



CSCE 580: Introduction to AI

Lecture 28: AI for the Real World

PROF. BIPLAV SRIVASTAVA, AI INSTITUTE

5TH DEC, 2024

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

Credits: Copyrights of all material reused acknowledged

Organization of Lecture 28

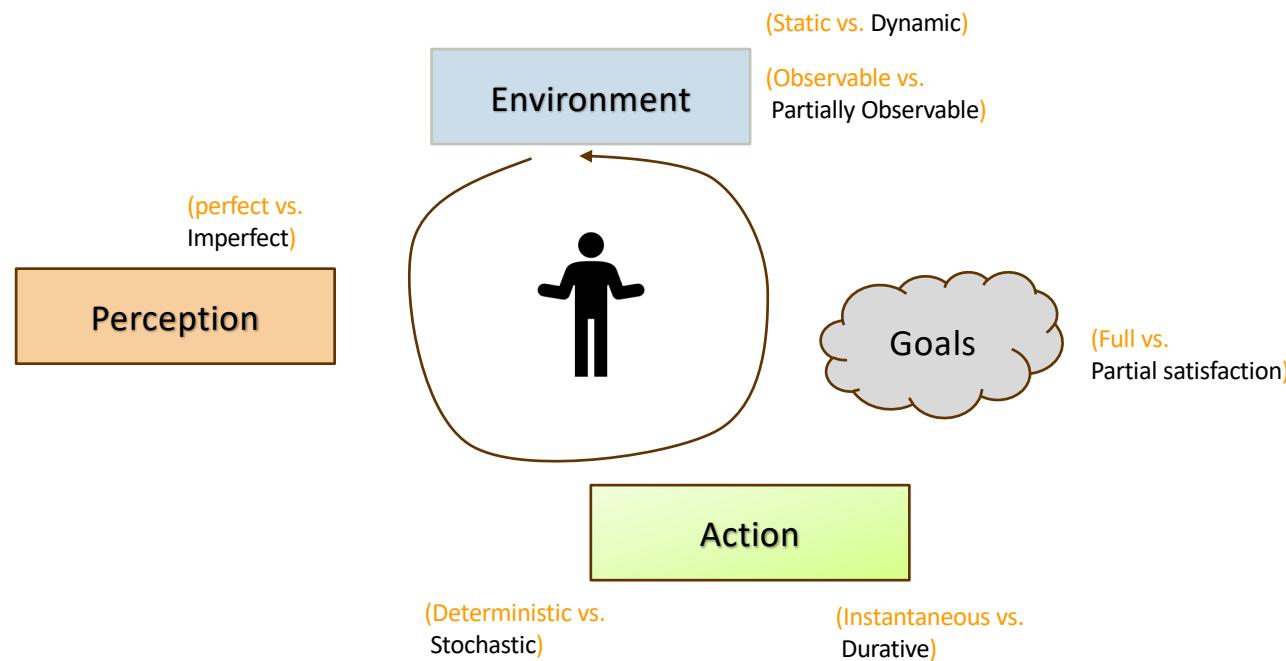
- Introduction Segment
 - Recap of Lectures 26, 27
- Main Segment
 - Bringing All Together
 - Real world problems
 - Smart city – setting goals for improvement
 - Framework for identifying opportunities to solve problems with AI
 - Case studies in smart city (public health) and business (Clarity - business intelligence)
- Concluding Segment
 - Ask me anything

