



北京城市实验室  
Beijing City Lab



ID of the slides

75

## Slides of BCL

[www.beijingcitylab.com](http://www.beijingcitylab.com)

### How to cite

Author(s), Year, Title, Slides at Beijing City Lab, <http://www.beijingcitylab.com>

E.g. Long Y, 2014, Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest, Slides at Beijing City Lab, <http://www.beijingcitylab.com>

# 大数据，新范式和大协作： 决策导向的大数据应用范式和案例

(大思小事)

周江评

澳大利亚

昆士兰大学

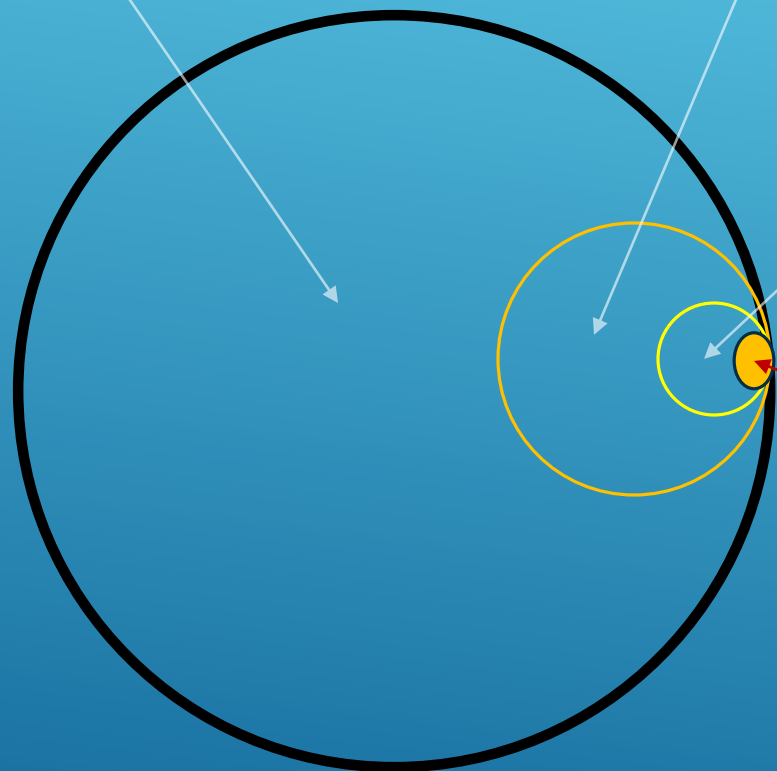
jp.zhou@uq.edu.au

- 数据优先，还是决策(规划)优先？
  - 我们为何决策？
  - 我们该怎样决策？
  - (大)数据应可以怎样帮助决策？
  - 决策导向型的大数据应用范式
- 以上，大思考
- 案例
- 以上，小事情
- 小结

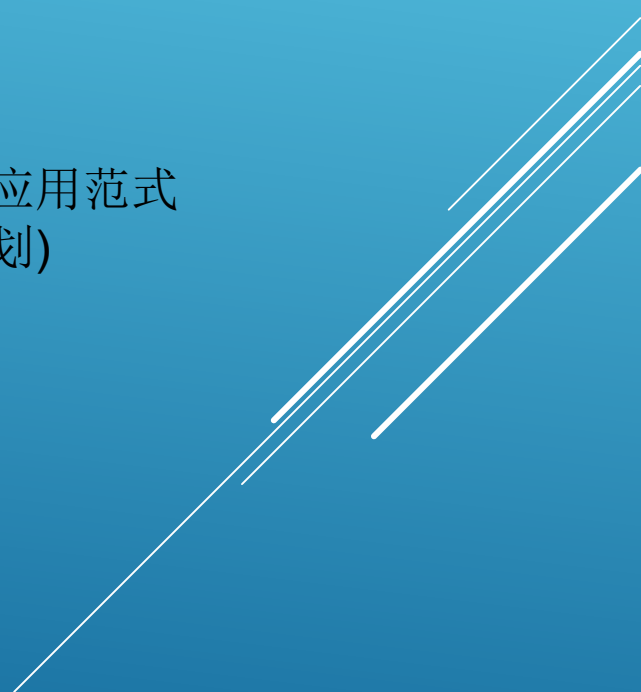
需要决策和规划的事情

涉及数据

涉及大数据



决策导向型的大数据应用范式  
(偏重于城市规划)



数据优先，  
还是决策(规划)优先？

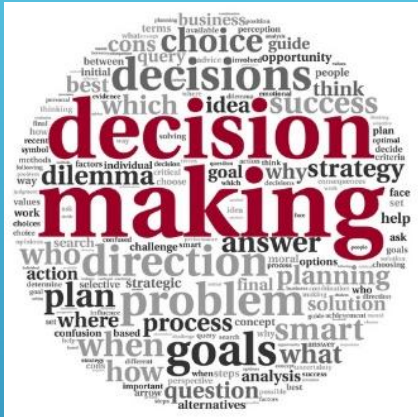


“你千万不要想着拿数据去改进一个业务，这不是大数据。你一定是去做了一件以前做不了的事情。”

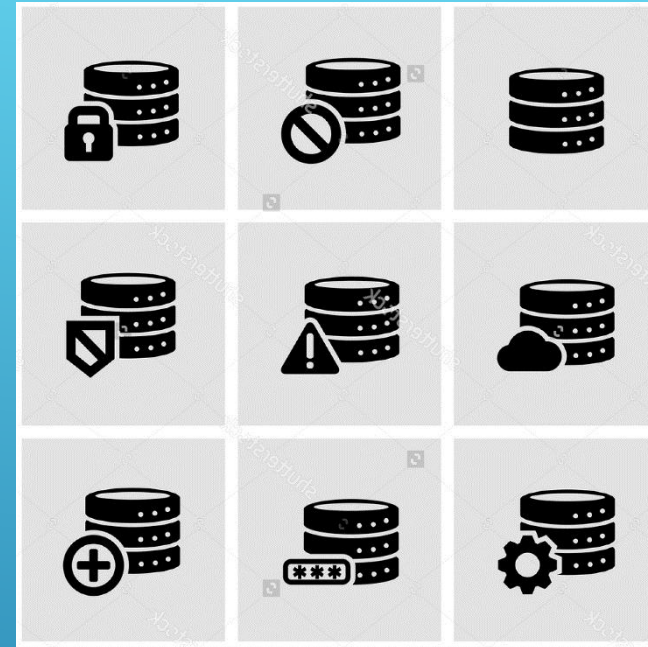
——王坚(阿里巴巴CTO)

事情/问题优先！

(决什么策？为何[谁]决策？)



Option 1



Big/Small Data



Option 2



No/Little Data




这些问题 无论多少数据 都无法告诉你正确的答案

you know, the problems where no amount of data  
or numbers will get you to the right answer



- ▶ 我发现放到总统桌上的问题总是难题——那些无论多少数据或数字都无法得出正确答案的难题……那些风险如此之高的选择，根本容不得一星半点的差错。
- ▶ 还有，作为总统，你会收到各种各样的人向你发出的各种各样的建议。
- ▶ 但是到最后，需要做出决定的时刻，作为总统，你所拥有的全部指引就是你的**价值观**，**判断力**，以及那些对你影响深远的**成长经历**。

我们(城市)碰到的大事情是？  
我们需要决关于什么的策？



No.	UNDP* (全球)	UNDP* (发展中国家)	Science**	中国***
1	失业	贫民(窟)	污染和废物处理	环境恶化
2	气候变化	非正规就业	可持续能源	住房
3	环境恶化	基础设施不足	健保和疾病预防	教育
4	-	无序发展	社区建设	气候变化
5	-	资源分配冲突	犯罪与安全	失业
6	-	-	食品安全	交通
7	-	-	创新	贫富差距

\*UNDP, 2013.

