能力。

**可提供资源：**

1-2 个城市地块的空间结构，属性职，职住人口分布及通勤出行时间/方式等群体画像

**科研目标：**

基于地块场景进行地块属性/连接的预测，提升准确率。

### 4.2. 城市素材库建设

**研究领域：**自动化建模。

**课题介绍：**

城市素材库建设需要对城市元素进行拆解，需要对于城市设计原理、城市规划设计、居住设计原理、建筑构造等有较深的理解和应用经验，才能产出更加符合真实世界的城市理解。现有团队更多是地图方面研发和设计，需要更多专业知识的输入来提升该部分能力，而建筑类高

校的引入可以提升该部分的知识，因此，希望可以引入高校专业老师和学生人才，实现建筑类、规划类知识的填充，并实现具体项目的落地。

**科研目标：**

完成具备一定风格特色的 A 城市的建筑风格分析，解构城市构建的数字化和抽象化分析，结合城市建设原则和指导方针，形成建设方案分析报告，并为 WeMap 自动化构建提供可用的城市地标和建筑素材库。

---

## 5. 自然语言处理与语音生成

### 5.1. 导航引导语音拟人化合成技术

**研究领域：** NLP，语音合成。

**关键词：** 自然语言处理，自然语言生成，深度学习。

**课题介绍：**

TTS: Text to Speech 技术，是一种将文字转化为语音的技术。该技术广泛应用于地图导航软件的导航语音播报中。现有的解决方案是：导航软件通过一系列规则，拼接出语音播报文案，再经由 TTS 引擎，将播报文案转换为对应音频。现有方案的主要问题是：语音播报生成内容格式固定，内容单一，长时间使用，使用户对播报内容产生疲劳感，进而影响产品功能使用体验。为了解决播报内容单调的问题，现有方案推出了主题语音功能，虽然可以设定风格不同的语音播报主题，但实现原理过于简单，只是对特定文本进行了简单替换操作，整体播报体验感觉仍然单调，缺乏灵活性。

**具体研究课题：**

1. 对规则生成的导航语音文本，进行单词级别，或者整句级别的同义替换，使文本更具备拟人化特点，同时解决规则生成文本内容单调的问题，以提升用户使用地图产品导航中的使用体验。

2. 根据特定的主题形象，使用深度学习等 AI 技术，对文本生成的音频进行个性化训练，使最终人机交互的语音文本，更符合用户设定的主题形象，以提升用户体验。

**可提供的资源：**

规则生成的导航语音文本，已有主题形象 TTS 引擎效果样本

**科研目标：**

基于业界前沿理论和实践，突破现有 TTS 存在瓶颈，基于深度学习等 AI 技术，提出一种拟人化的 TTS 生成方法，发表高质量专利 2 篇，并结合业务实践，在部分场景实验业务落地。

## 6. 未来网络

涉及研究领域包括：5g+、交通通信网络。

### 6.1. 面向未来交通的确定性网络机制研究

**研究领域：**5g+/交通通信网络。

**关键词：**时延敏感网络 TSN。

**课题介绍：**

目前主要面向有人驾驶的车路协同方案，通过 NTP server 方式对终端进行同步，可以达到 20ms 以内的同步精度，可以满足安全和效率类的应用需求。而远程驾驶要求网络的抖动不大于 10ms，需要借助 NTP server 以外的机制如 TSN 来实现。5G R16/R17 的 TSN 技术面向工业控制以低于 1ms 的同步精度为指标但成本代价太高，不完全适合车联网的需要。

基于以上考虑，需要研究的内容包括：研究提出轻量级的确定性网络/时延敏感网络技术解决方案，使得同步精度低于 10ms。轻量化的确定性网络技术方案，需要放宽设备和器件要求，如对时钟源的要求以及对底层无线网络的性能要求等。同时，轻量化的确定性网络技术方案，也需要在复杂的车联网网络（公网+专网，异构多种网络 4G/5G，V2X 等）环境下能提供性能保障。合作方开始算法研究和方案设计，通过仿真或者现网验证方案的性能。仿真验证：主要由合作方完成。现网验证：需腾讯投入相应的研发人力来配合。

