

CSCE 580: Introduction to AI *CSCE 581: Trusted AI*

Lecture 27: AI for the Real World

PROF. BIPLAV SRIVASTAVA, AI INSTITUTE

30TH NOV, 2023

Carolinian Creed: “I will practice personal and academic integrity.”

Credits: Copyrights of all material reused acknowledged

Organization of Lecture 27

- Introduction Segment
 - Recap of Lectures 26
- Main Segment
 - Bringing All Together
 - Real world problems
 - Smart city – setting goals for improvement
 - Framework for identifying opportunities to solve problems with AI
 - Case studies in smart city (public health) and business (Clarity - business intelligence)
- Concluding Segment
 - Course Project Discussion
 - About Next Lecture – Lectures 28 and 29
 - Ask me anything

Introduction Section

Recap of Lecture 26

- Topic discussed
 - Graduate paper presentations
 - Feedback by undergrads and 2 advanced reviewers
- Feedback shared with presenters

Final Exam – Based on Graduate Student Presentations

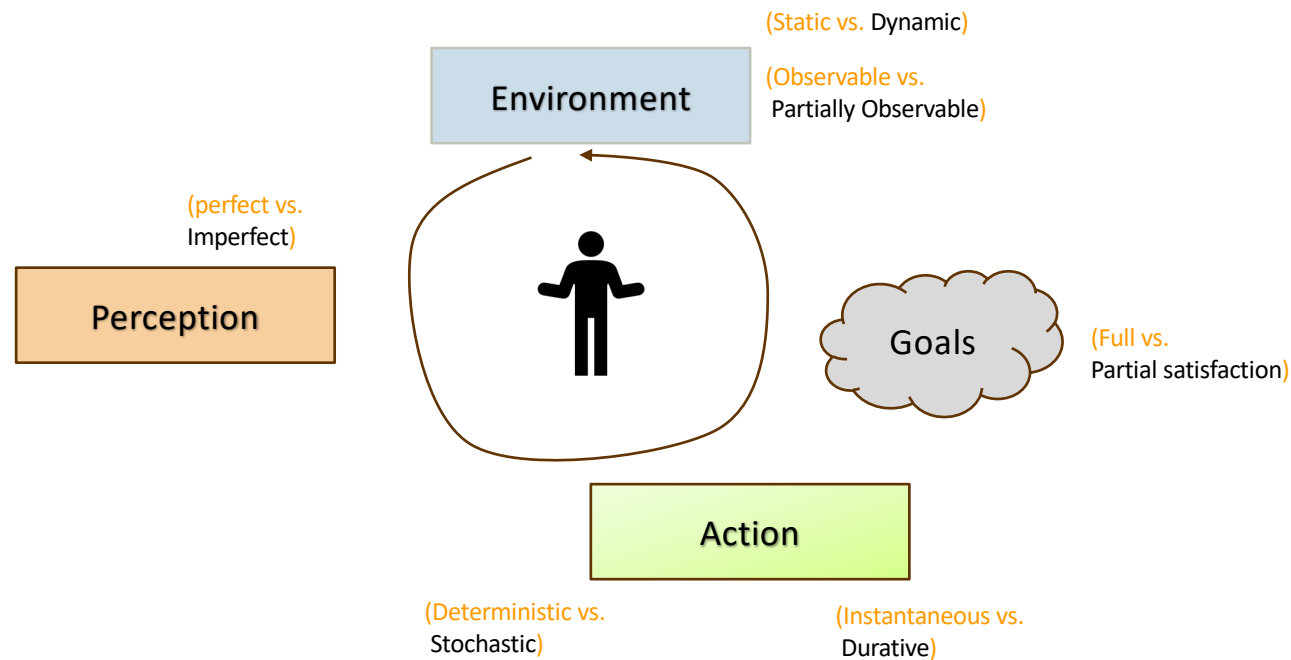
- Graduate students
 - Paper presentation [100 points]
 - Write about their paper presented [100 points]
- Undergraduate students
 - Write about 2 papers presented in class by graduate students [150 points]
 - Vote for the papers presented [50 points]
- Paper reports due by Dec 5, 2023 (Tuesday)

Final Exam	200	100
Additional Final Exam – Paper summary, in-class presentation		100

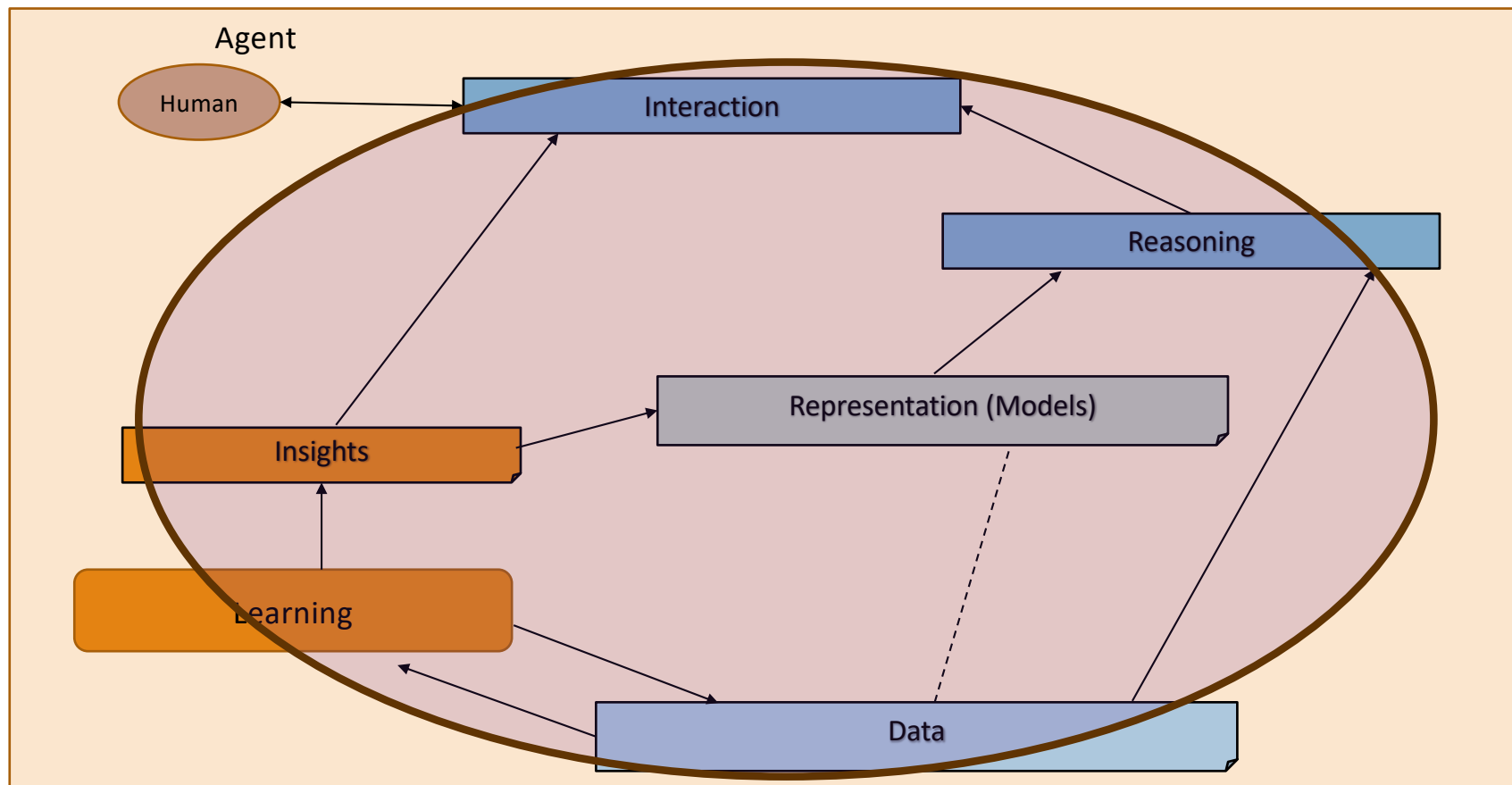
Final Exam – Based on Graduate Student Presentations

- Evaluation
 - By undergrads as well as 2 additional (instructor and TA)
 - All undergraduates to attend and give survey response; link to be shared
 - Those undergrads not giving inputs will be given negative marks as part of the final score [-10 point per presentation]
- What to have in the report – minimum 1 page per paper (<500 words).
 - Paper summary
 - Key contributions
 - Your critique about the paper.

Intelligent Agent Model



Relationship Between Main AI Topics



Where We Are in the Course

CSCE 580/ 581 – In This Course

- Week 1: Introduction, Aim: Chatbot / Intelligence Agent
- Weeks 2-3: Data: Formats, Representation and the Trust Problem
- Week 4-5: Search, Heuristics - Decision Making
- Week 6: Constraints, Optimization – Decision Making
- Week 7: Classical Machine Learning – Decision Making, Explanation
- Week 8: Machine Learning - Classification
- Week 9: Machine Learning - Classification – Trust Issues and Mitigation Methods
- Topic 10: Learning neural network, deep learning, Adversarial attacks
- Week 11: Large Language Models – Representation, Issues
- Topic 12: Markov Decision Processes, Hidden Markov models - Decision making
- Topic 13: Planning, Reinforcement Learning – Sequential decision making
- Week 14: AI for Real World: Tools, Emerging Standards and Laws; Safe AI/ Chatbots

Main Section

Credit: Retrieved from internet

Real World Problems

- Domains
 - Health
 - Water
 - Traffic
 - Food
 - Energy
 - ...
- What can AI do here?