\*\*Science, 2016.

\*\*\*来源于笔者2016年5-6月开展的网上调研，N=233。

## 城市的大事情

	问题	.....	研究火	研究特点
1	环境恶化	.....	个人行为及其时空差异	What not how
2	住房	.....	城市与区域的形态和结构	What not how
3	教育	.....	规划制定和管理	What; How
4	气候变化	.....	空间模拟和公共参与	What; How
5	失业	.....	.....	.....

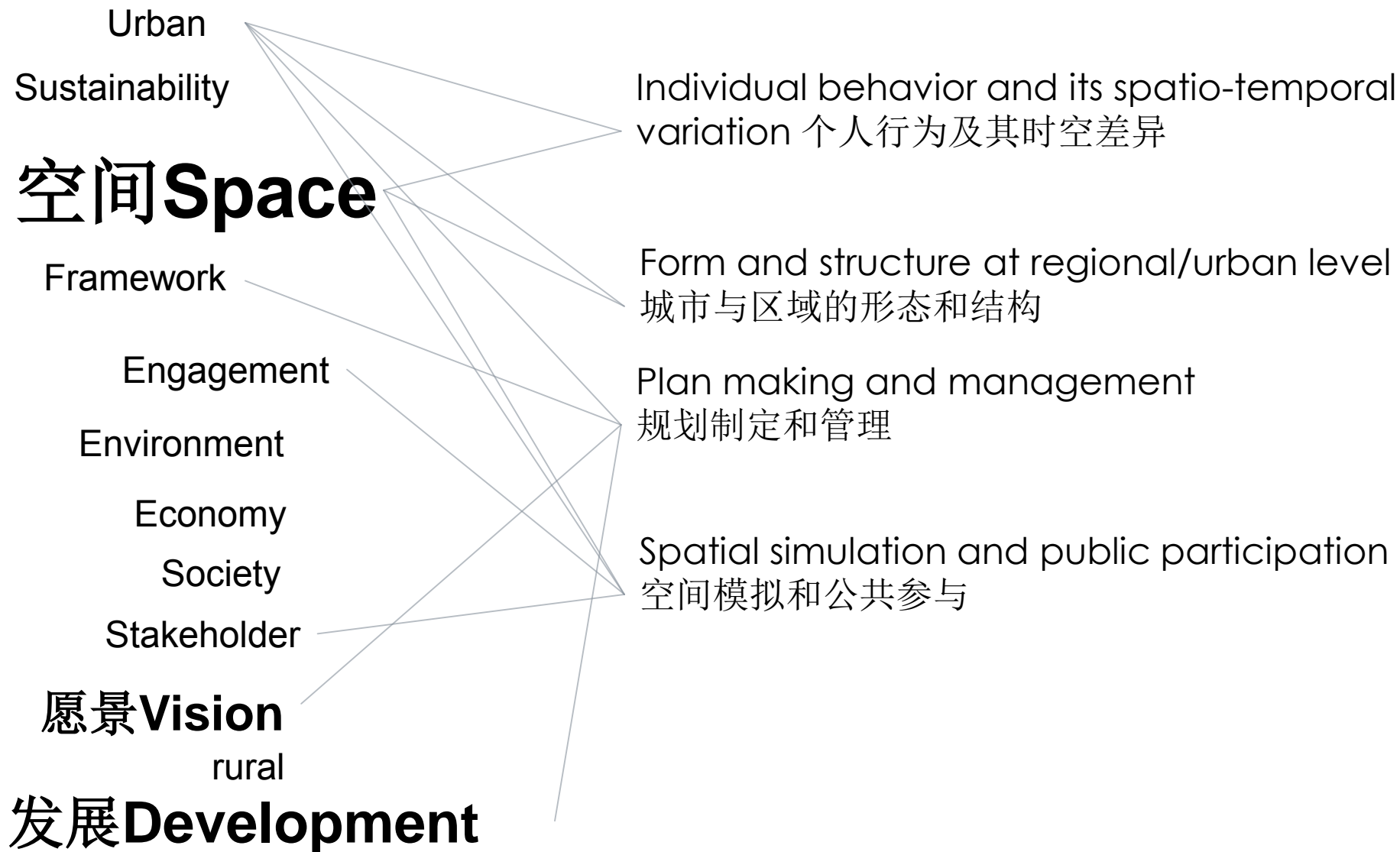
\* Hao et al., 2015

# 大数据研究VS.大事情

- ▶ **Urban Planning** is not about images but is a way to make a difference; it is a framework that helps leaders transform a vision into reality using space as a key resource for development and engaging stakeholders along the way (UNDP, 2013)
- ▶ 协调城乡空间布局，改善人居环境，促进城乡经济社会全面协调可持续发展(《城乡规划法》)

## 大数据研究VS.大事情(2)

# 大数据研究VS.大事情(3)



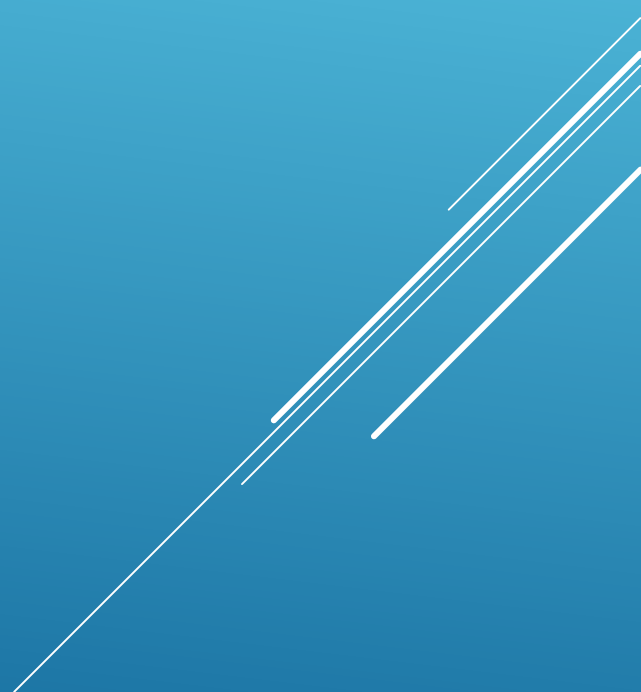
- ▶ 个人行为及其时空差异
- ▶ 城市与区域的形态和结构
- ▶ 规划制定和管理
- ▶ 空间模拟和公共参与