Previous Classes

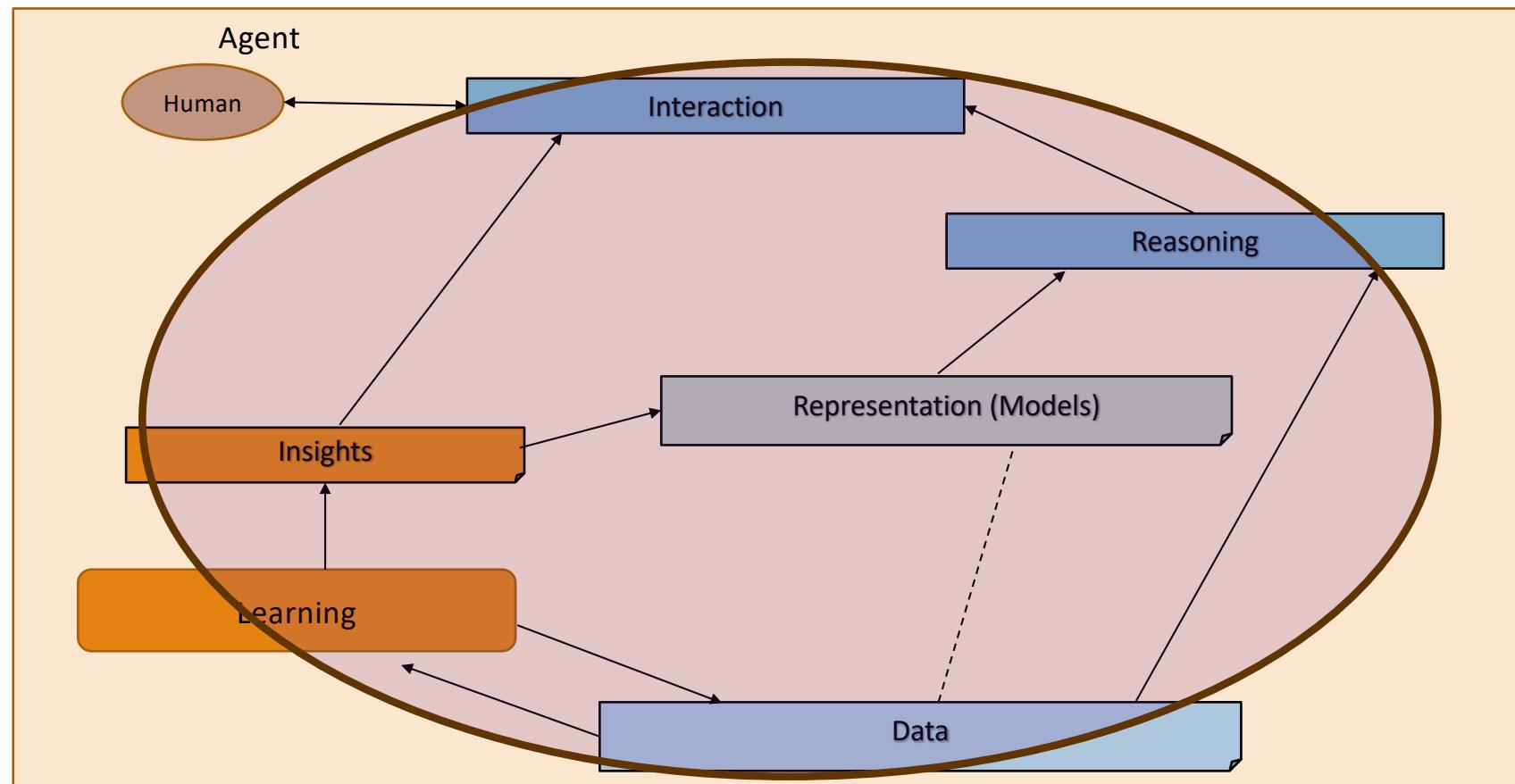
- Lecture 26: graduate student papers
 - Presentation marks posted (grad students)
 - Report marks tbd (grad students)
 - Undergrad attendance bonus marks – as applicable (undergrad)
- Lecture 27: project #2 presentations
 - Difference in context, marks given for attempting

Introduction Section

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
 - **Governance**
- What can AI do here?



Major Community Challenges

Challenges

Growing population

- Access to safe water
- Need jobs

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

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



Reducing resources

- Land
- Potable water

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?

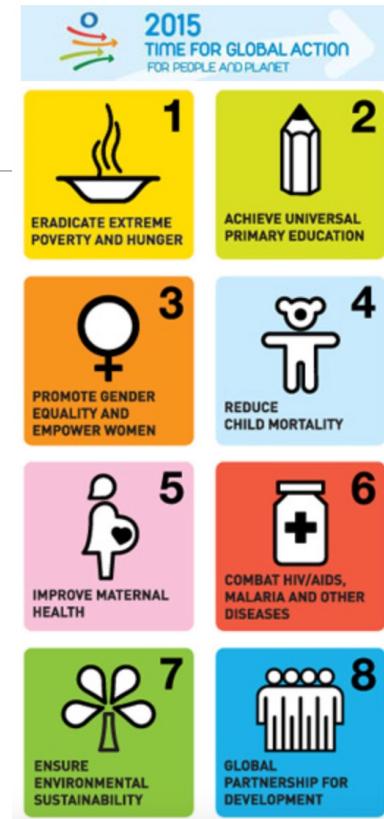
Smart city can mean one or more of the following:

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.