Major Community Challenges

Challenges

Growing population

- Access to safe water
- Need jobs

Reducing resources

- Land
- Potable water

Environmental degradation

- Air pollution
- Animal management

Photo by: Jennifer Wolf/ WolfHart Image/ Marine Photobank - child's toy left on the beach.

Credit: <https://scdhec.gov/environment/your-water-coast/ocean-coastal-resource-management-ocrm/marine-debris-abandoned-vessels>



Credit:

<https://www.greenvilleonline.com/story/news/2022/12/29/fatalities-on-sc-roads-decline-in-2022-greenville-roads-dangerous-spartanburg-anderson-traffic/69763182007/>



South Carolina's traffic deaths rose by 26 percent from 1994 to 2020, **marking the biggest increase by any state over those 26 years.** In 2022, **1,035 people died on SC roads**, per preliminary data reported by the SC Department of Public Safety

AI in the Real World

- AI is eventually like any technology for problem solving
- Key question: Why will someone use AI? What value does it provide?

AI builds on trends

Lots of data
Cheap compute
Access to online information

Basics: Smart City

Goals: Smart City? Millennium Development Goals?

Smart city can mean one or more of the following:

As a **resource optimization objective**, it is to know and manage a city's resources using data.

As a **caring objective**, it is about improving standard of life of citizens with health, safety, etc indices and programs.

As a **vitality objective**, it is about generating employment and doing sustainable growth.

A city leadership can choose among these or define their own objective(s) and manage with measurements to pro-actively achieve it



See other FAQs at: <https://sites.google.com/site/biplavsrivastava/research-1/intelligent-systems/scfaqs>

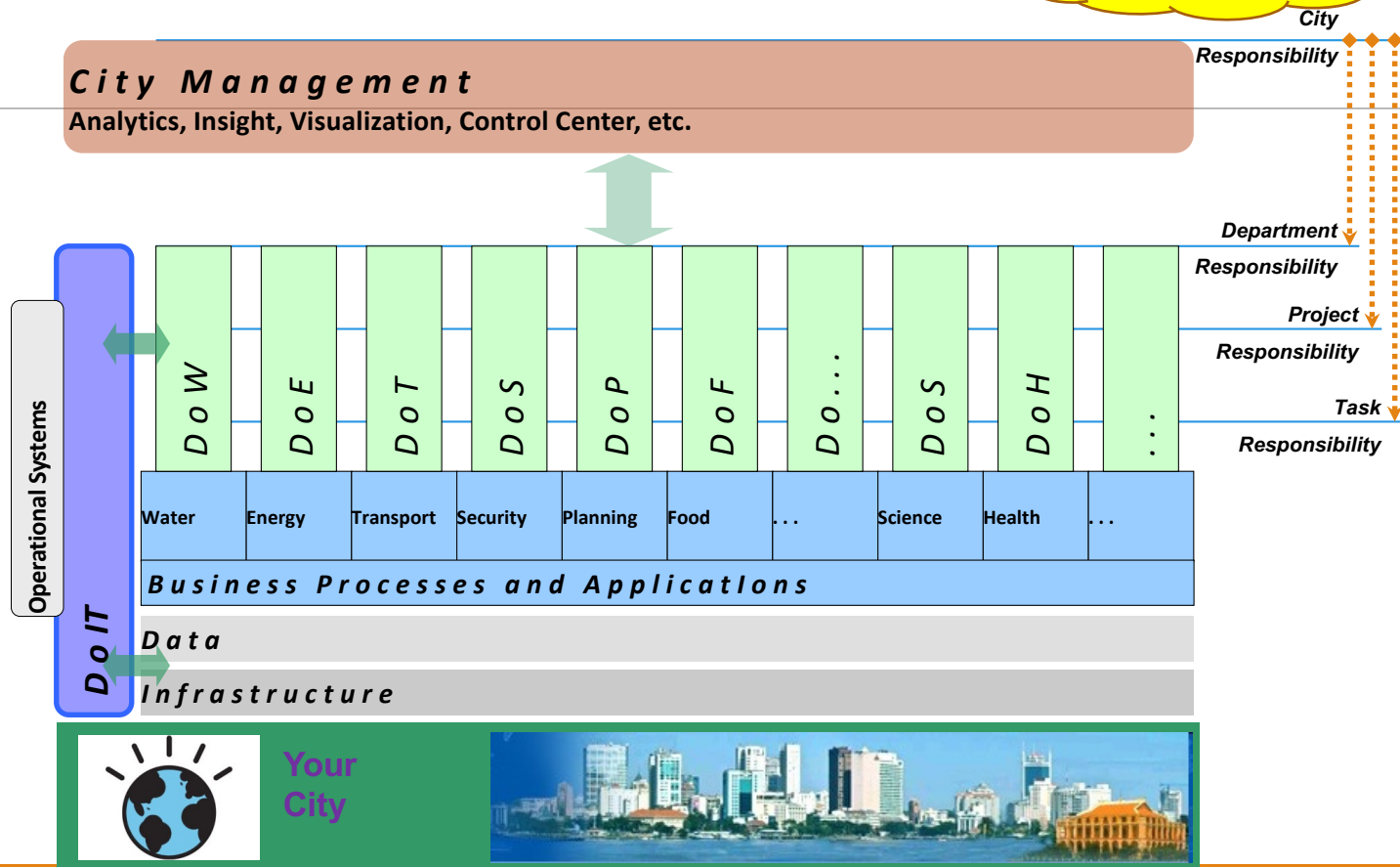
Credits: See: <https://www.un.org/millenniumgoals/bkgd.shtml>

Before

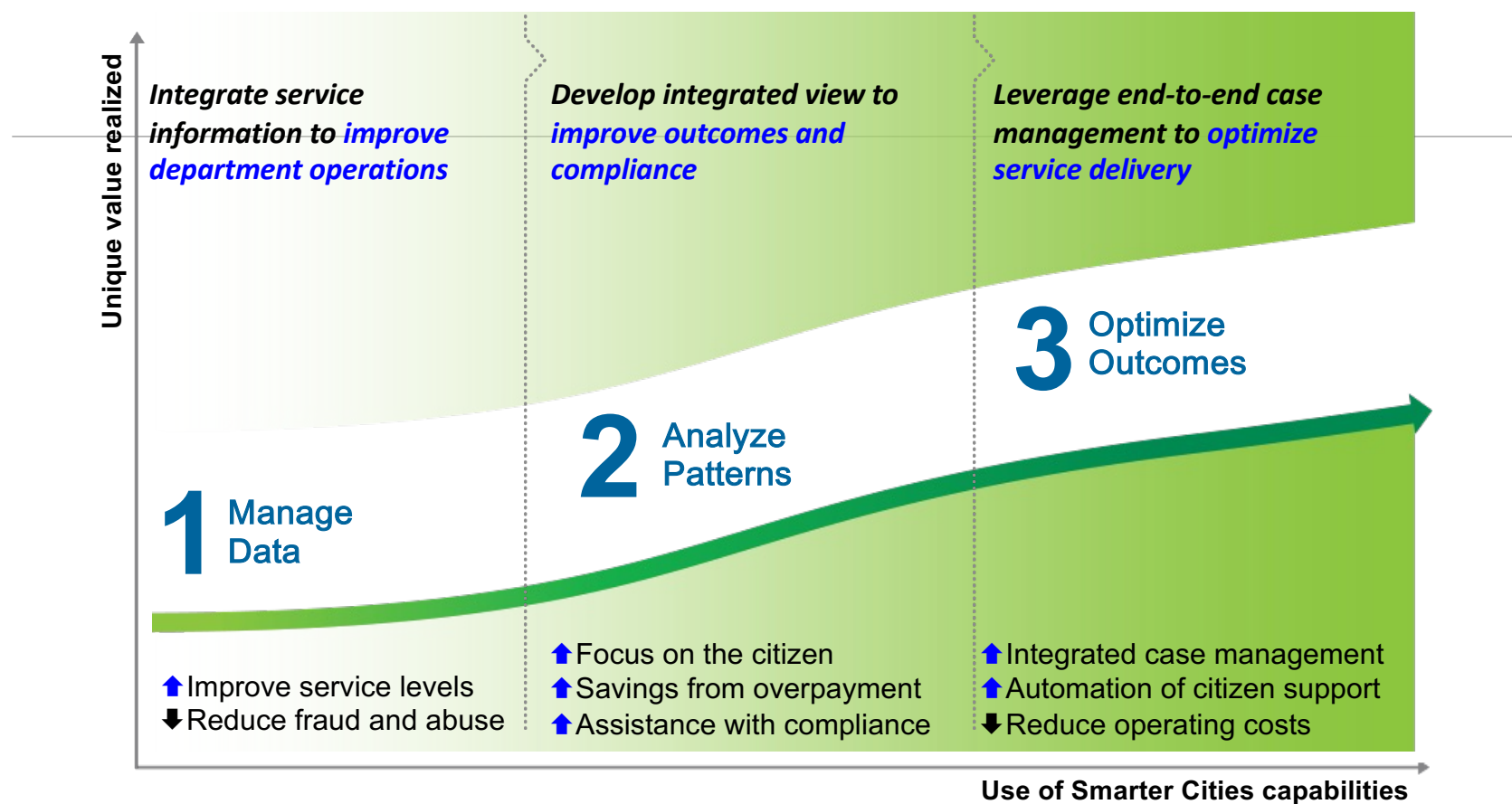


An integrated Smarter City Framework – a comprehensive management system across all core systems, will anchor the vision to executable steps