对照王坚所说，我们主要还在强化业务，还是没有做“以前做不了的事情”？！

# 大事情？！



基于以上，更多的大思考？





有待强化的方向

## STRATEGIC/PRINCIPAL CONSIDERATIONS

- ▶ What!
  - ▶ How!
  - ▶ Value! Vision! Valor! Validity! (c.f.: Volume, velocity, variety, veracity)
- 
- A decorative graphic consisting of several parallel white lines of varying lengths, slanted upwards from left to right, located in the bottom right corner of the slide.

II



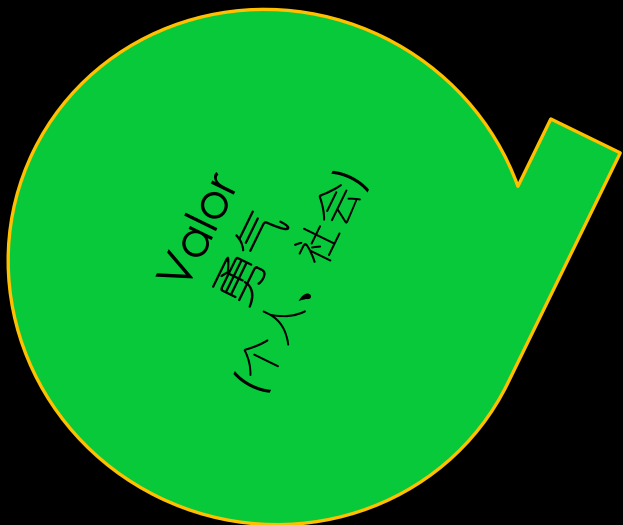
I



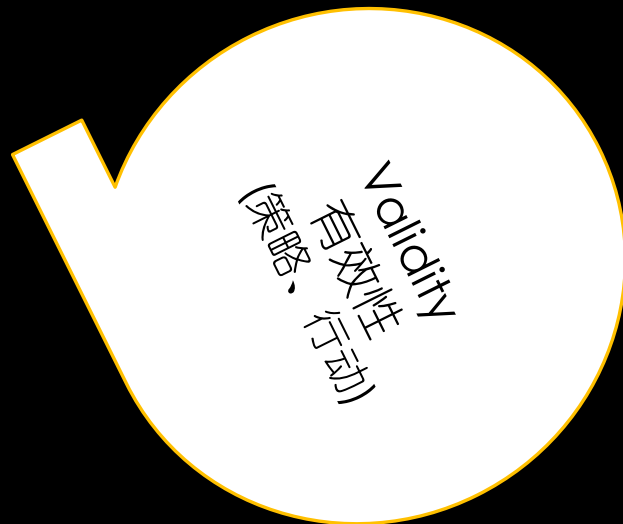
(城乡规划)

决策导向型的大数据应用

III



IV



## 40 ZETTABYTES

[ 43 TRILLION GIGABYTES ]  
of data will be created by 2020, an increase of 300 times from 2005

6 BILLION PEOPLE  
have cell phones



WORLD POPULATION: 7 BILLION

## Volume SCALE OF DATA

2005

2020

It's estimated that  
**2.5 QUINTILLION BYTES**

[ 2.3 TRILLION GIGABYTES ]  
of data are created each day



Most companies in the U.S. have at least  
**100 TERABYTES**  
[ 100,000 GIGABYTES ]  
of data stored



# The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015  
**4.4 MILLION IT JOBS**  
will be created globally to support big data, with 1.9 million in the United States



As of 2011, the global size of data in healthcare was estimated to be

**150 EXABYTES**  
[ 161 BILLION GIGABYTES ]



**30 BILLION PIECES OF CONTENT**  
are shared on Facebook every month



## Variety DIFFERENT FORMS OF DATA

By 2014, it's anticipated there will be

**420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

**4 BILLION+ HOURS OF VIDEO**  
are watched on YouTube each month



**400 MILLION TWEETS**  
are sent per day by about 200 million monthly active users



The New York Stock Exchange captures

**1 TB OF TRADE INFORMATION**  
during each trading session



## Velocity ANALYSIS OF STREAMING DATA

Modern cars have close to  
**100 SENSORS**  
that monitor items such as fuel level and tire pressure