**可提供资源：**

腾讯提供需求及现网测试的人力资源。

**科研目标：**

基于学术界和业界的最新研究，发表 TSN 相关领域的顶级会议论文 1 篇，专利 1 篇，技术报告 1 篇。

### 6.2. 面向未来交通的通信感知融合机制研究

**研究领域：**5g+/交通通信网络。

**关键词：**通信感知融合。

**课题介绍：**

通信感知融合技术是 5.5G 的一个重要方向，从通信的角度，扩展已有通信信号的功能，实现感知能力的延迟，可以实现基于通信终端的测距、测方向、成像等功能，有助于实现通信和感知的协同设计，实现小型化和低成本，在车端以及路侧感知以及交通以外的其他领域也有一定的应用潜力。

基于以上考虑，需要研究的内容包括：通信感知融合的技术架构、算法、方案及其性能评估，包括基于 5.5G 技术以及其他短距通信技术，实现通信感知融合的 PoC 系统，并结合未来交通场景进行落地场景的分析。通过这个研究项目的成果，在感知通信一体化方向的进行前瞻性技术布局。

**可提供资源：**

腾讯需要提供研究的场景需求。

**科研目标：**

基于学术界和业界的最新研究，发表通信感知融合相关领域的顶级会议论文 1 篇，专利 1 篇，技术报告 1 篇及 1 套 PoC。

---

## 7. 大数据分析

涉及研究领域包括：自动化分析、NLP。

### 7.1. 大数据波动归因算法研究

**研究领域：**自动化分析。

**关键词：**归因算法，多维分析，高维数据，根因。

**课题介绍：**

归因分析是用户增长的重要数据工具。特性发布，版本更新，运营活动，天气环境等诸多因素都有可能对业务指标的波动带来直接或者间接的影响，由于现实的因素错综复杂，为真正的根因搜索带来了很大的挑战。业内有不少公开发表的归因算法（Attributor, IDice, HotSpot, Squeeze）每个算法都有其局限性。另外鲜有公开的归因系统，归因系统要对业务产生价值，既需要高效的算法，又需要可扩展的系统架构和科学的属性收集管理方法论，这些都有研究落地的空间。

随着数据驱动的增长模式建立，数据，算法，算力成为了新时代生产力的三驾马车，越来越多的系统，子系统，模块产出了海量的高维度数据。如何正确，快速的理解这些数据成为了学术界和工业界的一个共同的问题。

**待研究问题：**

1. 高效的从高纬度的数据中，找到核心指标的波动根因；
2. 有机结合人工归因，自动归因，产出可理解，可解释的归因结果。

**可提供资源：**

1. 脱敏后的多维度业务数据；

2. 部署算法的相关计算、存储资源。

**科研目标：**

1. 调研和优化学术界，工业界各种主流前沿归因算法包括并不限于 Attributor, IDice, HotSpot, Squeeze;
2. 发表高质量至少 A 类论文一篇，专利两个；
3. 搭建可扩展的归因算法平台。

## 7.2. 交通出行领域个性化资讯的自然语言生成模型

**研究领域：** NLP。

**关键词：** 自然语言处理，自然语言生成。

**课题介绍：**

随着社会发展，出行服务领域用户规模已经突破 10 亿，导航出行 App 提供的个性化服务和资讯越来越多渗透到用户生活中的方方面面。为了给海量用户提供交通出行领域个性化资讯的信息，需要耗费大量的运营人力和财力。对于这个情况，提出该题目：通过自然语言生成技术生成交通出行领域个性化资讯的自然语言文本，为用户提供即时的个性化资讯服务。高质量，生成高质量交通出行领域个性化资讯的自然语言文本。高性能，支持上亿用户级别的个性化资讯即时生成。