After



Smarter Cities solution paths leverage a similar approach



Major Community Challenges and Opportunities Today

Challenges

Growing population

- Access to safe water
- Need jobs

Reducing resources

- Land
- Potable water

Environmental degradation

- Air pollution
- Animal management

AI Opportunities - many avenues for impact

- Water management – e.g., clean plastics [1]
- Disease – detecting from photos [2]
- Video-based animal management
- Intelligent traffic control
- Promoting sustainable tourism
- Journalism - Media coverage of homicides and role of race [2]
- **Vision:** create participative and trustable technology solving real-world problems

References:

1. <https://theconversation.com/new-pfas-guidelines-a-water-quality-scientist-explains-technology-and-investment-needed-to-get-forever-chemicals-out-of-us-drinking-water-201855>
2. <https://www.prnewswire.com/news-releases/students-win-more-than-1-8-million-at-2023-regeneron-science-talent-search-for-remarkable-scientific-research-on-rna-molecule-structure-media-bias-and-diagnostics-for-pediatric-heart-disease-301772440.html>

2015 IJCAI Tutorial - **AI for Smart City Innovations with Open Data**
<https://www.linkedin.com/pulse/tutorial-ai-smart-city-innovations-open-data-biplav-srivastava/>

Open Data

- 2015 International Open Data Charter – established six norms for the publication of public data stating that data should be:
 - open by default
 - timely and comprehensive
 - accessible and usable
 - comparable and interoperable
 - for improved governance and citizen engagement
 - for inclusive development and innovation.
- Culmination of open data movements around the world, as discussed in first few lectures

Reference: https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf

A Framework

A Systematic Approach

- **Identify:** Value of decision: before and after
- **Assess:** Data-needed
- **Explore:** Methods
- **Conduct:** Evaluation
- **Integrate:** solution with overall process



IJCAI 2015 Tutorial: <https://sites.google.com/site/aismartcitytutorial/>

Real-World Applications of Technology Follow a Pattern

- **Value (from Action, Decisions)** – Providing benefits that matter to people most in need of, and in a timely and cost-efficient manner. Going beyond technology to process and people aspects.
- **Data + Insights** – Available, Consumable with Semantics, Analysis, often including Visualization too
- **Access** - Apps (Applications), Usability - Human Computer Interface, Application Programming Interfaces (APIs)

Example – River Water Pollution

- Value – To individuals, businesses, government institutions
 - Example – Can I take a bath? Will it cause me dysentery?
 - Example – How should govt spend money on sewage treatment for maximum disease reduction?
- Data – Quantitative as well as qualitative
 - Dissolved oxygen,
 - pH,
 - ... 30+ measurable quantities of interest
- Access –
 - Today, little, and that too in water technical jargon
 - In pdf documents, website
- **Key Idea:** Can we make insights available when needed and help people make better decisions?

Example –Traffic Management

- Value – To individuals, businesses, government institutions
 - Example – Can I reach office on time? Where to park if I take my car?
 - Example – How much overt-time does the city need to give today? Where should I deploy my traffic cops today?
 - Example – When to service city's buses?
- Data – Quantitative as well as qualitative
 - Volume – traffic count
 - Speed on road
 - City events
- Access –
 - Today, little and on city websites
 - Facebook sites
- **Key Idea**: Can we make insights available when needed and help people make better decisions?

Journey Planning with Frugal Data

Promoting Public Transportation: Before and After We Seek

Many cities around the world, and especially in India and emerging ones, are getting their transportation infrastructure in shape.

- They have multiple, fragmented, transportation agencies in a region (e.g., city)
- They do not have instrumentation on their vehicles, like GPS, to know about their operations in real-time
- Schedule of public transportation is widely available in semi-structured form. They are also beginning to invest in new, novel, sensing technologies
- Cities give SMS-based alerts about events on the road.

Our approach seeks to accelerate time-to-value for such cities.

Kind of Information	Today Available to Bus User	With IRL-Transit+	Benefit
Bus Schedule (static)	Available online and pamphlets	Available from IT-enabled devices(low-cost phones, smart phones, web)	Increase accessibility
Bus Schedule Changes (dynamic)	No information	Infer from city updates	Increase information
Analytics (Bus Selection Decision Support)	No information	Will be available (Transit)	Increase information
Standardization of information	No support	Will be supported (SCRIBE, Transit)	Increase information's interoperability

Prior Work

- **San Francisco Bay Area, USA** has : <http://511.org>
 - Multi-agency public authorities consortium, has advanced instrumentation
 - It is the model to replicate
- **Google** has state-of-the-art from any non-public organization. It has separate services
 - Maps for driving guidance
 - Transit for public transport, more than 1 mode
 - Gaps:
 - Considers only time, not other factors like frequency, fare and waiting time
 - Does not integrate across their services for different mode categories
 - Does not publish their data
 - Acknowledgement: We use their GTFS format to consolidate schedule data
- **Many experimental systems** with capabilities less than Google,
 - Delhi: Disha on DIMTS website - <http://61.16.238.196/disha/index.php>
 - Mumbai Navigator: <http://www.cse.iitb.ac.in/navigator1/index.html>
 - Mumbai: Go4Mumbai (portal)- A <http://www.go4mumbai.com/>
- **Shortest route finding algorithms** from mapping companies

Public Transport Journey Problem

- Invariant Inputs:
 - The person
 - has a vehicle (e.g., car), and
 - can also walk short distances
 - The city has taxis, buses, metros, autos, rickshaws
 - Buses and metros have published routes, frequency and stops
 - Autos and rickshaws can be available at stands, or opportunistically, on the road
 - Taxis can be ordered over the phone
- Input:
 - A person wants to travel from place A to B
- Output
 - Suggest to the person which mode or combination of modes to select
- Observation: Using preferences over factors that matter to users to keep commuting convenient, while making best use of available public and para-transit commute methods

Background: Public Transportation Schedule Information

- Is widely available for public transportation agencies around the world
- Gives the basic, static, information about transportation service
- Usually in semi-structured format with varying semantics
- Can have errors, missing data

Input : Bus Stop

Bus id	256
Source	Indira Puri (Loni)
Destination	Shahdara (T)
Bus Stop	1. INDIRA PURI LONI BORDER 2. P.S. INDIRA PURI 3. JAWAHAR NAGAR 4. INDIRA PURI XING 5. SHAHDRA.

Input : Bus Time Table

Type of Bus	DEPOT : (Low Floor Buses),...
Running Time	"RUNNING TIME : 120 Minutes
Bus id	ROUTE NO.,DEPARTURE TIME,...
Bus Starting Time	34, Mehrauli Terminal, Noida Sector-32 ,05.10,05.15,

Input : Metro Stop

Train id	,LINE-1,...
Stop Name	STATION,TIME,
First Metro Time	Dilshad Garden,5:30:00, JHILMIL,5:23:40, Mansarovar Park,5:21:20,

Input : Metro Time Interval

Train id	Line-1,
Frequency Time	NP Frequency,7, minute
Time	P Frequency,4, minute 0600,NP 0800,P

Solution Steps

- Use the widely available schedule information from individual operators (agencies)
- Clean and consolidate it across agencies and modes to get a multi-modal view for the region
 - Optionally: Convert it into a standard form
 - Optionally: Enhance (fuse) it with any real-time updates about services for the region
- Perform **what-if analysis** on consolidated data
 - Path finding using Dijkstra's algorithm
 - Analyses can be pre-determined, analyses can also be user-created and defined
- Make analysis results available as a service
 - On any device
 - To any subscriber

Multi-Mode Commuting Recommender in Delhi And Bangalore

Timeframe:
2012-2014

Get public transit directions

City:

From:

To:

Start: End:

Mode:

No. of choices:

No. of hops:

13 solutions found. Please click on the links to

🔍 **Solution 1** [# Hops = 0; Modes used = metro]

Source: aiims
Destination : rajiv chowk
Route Id : LINE-2
Starting Time : 17:00
Reaching Time : 17:22
Frequency : 005
Mode : metro

⬅️ **Solution 2** [# Hops = 1; Modes used = bus , metro]

⬅️ **Solution 3** [# Hops = 1; Modes used = bus , metro]

Highlights

- Published data of multiple authorities used; repeatable process
- Multiple modes searched
- Preference over modes, time, hops and number of choices supported; more extensions, like fare possible
- Integration of results with map as future work; already done as part of other projects, viz. SCRIBE-STAT