By 2016, it is projected there will be

**18.9 BILLION NETWORK CONNECTIONS**

— almost 2.5 connections per person on earth



**1 IN 3 BUSINESS LEADERS**

don't trust the information they use to make decisions



Poor data quality costs the US economy around

**\$3.1 TRILLION A YEAR**



**27% OF RESPONDENTS**

## Veracity UNCERTAINTY OF DATA

in one survey were unsure of how much of their data was inaccurate

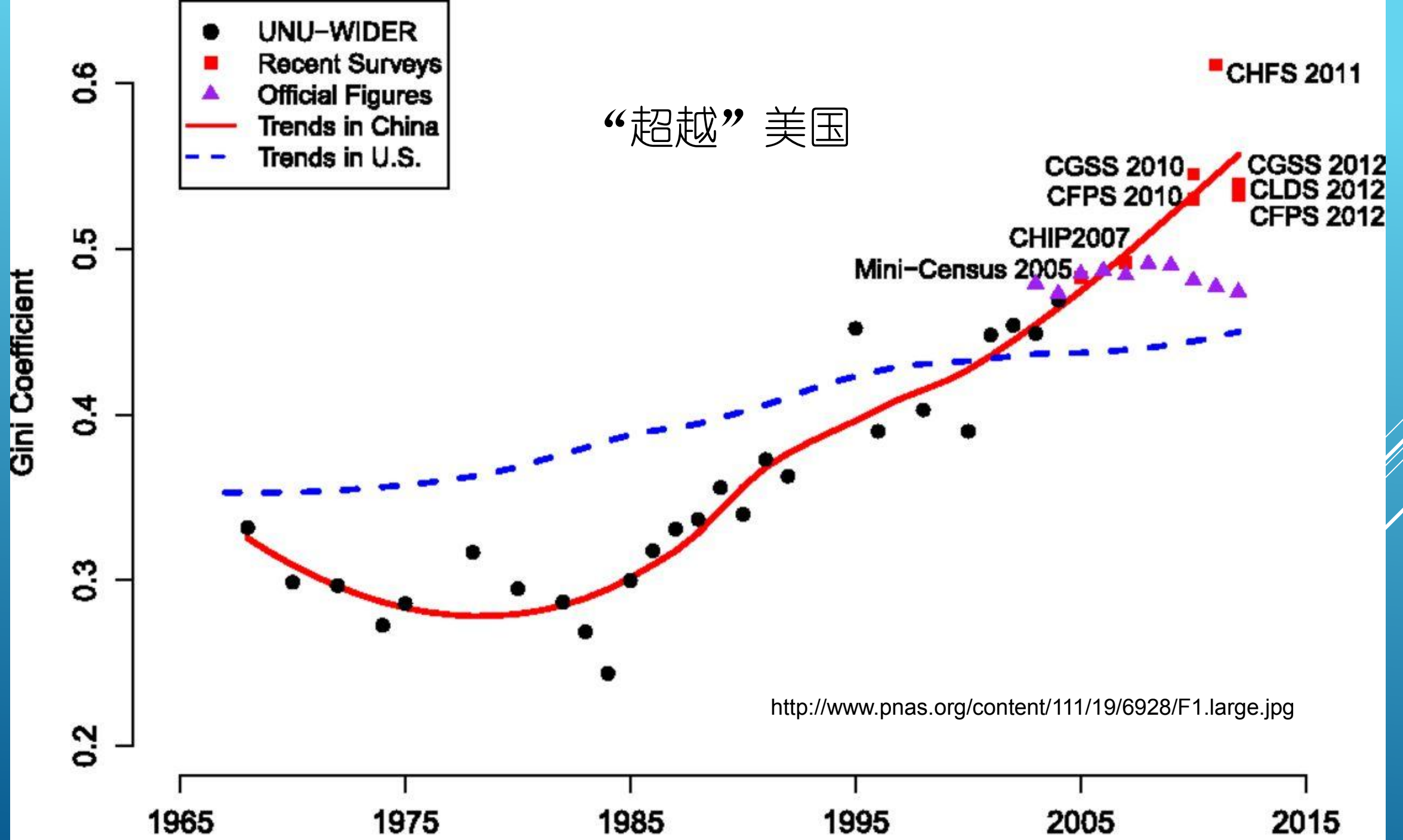
# 有待强化的子域

## TACTICAL/DOMAIN CONSIDERATIONS

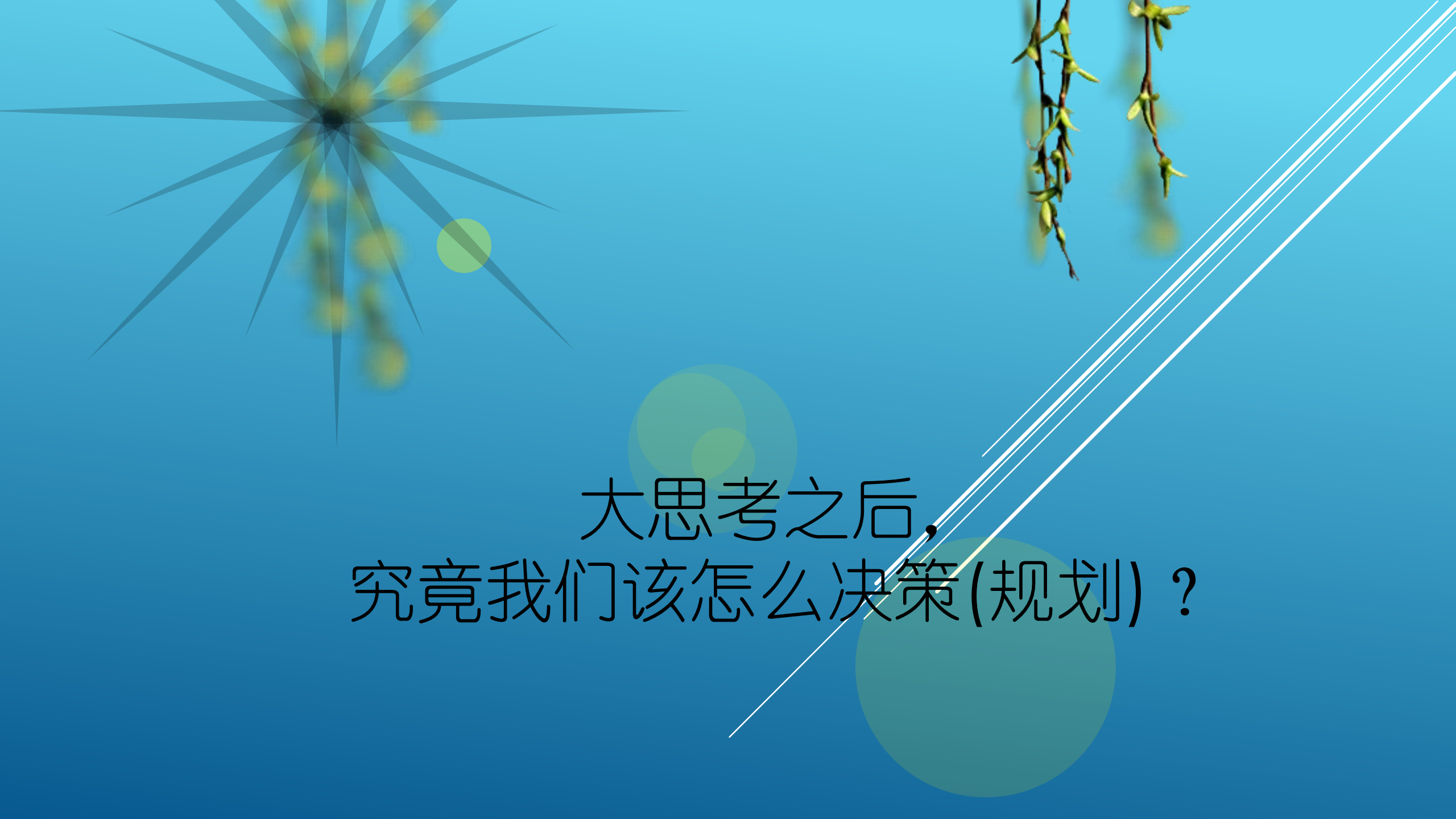
- ▶ 环境 (ep.空气, 水, 土壤, 废物处理)
- ▶ 社会发展(ep. 贫富差距)
- ▶ 农村
- ▶ 可持续性 (e.g.,资源的循环使用; 代际平衡)