**可提供资源：**

相关业务场景数据集、相关算法训练资源。

**科研目标：**

调研学术界与工业界的最新进展，研究交通出行领域资讯等个性化信息的自然语言文本生成模型。研究成果期望发表高质量论文和专利一篇，并在相关业务场景中落地应用。

---

## 8. 智慧空间

涉及研究领域包括：知识图谱、图数据库、计算机视觉、机器学习，碳中和。

### 8.1. 城市空间三维语义分析与计算模型研究

**研究领域：** 知识图谱、数据分析、图数据库。

**关键词：** 三维语义模型，数据建模。

**课题介绍:**

研究表明，地球上至少 60% 的信息都具备地理空间参考信息，地理元素已成为信息的重要组成部分。特别是在社会信息化和网络化高速发展的时代，地理信息已然是众多如环境保护、智慧城市、基于位置服务等应用重要的数据基础。与此同时，地理数据资源也越来越丰富，不仅数据覆盖范围变广(高空、地上、地表及地下全空间)，而且信息维度也不断增加(二维、三维和时空四维)。如何从丰富数据资源中获取正确的地理信息是满足日益增长且持续变化的应用需求的基本前提。通常，获取地理信息的方式有人工采集与模型计算两种。人工采集的地理信息通常是基础地理信息，而需要依赖模型计算提取的地理信息则是复合地理信息。在当前基础地理信息相对丰富的情况下，研究如何利用语义信息辅助计算机正确地执行提取模型，从丰富数据资源中获取准确的复合地理信息变得相对更重要。结合语义分析技术，提出语义支持的地理实例计算模型，研发相应的原型系统，验证地理语义应用于地理模型计算的可行性。

**可提供资源:**

部分园区及城市三维倾斜摄影数据、基础地理数据、BIM 模型数据、点云数据

**科研目标:**

1. 构建以地理概念为核心对象的语义网络，从语义内涵及语义关联表达地理概念语义，实现模型计算的理论研究基础；
2. 重点分析与形式化表达地理概念、概念实例与语义约束的语义内涵及关联；
3. 基于地理知识关联数据库，提出语义支持的地理实例计算模型。通过分析并形式化表达地理概念的语义内涵及语义约束，构建地理知识关联数据库，并以此为知识基础研发语义支持的地理实例计算引擎，提供基于语义信息的地理实例提取模型计算。
4. 基于概念语义内涵(以本体属性表达)的数据语义有效性验证模型研究；
5. 设计并研发基于计算模型的原型系统，验证模型设计的合理性。原型系统实现包括语义形式化表达、语义约束构建、知识链构建与计算等功能的实现；

## 8.2. 数据驱动的城市生成与评估指标体系

**研究领域:** 计算机视觉、计算设计、城市数据模型。

**关键词:** 参数化设计，程序化城市模型，图形算法。

**课题介绍:**

城市规划与城市设计导则方案设计中，涉及大量规范约束的城市设计生成。如获得某地建筑设计规范（如防火分区、护栏高度、女儿墙高度等）、场地设计规范（如消防通道、防火间距、疏散面积等）、城市道路设计规范（如场地出入口、公交站点路口间距等）。因此由经济发展或生态环保为指标导向的城市规划方案设计中，需要用经济指标（容积率）或生态指标（绿

化率)等做为方案生成指向型参数,并自动满足各类用地属性中的建筑设计规范、场地设计规范等,快速生成城市设计方案并可视化。其难点在于建筑间的布局方式、设计美感、建筑平面与采光面积平衡、日照间距约束等。

同时,量化城市设计也需要一套评估指标体系进行多场景规划比选的依据,因此数据驱动生成建筑与用地的成果指标(建筑总面积、建筑日照立面面积等)也应为一套评价标准。在城市规划设计院、设计事务所、房建企业拿地等日常项目中各专题规范已有指标体系的探索。并通过设计师使用过程的用户反馈与干预、沉淀行业机理模型。

**可提供资源:**

部分城市基础地理数据、BIM 模型数据。

**科研目标:**

1. 构建以规范数据驱动的建筑生成算法工具链原型;
2. 构建城市专题规划指标评价体系;

### 8.3. 精确建筑能耗与用地属性关联模型

**研究领域:** 机器学习。

**关键词:** 节能减排, 机器学习, 碳中和。

**课题介绍:**

我国将推进 2030, 2060 碳达峰、碳中和双碳目标,并强制推进建筑能效优化与主被动节能、超低功耗建筑节能措施。除此之外,用地属性与能耗的关联一直是学术界难以解决的课题,原因是拥有准确建筑各系统能耗的业主与拥有区域用能数据的国企数据并不打通,并且各数据模型建立时采用了不同的数据对象规范。

对建筑能效(分系统)与用地属性关联进行机器学习,可得出初步建筑碳排放与用地属性关联,并提供城市规划进行经济发展与生态保护之间的平衡,避免出现无法精准决策两者城市指标体系的困难。可联合权威科研院所,发布建筑能耗与用地属性等关联模型,并绑定行业机理模型的规范性。

**可提供资源:**

部分建筑、园区的精准用能情况、区域用能数据、部分城市基础地理数据、BIM 模型数据。

**科研目标:**

1. 发布建筑能耗与用地属性关联模型与研究报告;
2. 输出碳排放与经济发展优化算法工具包原型;