IRL - Transit - Search service details page

Plan trip understanding the best means of transportat

Get public transit directions

City:

From:

To:

Start: End:

Mode:

No. of choices:

No. of hops:

2 solutions found. Please click on the links to explore

🔍 **Solution 1** [# Hops = 1; Modes used = bus , bus]

Source: bial Bgl
Destination : majestic (kbs) Bgl
Route Id : BIAS-1
Mode : bus

Source: majestic (kbs) Bgl
Destination : jn. of nice rd Bgl
Route Id : G-7
Mode : bus

⬅️ **Solution 2** [# Hops = 1; Modes used = bus , bus]

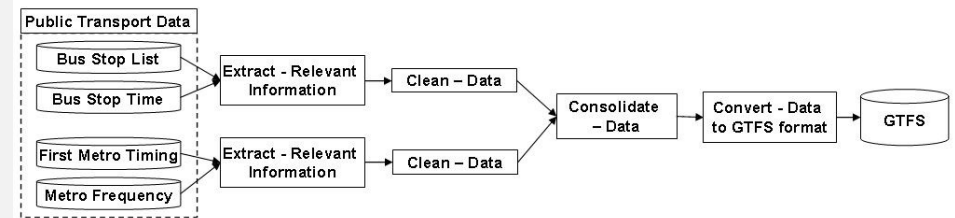
Technical Details – Scope, Accuracy, Future

- Factors which affect/ can improve accuracy

- Quality of schedule published by public transportation operators (bus and metro for us)
 - Names, spelling and conventions in stops by different agencies
 - We correct and can do more - If we correct too much, we remove the traceability to original published schedules
- Lack of co-relationship across stop names and location
 - Affects what the user sees when they select
 - We can include geo-spatial analysis when we offer choice of locations
- Increase inter-operability across agencies
 - Make traffic data into linked open data format
- Integrate with geo-spatial analysis of tools like STAT (IBM Research) or ESRI

- Future extension possibilities

- Technical
 - Include more cities; we currently have Delhi and Bangalore
 - Include more mode agencies
 - Include more data inputs as they become available, especially dynamic updates
 - Include fare, geo-spatial analysis and improved plans
- Business
 - Support a free or paid service to citizens in select cities by a client/ partner

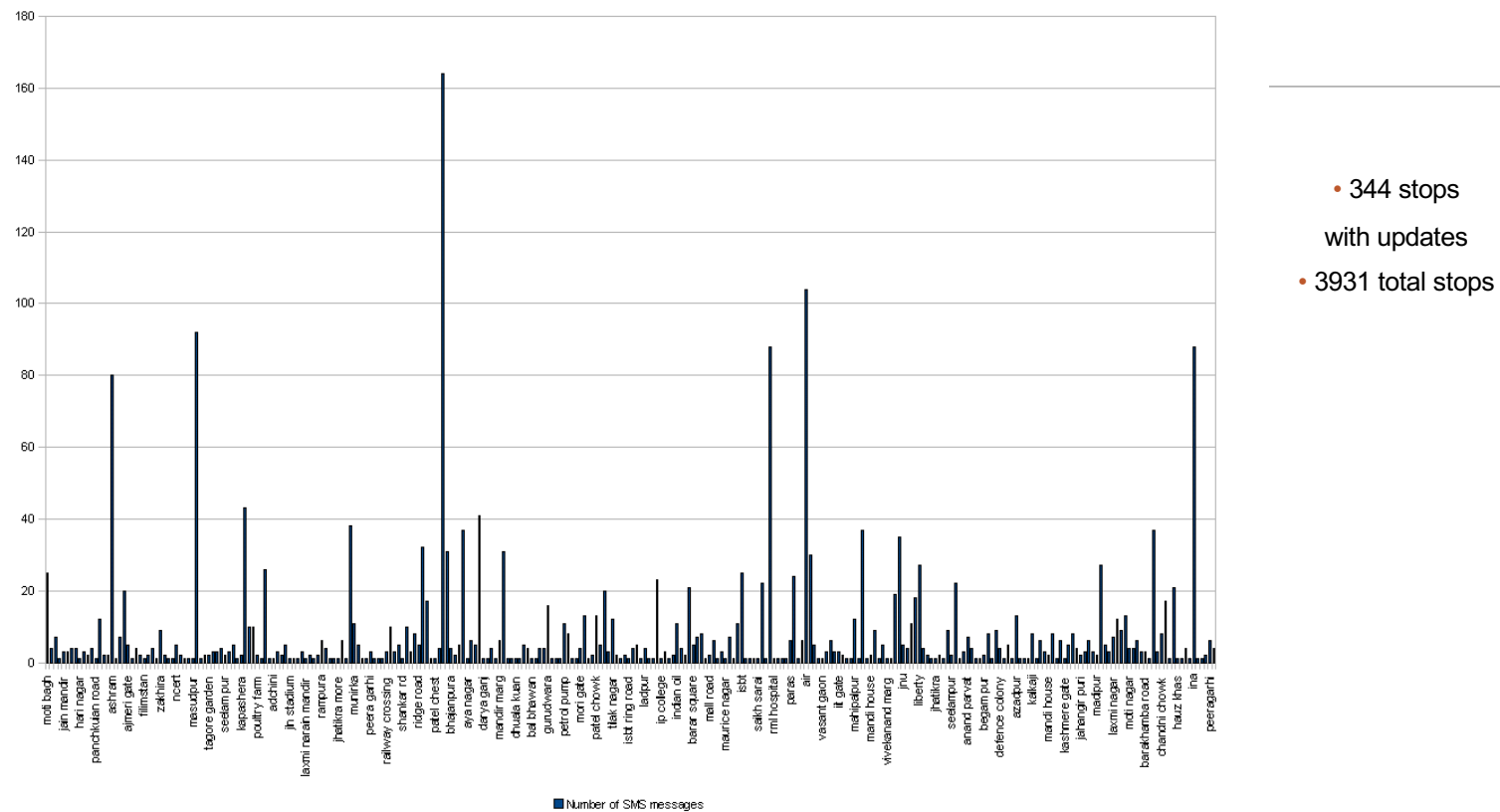


IRL-Transit is only one type of decision support
We built others for private para-transit
agencies (e.g., radio taxis)

Extended Work* - Dynamic Updates

- Invariant Inputs:
 - The person
 - has a vehicle (e.g., car), and
 - can also walk short distances
 - The city has taxis, buses, metros, autos, rickshaws
 - Buses and metros have published routes, frequency and stops
 - Autos and rickshaws can be available at stands, or opportunistically, on the road
 - Taxis can be ordered over the phone
- Input:
 - A person wants to travel from place A to B
 - [Optional] City provides updates on ongoing events, some may affect traffic
- Output
 - Suggest to the person which mode or combination of modes to select
- Observation: Using preferences over factors that matter to users to keep commuting convenient, while making best use of available public and para-transit commute methods
- * City Notifications as a Data Source for Traffic Management, Pramod Anantharam, Biplav Srivastava, in 20th ITS World Congress 2013, Tokyo

Number of SMS messages for bus stops in Delhi for 2 years (Aug 2010 – Aug 2012)*



* using Exact Matching

IRL-Transit

A Preference-aware Plan Selection for Mixed-Mode Commuting Powered With Just Schedule Data

IRL – Transit in Aug 2012

Get public transit directions

Examples

Delhi/AIIMS/Rajiv Chowk Bangalore/Majestic/Marathalli

City:

Search From:

Search To:

Start:

Mode:

City Update:

No. of choices:

No. of hops:

Key Points

- SMS message from city
- Event and location identified
- Impact assessed
- Impact used in search

100 solutions found for source **Aiims** and destination **Ina**
Please click on the links to explore

Group Hop 0

[# Hops = 0; Modes used = Bus]

[# Hops = 0; Modes used = Bus]

[# Hops = 0; Modes used = Bus]

Source: Aiims
Destination : Ina
Route Id : 465EXT
Mode : Bus

[# Hops = 0; Modes used = Bus ; Update used]

Source: Aiims
Destination : Ina
Route Id : 543
SMS : Traffic is affected in th
SMS Time: 2012-07-05 10:08:29

Details :

Show <input type="text" value="10"/> entries		Search: <input type="text"/>				
ID	Message	Time	Name	Start Location	End Location	On Location
175	Traffic is moving in one lane only on Burari road due to MCD work in front of Delhi Jal Board office.msg@10.46am,230612.	23 Jun 2012 04:38:38 GMT	RepairWork	null	null	null
176	Traffic is affected at IGI stadium gate no.-16, from Raj ghat towards ITO.Due to break down of an container. Massage at 07:45 am Date-23/06/2012.	23 Jun 2012 02:16:30 GMT	BreakDown	Raj ghat	ITO	null
177	From AIIMS crossing, the traffic will not go on Aurobindo Marg due to Jagannath Yatra.msg@3.52pm,210612.	20 Jun 2012 22:22:45 GMT	Procession	null	null	null
178	Traffic will remain closed from IIT to AIIMS between 2pm to 3 pm and 4pm to 4.30pm And from AIIMS to IIT between 3pm to 4pm due to Jagannath Yatra.msg@12.30pm,210612.	21 Jun 2012 07:06:55 GMT	Procession	IIT	IIT between 3pm to 4pm	null