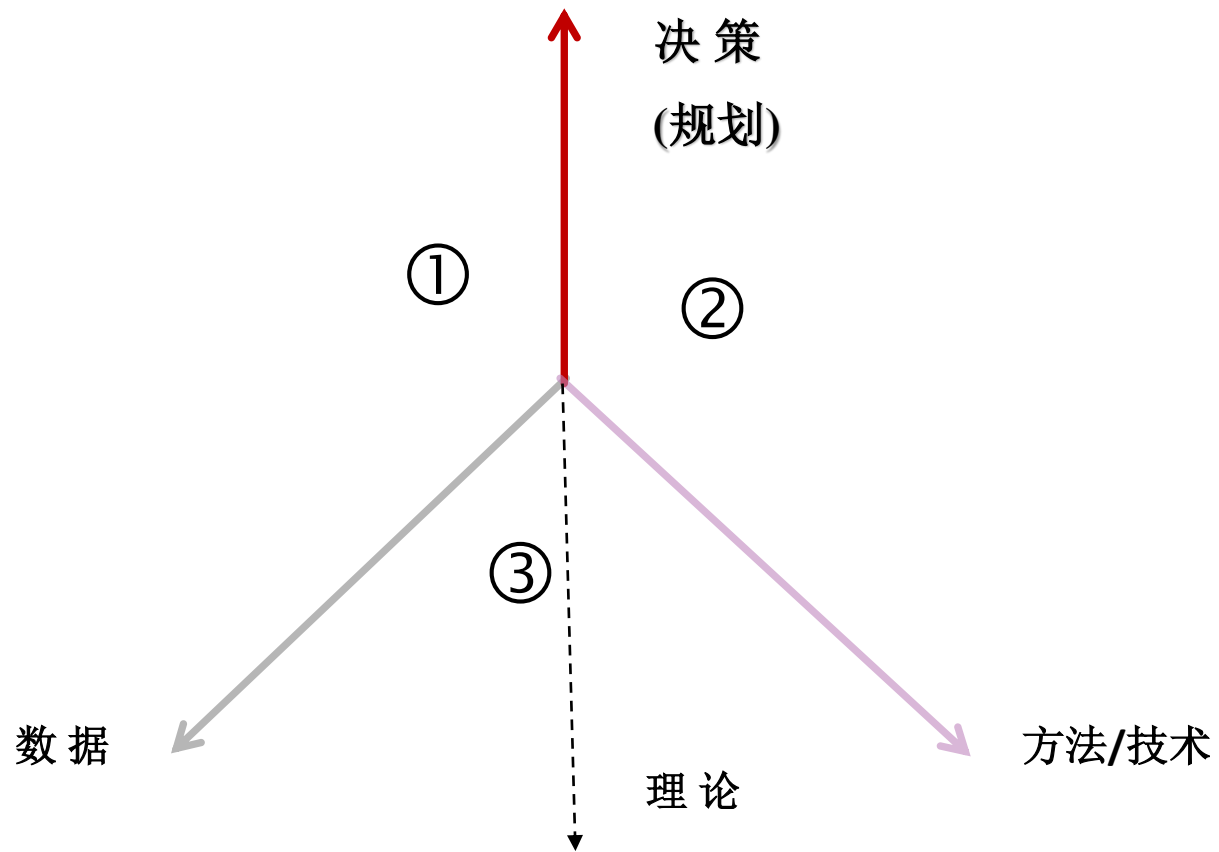
“超越” 美国



<http://www.pnas.org/content/111/19/6928/F1.large.jpg>



大思考之后，  
究竟我们该怎么决策(规划)？



the  
scholar

---

- ▶ 城市逐渐代替国家成为世界舞台上的主要角色。
- ▶ 城市就和国家一样，必须要考虑如何平衡利益与摩擦。



战略家的视角



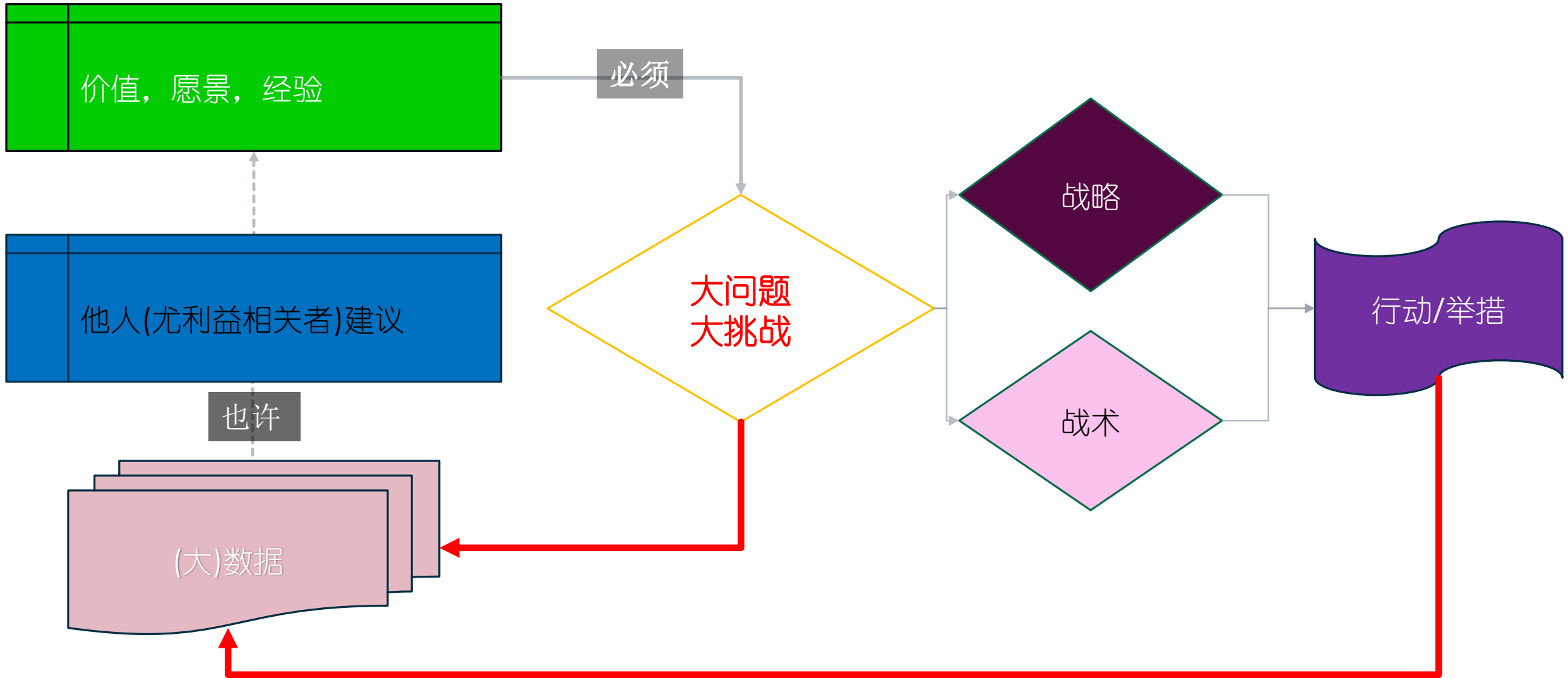
- ▶ 明确主要矛盾(角色) 在哪里？
- ▶ 利益是什么？
- ▶ 摩擦在哪里？
- ▶ 涉及谁？
- ▶ 他们的价值和愿景(诉求)是什么？