CSCE 580, 581 - FALL 2023

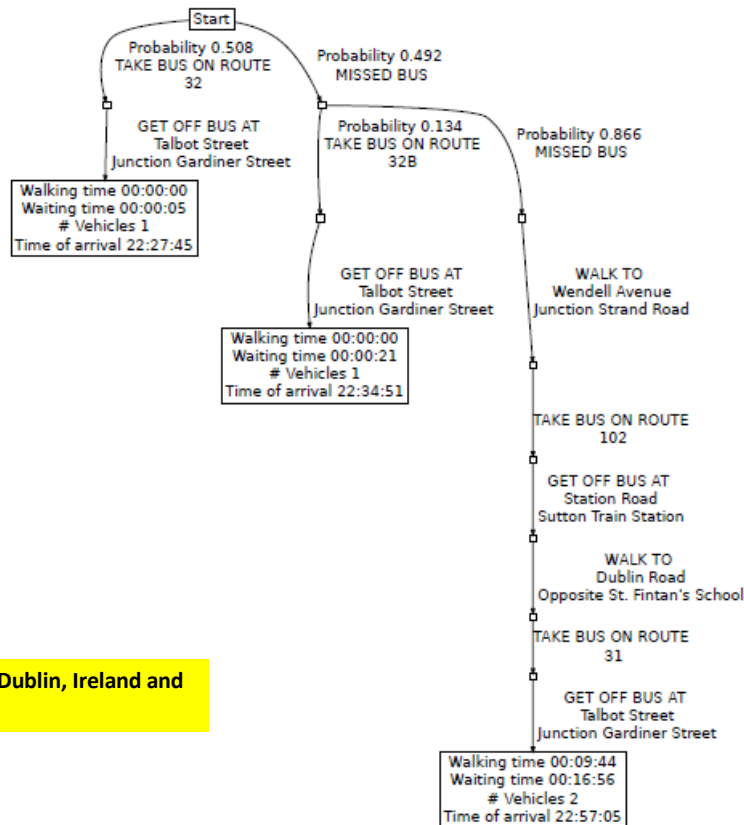
Increase Accessibility and Availability of Bus Information to Passengers (2013-end)

Kind of Information	Today Available to Bus Users	With Project in Bangalore	Mysore ITS (for reference)*	Benefit
Bus Schedule (static)	Available online and pamphlets	Available from low-cost phones (Spoken Web – Static)	Available online and pamphlets	Increase accessibility
Bus Schedule Changes (dynamic)	No information today	Will be available (Spoken Web - Human)	No information but in plan	Increase information
Bus Location	No information today	Will be available (GPS)	Will be available (GPS)	Increase information
Bus Condition	No information today	Will be available (Spoken Web - Human)	No information today	Increase information
Analytics (Bus Selection Decision Support)	No information today	Will be available (Transit)	No information but in plan	Increase information
Last –mile Connectivity to/ from nearest stop	No information today	Will be available (Spoken Web - Human)	No information today	Increase information
Standardization of information	No support	Will be supported (SCRIBE, Transit)	Some support due to GPS	Increase information's interoperability

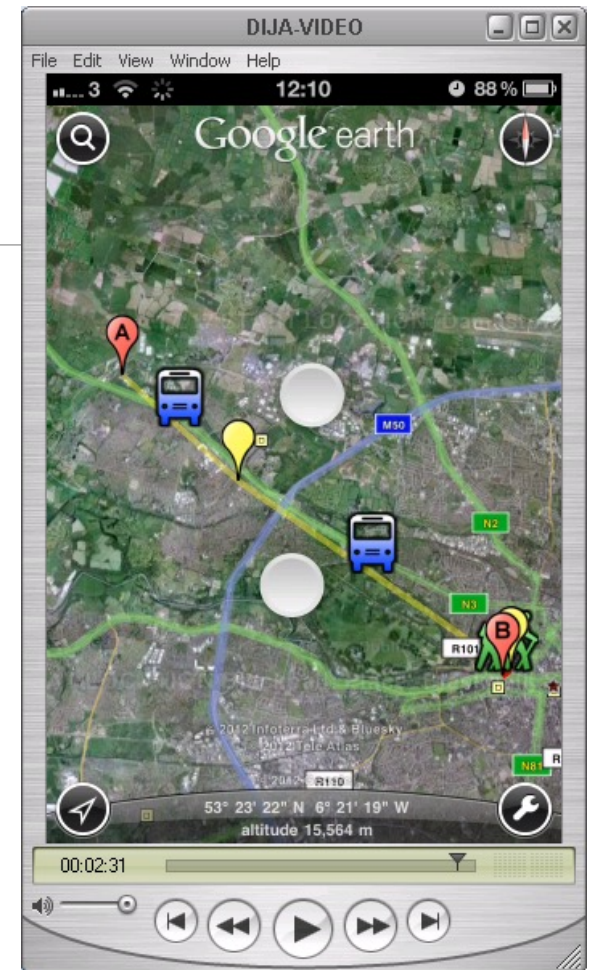
* Opinion based on only public information

End Vision: Information to Commuters to Reach Destination in All Eventuality

A Flexible Journey Plan

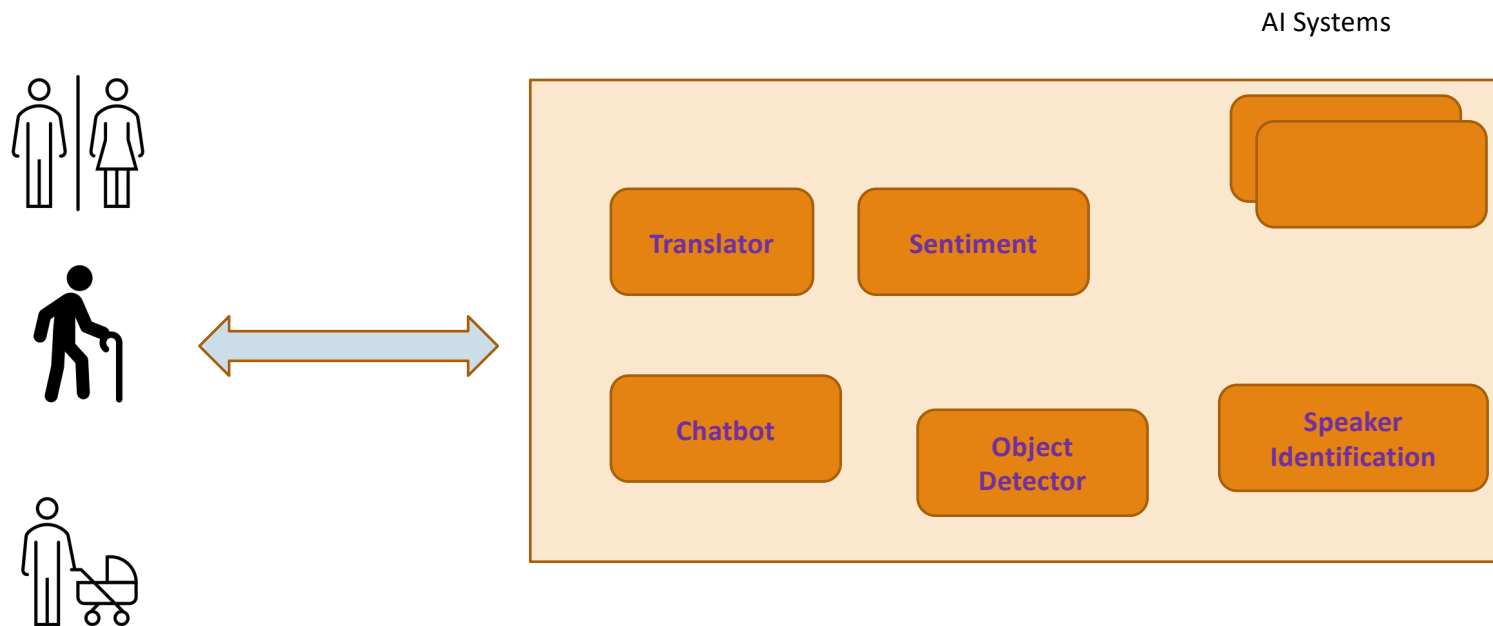


Pilots completed in Dublin, Ireland and
Montpellier, France



Docit: An Integrated System for Risk-Averse Multi-Modal Journey Advising, Adi Botea, Michele Berlingiero, Stefano Braghin, Eric Bouillet, Francesco Calabrese, Bei Chen, Yiannis Gkoufas, Rahul Nair, Tim Nonner, Marco Laumanns, IBM Technical Report, 2014

Technology and People



Trust: *Can people trust AI systems to perform capably, consistently, and with human values?*

Better Health

Two Tales from (Public) Health

Cutting-edge Technical Progress

- Enormous improvement in our understanding of diseases. E.g., Computational epidemiology
- Enormous advances in treating diseases are being made
 - We are living longer - A baby girl born in 2012 can expect to live an average of 72.7 years, and a baby boy to 68.1 years. This is 6 years longer than the average global life expectancy for a child born in 1990. (Source: WHO 2014 Health Statistics)
- Data on disease outbreaks is more available than ever before thanks to open data movement (E.g., data.gov, data.gov.in)

Stone-age Ground Reality

- Half of the top 20 causes of deaths in the world are infectious diseases, and maternal, neonatal and nutritional causes, while the other half are due to noncommunicable diseases (NCDs) or injuries. (Source: WHO 2014 Health Statistics)
- Worse – Indifference, mismanagement in response to communicable diseases - **late response to known diseases, in known period of the year**
 - E.g.: Japanese Encephalitis (JE) has been prevalent for ~3 decades in some parts of India killing 600+ every year
 - District level health experience is not reused over time and in similar regions

Ebola Data

International Bodies

The screenshot shows the eboladata.org website with the following content:

- World Health Organisation**
Ebola data and news maintained by the World Health Organisation.
4x [html](#)
- Ebola Treatment Centres, Isolation Wards, Hospitals and Transit Centres**
INACTIVE: Latest version of data is in <http://eboladata.org/dataset/ebola-treatment-centers-or-units>
3x [html](#)
- Ebola Treatment Centers or Units**
This dataset represents the best-known collection of status and location of the facilities known as Ebola Treatment Centers or Ebola Treatment Units in Guinea, Liberia and Sierra Leone.
- 1x [html](#)
- Sierra Leone Ministry of Health and Sanitation**
Updates and situation reports on disease incidence, treatment statistics, contact tracing
2x [html](#)
- Standby Task Force**
Comprehensive crowd-sourced maps of health facilities and related data: hospitals, laboratories, Ebola treatment centers, health centers, pharmacies
1x [html](#)
- EbolaBigData.org Case Counts**
Simple time series data of overall cases, deaths, and number of days.
1x
- Wikipedia Case Counts**
Wikipedia record of major Ebola virus outbreaks by country and by date – 3 September to most recent WHO / Gov update.
1x [html](#)

National Government

Crowd sourced

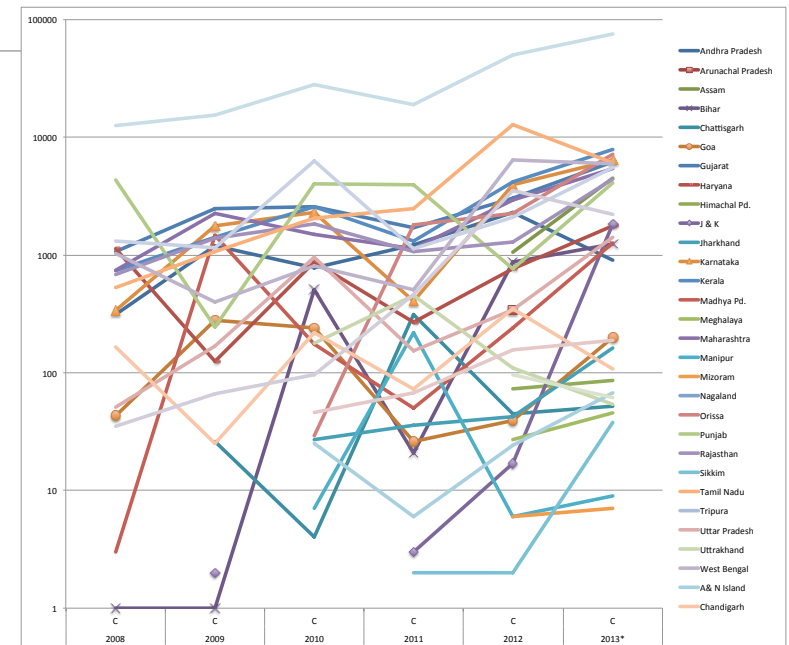
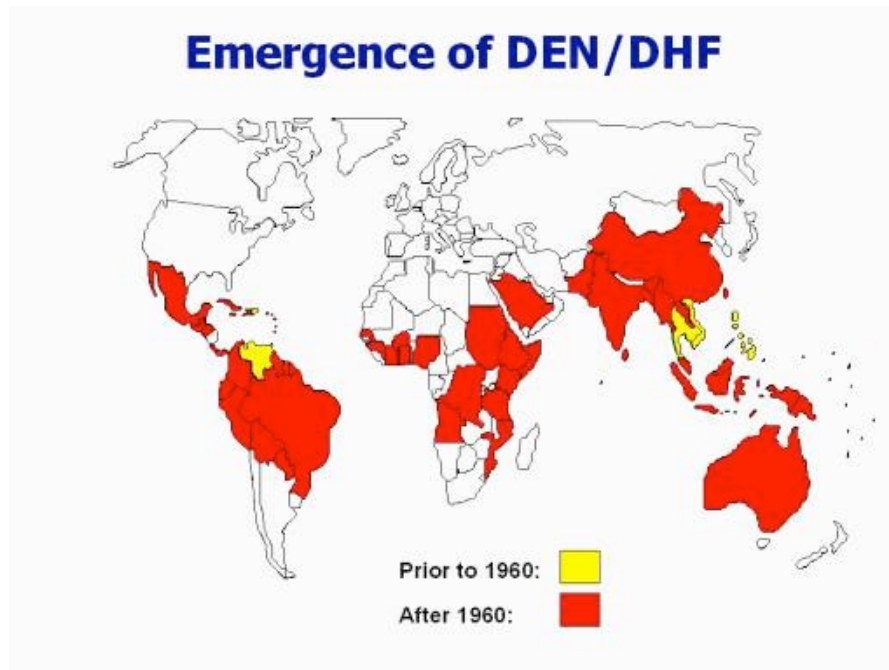
Online

The screenshot shows the EBOLA DATA JAM website with the following content:

- EBOLA DATA JAM**
- Datasets Visualizations search
- Welcome to DKAN site for the Ebola Open Data Jam.
- This project will take place in person on Saturday, October 18, 2014 at Open Space 156 5th Avenue, 2nd Floor, New York, NY (map)
- Datasets**
Below are a list of datasets that are available to make tools with. [csv](#) files can be downloaded or accessed through the DKAN Data Jam.
35 datasets
- World Health Organisation**
Ebola data and news maintained by the World Health Organisation.
4x [html](#)
- Ebola Treatment Centres, Isolation Wards, Hospitals and Transit Centres**
INACTIVE: Latest version of data is in <http://eboladata.org/dataset/ebola-treatment-centers-or-units>
3x [html](#)

Dengue

So, Do We Control Dengue Effectively? **NO**



Data for India

- Increasing number of states every year
- No consistent reduction of cases

Source: <http://nvbdcp.gov.in/den-cd.html>

Decisions in Public Health - Dengue

- **Identify:** Value of decision
 - What specific steps can we take to reduce Dengue and when?
 - How to choose among a set of possible options?
- **Assess:** Data-needed
 - Directly available
 - Available by proxy
- **Explore:** Methods
 - Decision theory (benefit v/s cost), minimize risks (deaths)
- **Conduct:** Evaluation
 - What metrics are relevant?
- **Integrate:** solution with overall process
 - How can the solution be integrated in overall process at a City?

Case Study: Dengue (Mosquito-borne)

Overall cost of a Dengue case is US\$ 828 (Sabchareon et al 2012).

From 9 countries in 1960s, it has spread to more than 110 countries now

Prevention methods

COMMUNITY

1. **Mosquito Coils & Candles:** The use of mosquito coils, candles & vapor mats indoors and outdoors of homes to combat mosquitoes.
2. **Window screens & Bed Nets:** The use of window screens in homes and bed nets in bedrooms to keep mosquitos out.
3. **Insecticide Application:** Application of insecticide to kill mosquitos that invade homes and surrounding areas.
4. **Larviciding at Home:** Application of larvicide in homes to kill larvae that live in stagnant water breeding sites like small ponds, gutters, cisterns, barrels, jars, and urns.
5. **Household/Community Cleanup:** Organize cleanups within communities in the surrounding housing areas and individual homes to recycle potential breeding sites like discarded plastic bottles, cans, old tyres, and any trash that can hold water for mosquitoes to breed in.

GOVERNMENT

6. **Surveillance For Mosquitoes:** Conduct periodical surveillance in hotspot areas and other communities to look for signs of mosquitoes.
7. **Medical Reporting:** To collate and compile reports of dengue cases and statistics to prioritize and focus dengue and vector mosquito control efforts and actions for best results.
8. **Effective Publicity & Campaigns:** To foster and champion effective campaigns amongst communities and create adequate public awareness of combating dengue.
9. **Enforcement:** Support and enforce the public and communities to practice effective dengue vector elimination under existing laws and implement new laws as appropriate for public health.
10. **Insecticide Fogging:** Conduct fogging in areas that have mosquitoes and dengue outbreak hotspots to kill adult mosquitoes.
11. **Public Education:** Foster, promote, and participate in public education in schools and all possible public meeting places to inform communities how to eliminate dengue vector mosquitoes, recognize early symptoms of the disease, and proper medical care and reporting.