针对以上，需要数据么？如果需要，是什么(大)数据，它们可以怎么采集、分析？需要最好地采集和分析，我们需要什么样的(大)协作？

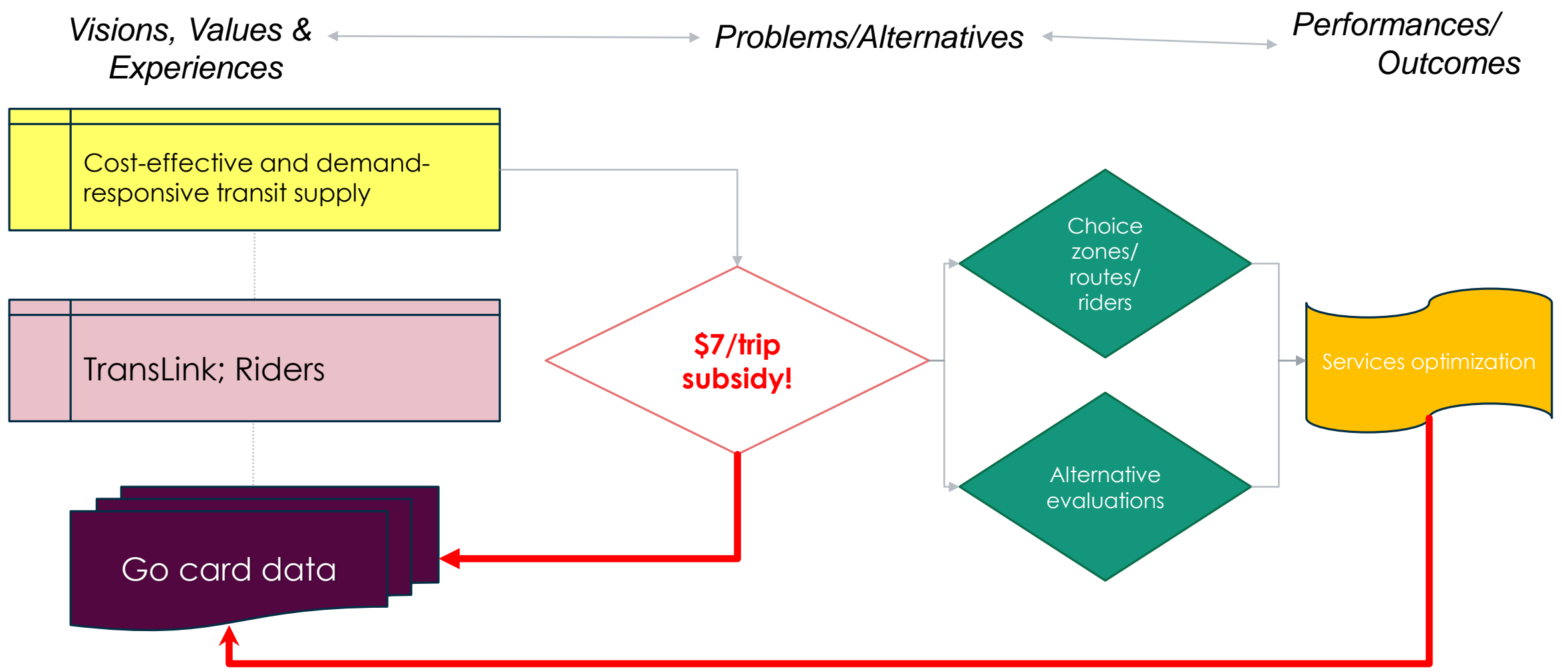
## 怎么决策(规划)？

(城乡规划)

# 决策导向型的大数据应用范式



# “挑剔顾客” 识别和服务优化





# TransLink Tracker

October 2015–December 2015 Q2

大事情/大问题

- ▶ 合理性价比 (TransLink; The Public)
- ▶ 需求及时响应 (Riders)
- ▶ 经验(短期数据, 零星感受——Planners; Operators)