CORPORATE

12. **Education:** To undertake community service initiatives and campaigns through marketing expertise and the media of TV, radio, and newspapers.
13. **PR/CSR:** To use public relations and customer service relations to reach communities on the fight against dengue.
14. **Adult Mosquito Traps:** To provide adult mosquito traps and other measures within the work areas to protect employees and workers from mosquitoes bites that transmit dengue.
15. **Mosquito Repellants:** Provide mosquito repellants to employees and workers within the work areas for further protection.
16. **Mosquito Control Materials, Methods, and Agents:** To provide the tools to the public and government that are necessary for dengue mosquito vector control like pesticides, biocontrol agents, mosquito traps, repellants, and other means to prevent dengue by eliminating the mosquito vectors.

WHO, 2013, Dengue Control. At <http://www.who.int/Denguecontrol/research/en/>, Accessed 21 June 2013.

Entogenex, 2013, Integrated Mosquito Management. At <http://www.entogenex.com/what-is-integrated-mosquito-management.html>,

Accessed 21 June 2013.

(ROI) Metrics - Illustration

Expense for disease control

- \$/person spent: How much money (in \$) is spent for a given method divided by the population of the region. Lower is better.

Impact of a disease control method

- Reduction: What is the magnitude of reduction in disease cases due to a method, expressed as a percentage, in a time period (e.g., year, disease season)? Higher is better.
- Cases/ person: How many reported cases of a disease occurred in a time period divided by the population of the region when a method was adopted? Lower is better.

Cost-effectiveness:

- Cases / \$: how many cases were reported for a disease per dollar spent on controlling it in a given time period? Lower is better.

Major Methods to Tackle Dengue

M1: Public awareness campaigns: to prevent conditions conducive to disease propagation, to improve reporting

M2: Chemical Control: Aerosol space spray

M3: Biological Control: Use of biocides

M4: Distributing equipments: bednets, insecticide- treated curtains

M5: Vaccination against the disease

Dengue Control Case Studies from Literature

Approach	Methods used (Mi)	Nature (Region, Population, area, year)	Expense per person	Reduction in number of cases
A1	M1, M2, M3	Sau Paulo, Brazil; 10,927,985; 2005	US \$1.14	34%
A2	M3	Puerto Rico; -; 2003	< US\$ 2.50	50% (in Dengue transmission)
A3	M2	Songkhla, Thailand; 162,645; 2009	US\$ 1.24	
A4	M5	Bang Phae, Thailand; 207,000; - AND Thailand; 4002; 2009 - 2014		0-70%, 30.2%

- An approach may use 1 or more method(s)
- They incur different costs per person
- Their efficacy is subject to various factors

Still, can we reuse these results in new areas?

Details:

Vandana Srivastava and Biplav Srivastava, Towards Timely Public Health Decisions to Tackle Seasonal Diseases With Open Government Data , International Workshop on the World Wide Web and Public Health Intelligence (W3PHI-2014), AAAI 2014

Challenge: Prescribe Methods to Use for a Hypothetical, Illustrative Area - Sundarpur

- City is Sundarpur
 - Made up of 10 districts
 - 10,000 people in each district.
- Disease control
 - Each district allocates \$10,000 per annum to prevent disease.
 - The city has a district-level health administrator per district and then an overall citywide public health administrator.
- What approach/ method should the district health officer use? What should the city health officer recommend?
 - a mix of control methods to produce the maximum reduction feasible.
 - Default option is to do nothing. This is unfortunately followed a lot!

Cost-benefits for Different Approaches

Approach Option	Population P	Amount available for expenditure (in USD) (a)	Expense per person for each method (in USD) (b)	Number of people exposed to the given method in the given amount $c = (a)/(b)$	Reduction in number of cases for each method (d)	Reduction in number of cases among exposed persons $(e) = (c)*(d)$	Effectiveness of the method $E = (e) / P$
O_def	10,000	10,000	0	0	0%	0	0%
O1_A 1	10,000	10,000	1.14	8772	34%	2982	30%
O2_A 2	10,000	10,000	2.5	4000	50%	2000	20%
O3_A 3	10,000	10,000	1.24	8065	10% *	806	8%
O4_A 4	10,000	10,000	8*	1250	70%	875	9%

* represents assumption made to compensate for missing data.

Prescription for Sundarpur

Best tactical option for administrators at Sundarpur (at district and the whole city level)

- is O1_A1 since it brings the maximum reduction.
- If the administrators are interested to cover the maximum number of people in the given budget, the best method is still O1_A1.
- If the administrators are interested to show maximum reduction in cases for a pocket of the city (sub-district level which may be more prone to the disease), they may choose O4_A4 but it costs maximum and thus can be perceived as taking resources away from the not- directed areas.

Strategic option

- Select top-2 (O1_A1 and O2_A2), and try them in 5 districts each in one year. It hedges risk of variability between Sundarpur and old location of previous studies.
- Based on efficacy, decide the single best option for Sundarpur in subsequent year.
- She may also use the vaccine option only when the disease outbreak is above certain threshold.

Details:

Vandana Srivastava and Biplav Srivastava, Towards Timely Public Health Decisions to Tackle Seasonal Diseases With Open Government Data , International Workshop on the World Wide Web and Public Health Intelligence (W3PHI-2014), AAAI 2014

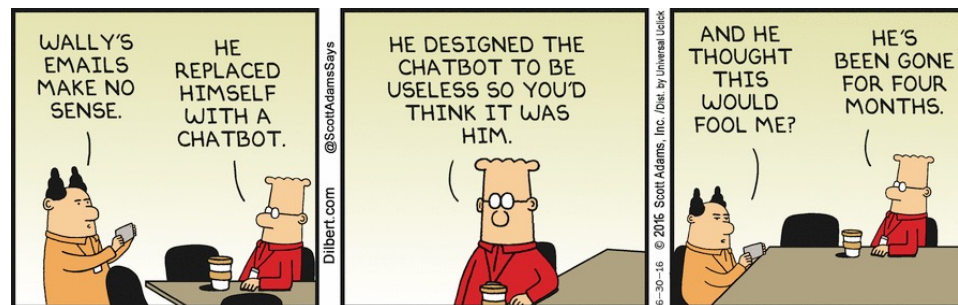
New Data Practices

- Find correlation among methods (positive or negative)
 - We assumed independence
 - Needs: Historic Data, Experiment Design
- Learn rate of return for approaches and methods (new combinations not tried in health literature)
 - Need: Collect data on efficacy of method individually
- Find similarity among regions
 - Data Need: Spatio-temporal modeling/ STEM
- Multi-objective optimization
 - Examples: Effectiveness of approach, Reduction of case, people coverage
 - Needs: Data about approaches tried historically

Request to Medical Community on Data

- Report both **cost** and **effectiveness** of approaches and methods
 - Overlooking one hampers reuse of results
- Interact with AI community to learn and try mixed approaches that reduce cost and improve overall effectiveness
 - All combinations cannot be tried on the ground due to practical constraints
 - Get more effective approaches rolled out faster targeted to new regions

Solving Real Business Problems



Credit: Dilbert – June 30, 2016

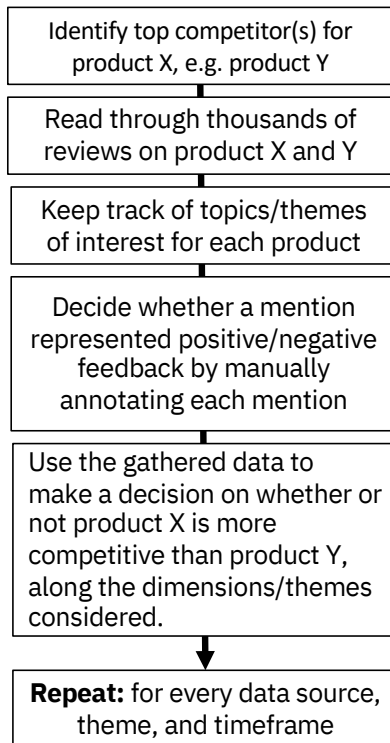
Case Study of Sentiment in Business

Clarity: Data-Driven Competitive Analysis

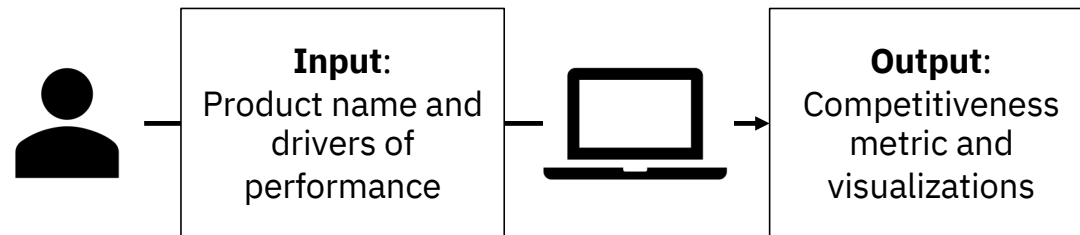
1. Sheema Usmani, Mariana Bernagozzi, Yufeng Huang, Michelle Morales, Amir Sabet Sarvestani, Biplav Srivastava, Clarity: Data-driven Automatic Assessment of Product Competitiveness, IAAI/AAAI 2020, **Deployed Application Award**
2. (Demo paper) Data-driven ranking and visualization of products by competitiveness, Sheema Usmani, Mariana Bernagozzi, Yufeng Huang, Michelle Morales, Amir Sabet Sarvestani, Biplav Srivastava, AAAI 2020
3. [Yufeng Huang](#), [Mariana Bernagozzi](#), [Michelle Morales](#), [Sheema Usmani](#), Biplav Srivastava, [Michelle Mullins](#), Clarity 2.0: Improved Assessment of Product Competitiveness from Online Content. [AI Mag. 42\(2\)](#): 59-70 (2021)

Competitive Analysis: Before & After

Today's Manual Process



New Process



Steps

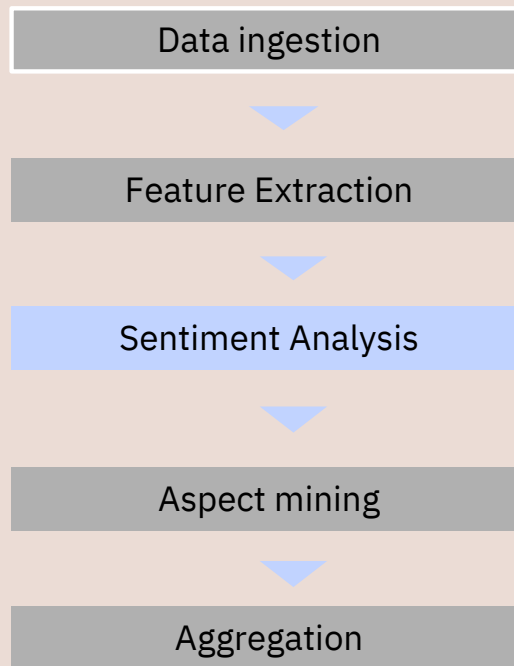
1. Prepare review data of products p_1 to p_N from sources d_1 to d_M (offline)
2. Process request for analysis for product p_i (online)
3. Visualize analysis results (online, optional)

Illustrative Output

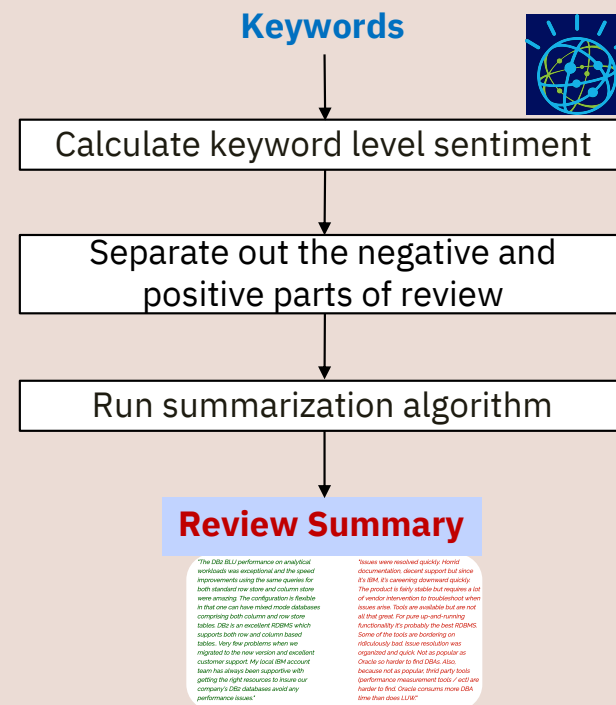


Clarity Score and Trends

Methodology



Sentiment Analysis



Course Project

Project Discussion: What Problem Fascinates You ?

- Data
 - Water
 - Finance
 - ...
- Analytics
 - Search, Optimization, Learning, Planning, ...
- Application
 - Building chatbot
- Users
 - Diverse demographics
 - Diverse abilities
 - Multiple human languages

Project execution in sprints

- Sprint 1: (Sep 12 – Oct 5)
 - **Solving**: Choose a decision problem, identify data, work on solution methods
 - **Human interaction**: Develop a basic chatbot (no AI), no problem focus
- Sprint 2: (Oct 10 – Nov 9)
 - **Solving**: Evaluate your solution on problem
 - **Human interaction**: Integrated your choice of chatbot (rule-based or learning-based) and methods
- Sprint 3: (Nov 14 – 30)
 - **Evaluation**: Comparison of your solver chatbot with an LLM-based alternative, like ChatGPT

Project Discussion: Dates and Deliverables

Project execution in sprints

- Sprint 1: (Sep 12 – Oct 5)
 - **Solving**: Choose a decision problem, identify data, work on solution methods
 - **Human interaction**: Develop a basic chatbot (no AI), no problem focus
- Sprint 2: (Oct 10 – Nov 9)
 - **Solving**: Evaluate your solution on problem
 - **Human interaction**: Integrated your choice of chatbot (rule-based or learning-based) and methods
- Sprint 3: (Nov 14 – 30)
 - **Evaluation**: Comparison of your solver chatbot with an LLM-based alternative, like ChatGPT

- Oct 12, 2023
 - Project checkpoint
 - In-class presentation
- Nov 30, 2023
 - Project report due
- Dec 5 / 7, 2023
 - In-class presentation

Skeleton: A Basic Chatbot

- Run in an infinite loop until the user wants to quit
- Handle any user response
 - User can quit by typing “Quit” or “quit” or just “q”
 - User can enter any other text and the program has to handle it. The program should write back what the user entered and say – “I do not know this information”.
- Handle known user query types // Depends on your project
 - “Tell me about N-queens”, “What is N ?”
 - “Solve for N=4?”
 - “Why is this a solution? ”
- Handle chitchat // Support at least 5, extensible from a file
 - “Hi” => “Hello”
 - ...
- *Store session details in a file*

Illustrative Project

1. **Title:** Solve and explain solving of n-queens puzzle
2. **Key idea:** Show students how a course project will look like
3. **Who will care when done:** students of the course, prospective AI students and teachers
4. **Data need:** n: the size of game; interaction
5. **Methods:** search
6. **Evaluation:** correctness of solution, quality of explanation, appropriateness of chat
7. **Users:** with and without AI background; with and without chess background
8. **Trust issue:** user may not believe in the solution, may find interaction offensive (why queens, not kings? ...)

Project Discussion: Illustration

1. Create a private Github repository called “CSCE58x-Fall2023-<studentname>-Repo”. Share with Instructor (biplav-s) and TA (kausik-l)
2. Create Google folder called “CSCE58x-Fall2023-<studentname>-SharedInfo”. Share with Instructor (prof.biplav@gmail.com) and TA (lakkarajukausik90@gmail.com)
3. Create a Google doc in your Google repo called “Project Plan” and have the following by next class (Sep 5, 2023)

1. **Title:** Solve and explain solving of n-queens puzzle
2. **Key idea:** Show students how a course project will look like
3. **Who will care when done:** students of the course, prospective AI students and teachers
4. **Data need:** n: the size of game; interaction
5. **Methods:** search
6. **Evaluation:** correctness of solution, quality of explanation, appropriateness of chat
7. **Users:** with and without AI background; with and without chess background
8. **Trust issue:** user may not believe in the solution, may find interaction offensive (why queens, not kings? ...)

Project Illustration: N-Queens

- Sprint 1: (Sep 12 – Oct 5)
 - **Solving**: Choose a decision problem, identify data, work on solution methods
 - Method 1: Random solution
 - Method 2: Search – BFS
 - Method 3: Search - ...
 - **Human interaction**: Develop a basic chatbot (no AI) as outlined
 - Deliverable
 - Code structure in Github
 - ./data
 - ./code
 - ./docs
 - ./test
 - Presentation: Make sprint presentation on Oct 12, 2023

Reference: Project Rubric - NEW

- **Project report – 60%**
 - Project description: problem, related work, approach, evaluation – 40%
 - Working system demo/ video – 10%
 - Well organized Github with code (./data, ./code, ./docs, ./test) – 10%
- **Project presentation – 40%**
 - Evaluation by peers, instructor and TA
- **Bonus**
 - Instructor discretion – 10%
- **Penalty**
 - Lack of timeliness as per announced policy (right) - up to 30%

Milestones and Penalties

- Oct 12, 2023
 - Project checkpoint
 - In-class presentation
 - **Penalty: presentation not ready by Oct 10, 2023 [-10%]**
- Nov 30, 2023
 - Project report due
 - **Project report not ready by date [-10%]**
- Dec 5 / 7, 2023
 - In-class presentation
 - **Project presentations not ready by Dec 4, 2023 [-10%]**

Evaluation of Presentation

1. An online form will be available during presentation
2. During a presentation, three students will be assigned to review along with instructor and TA
3. They will enter following survey questions:
 1. Their name
 2. Presentation number
 3. How useful is the system – will you use it? [1-5 scale]
 4. How well have you understood the project from the presentation? [1-5 scale]
4. Top and bottom scores will be removed. Average of remaining three will be used for final presentation marks

Lecture 27: Summary

- We talked about
 - Real world problems
 - Smart city – setting goals for improvement
 - Framework for identifying opportunities to solve problems with AI
 - Case studies in smart city (traffic, public health) and business (Clarity - business intelligence)

Concluding Section

About Next Lecture – Lectures 28 and 29

Course Project Presentation