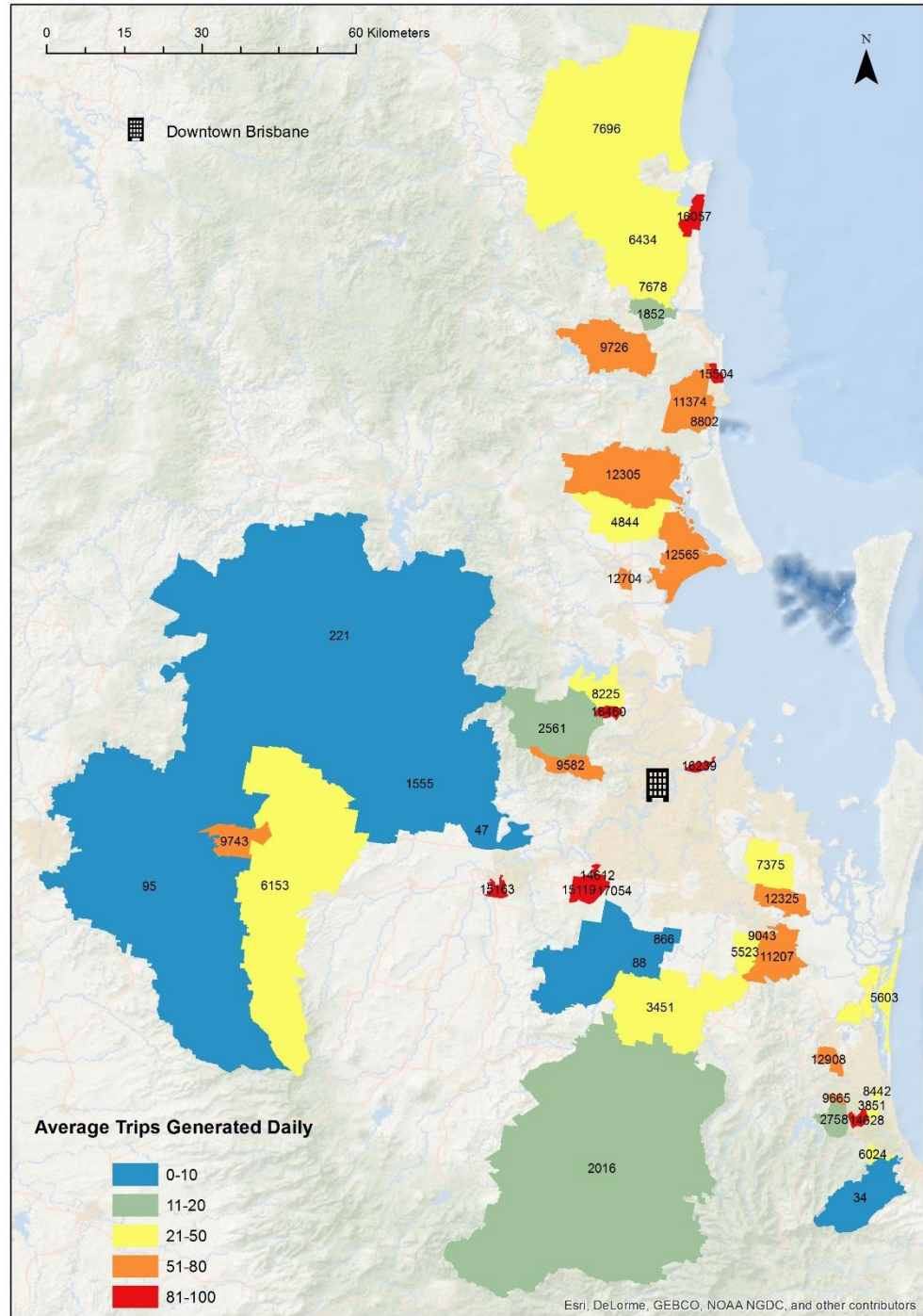
价值, 目标和经验

- ▶ 繁忙线路调整难度很大
- ▶ 郊区接驳铁路的公交有潜力
- ▶ 以上公交乘客个体需求主要还是换乘，各种方式的Rider-sharing(e.g.,小公交，合乘和Uber)均有可能满足其需求

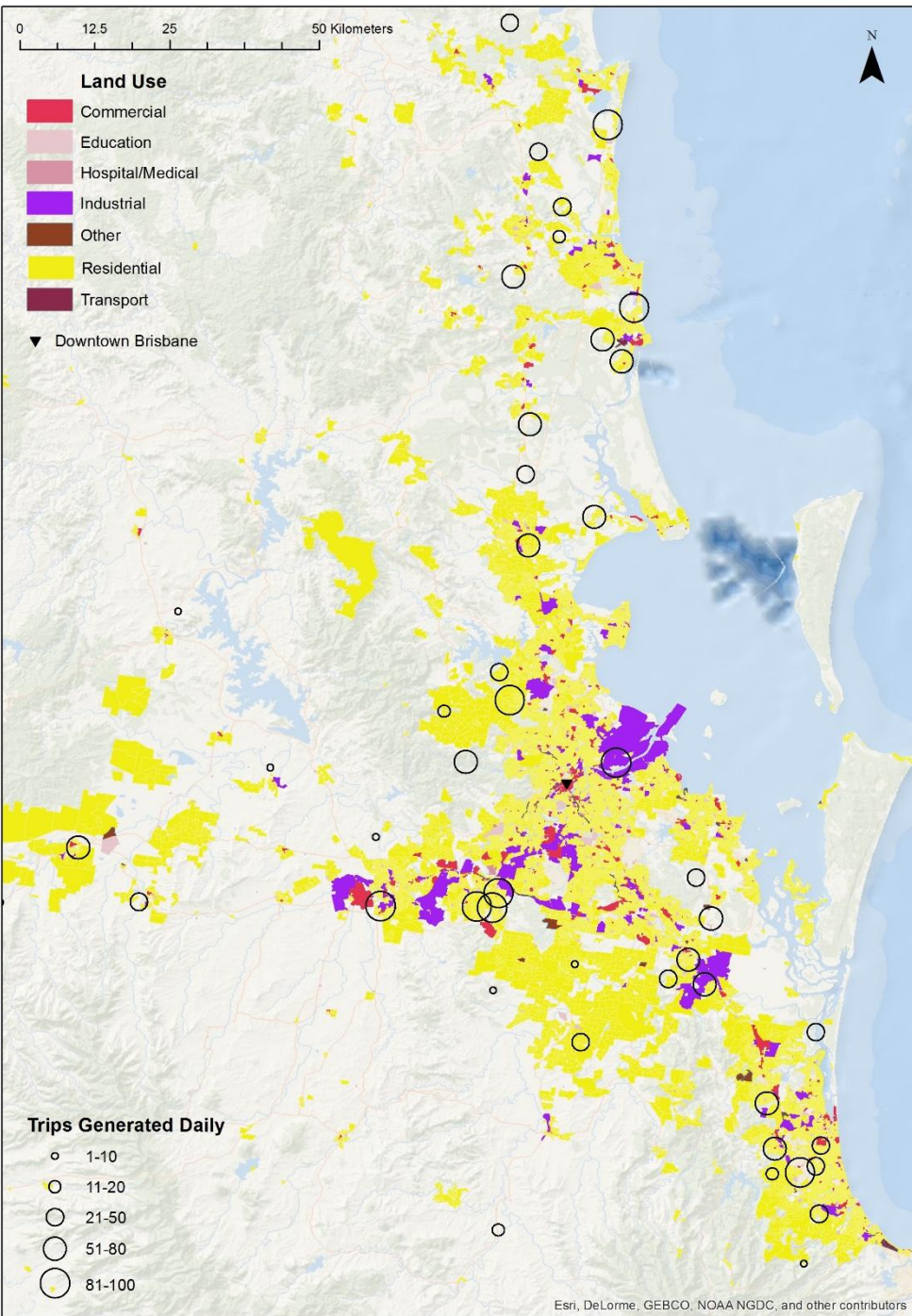
他人建议、想法  
(VALIDITY CHECK)

- ▶ Go card data: 分不同时段的需求和使用的空间分布
- ▶ Go card data+补充数据: 潜在的, 可优化的客户群, 线路和片区; 不同服务供给方案的评估
- ▶ 未来(大)数据改善建议: 以优化服务为目标, 调整数据采集周期、内容, 格式、分析、分享等!

## (大)数据







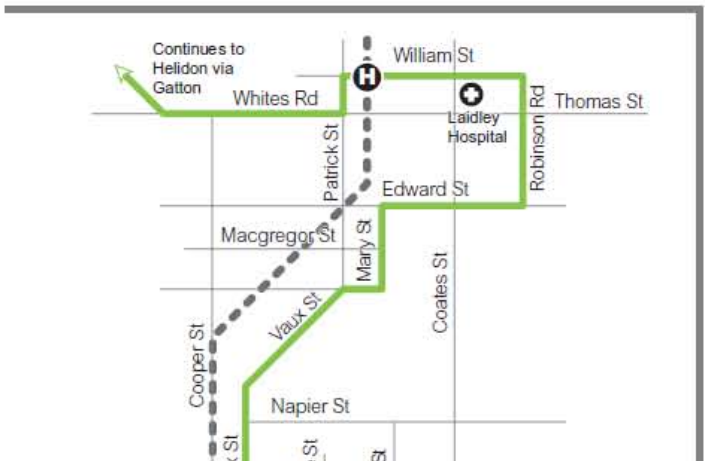
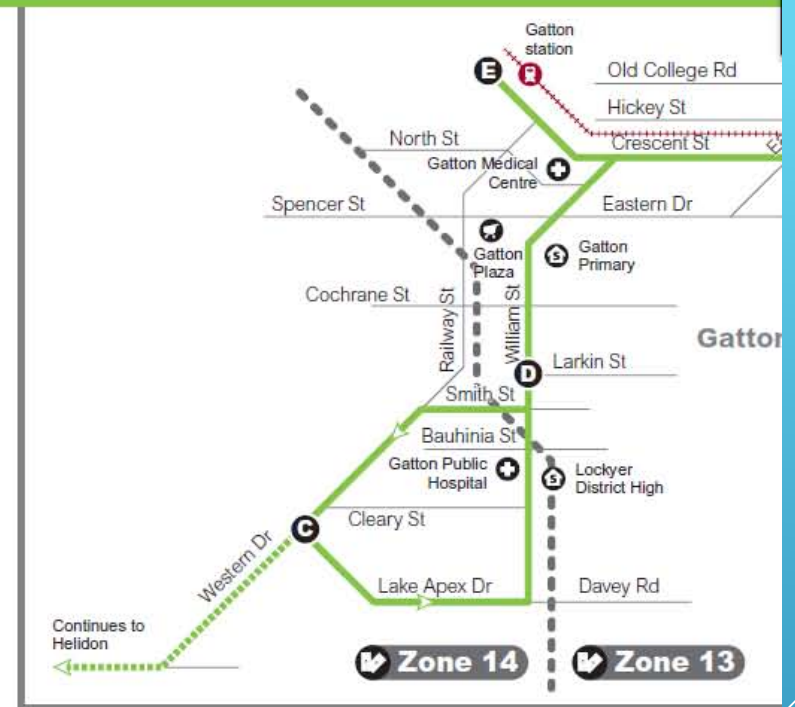
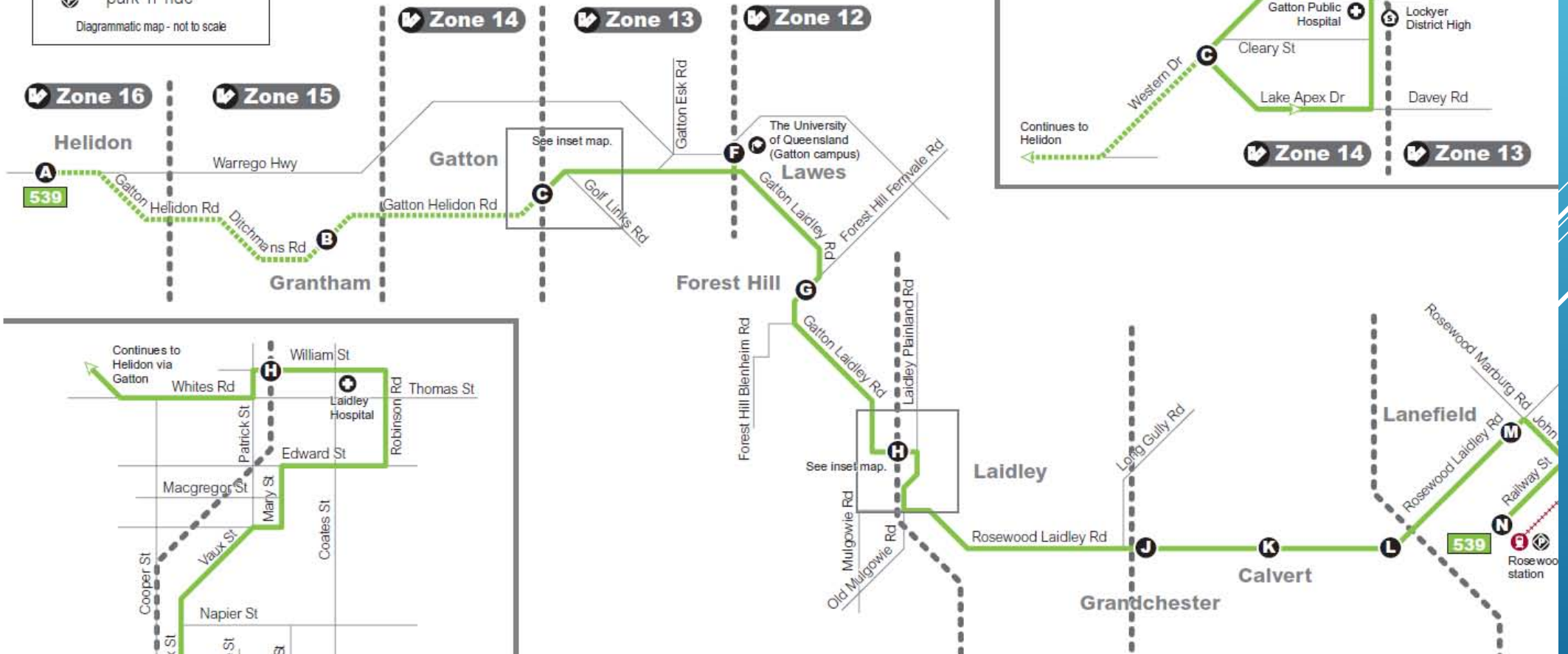
**Key**

- route 539
- route 539 (occasional service)
- zone boundary
- timing points
- train line & station
- school
- university/TAFE
- shops
- hospital
- park 'n' ride

Diagrammatic map - not to scale

**Route 539**

- Servicing**
- Helidon
  - Grantham
  - Gatton
  - Lawes
  - Forest Hill
  - Laidley
  - Grandchester
  - Calvert
  - Lanefield
  - Rosewood



Variable	Alternative 1 (Bus)	Variable	Alternative 2 (Uber)
Annual trips*	17768	Uber cost per km	\$1.56
Subsidy/trip	\$7	Average trip distance*	10.9(km)
Rider's cost/trip	\$3	Uber cost per trip	\$17.0
Annual Cost	\$177,680	Uber cost per trip*person (Assume two riders/vehicle)	\$8.5

\*Based on Go card data.

从单一公交供应到多元运输服务：

- ▶ 短期: Proactive, demand responsive and cost-effective ride service supply/sharing rather than passive bus-service provision!
- ▶ Ride services: Uber, carpool, vanpool & car-sharing
- ▶ 中长期: Accessibility/Life planning and management

服务优化(行动)



**TDOT 25-YEAR  
LONG-RANGE  
TRANSPORTATION  
POLICY PLAN**



**ACCESSIBILITY  
POLICY PAPER**



- ▶ 无论有无(大)数据，我们其实都要决策(e.g., Ms. Obama)
- ▶ 识别好，把握住大问题、大事情比大数据更重要! (e.g., 王坚)
- ▶ 大抱负(价值理念)、大经历和大判断是找好大问题、大事情的前提 (e.g., Parag Khanna)
- ▶ 以上再加上能(合)做(作)小事情，就可以有大作为

## 小结

# 小结(2)