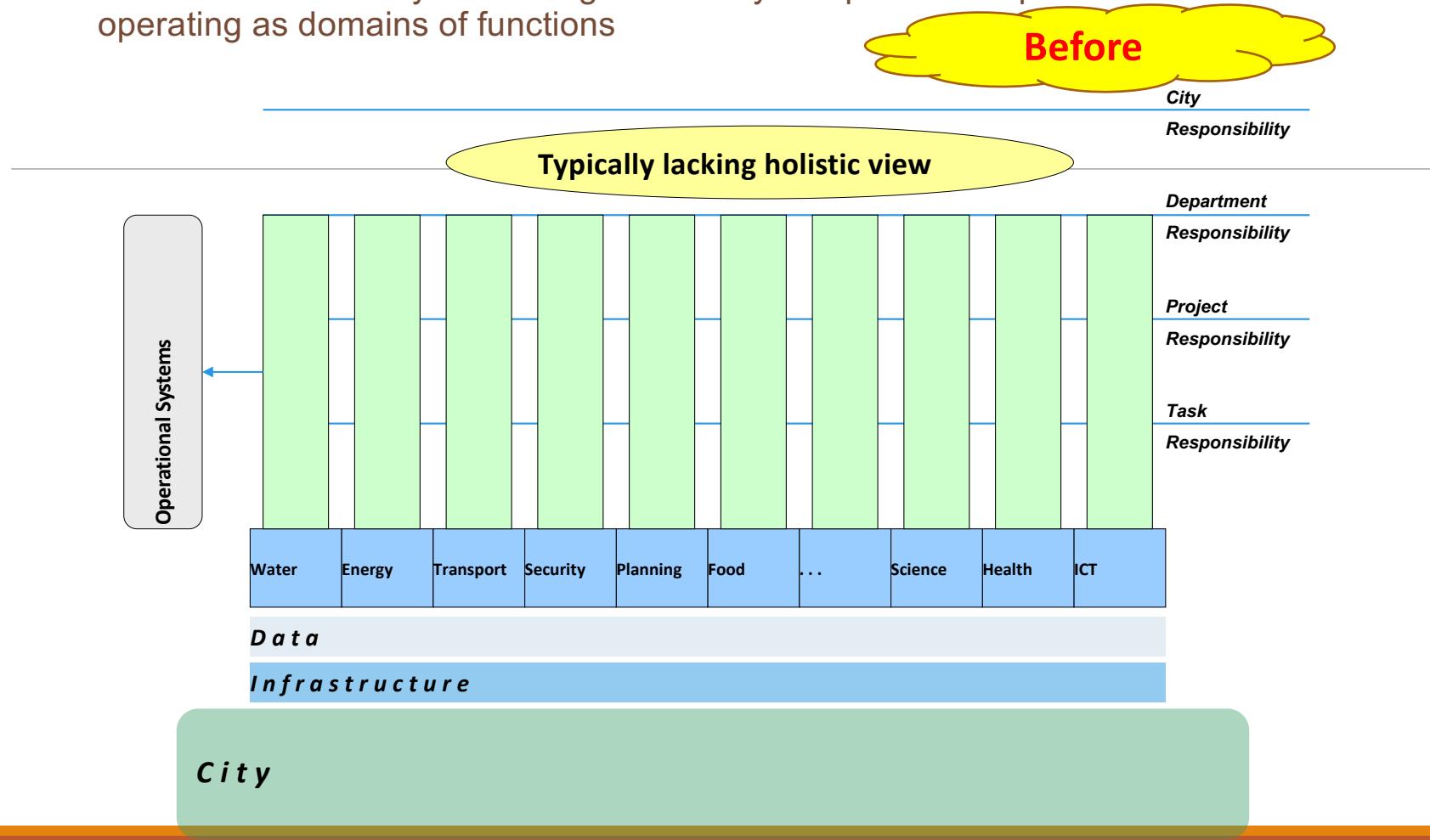
A city leadership can choose among these or define their own objective(s) and manage with measurements to pro-actively achieve it



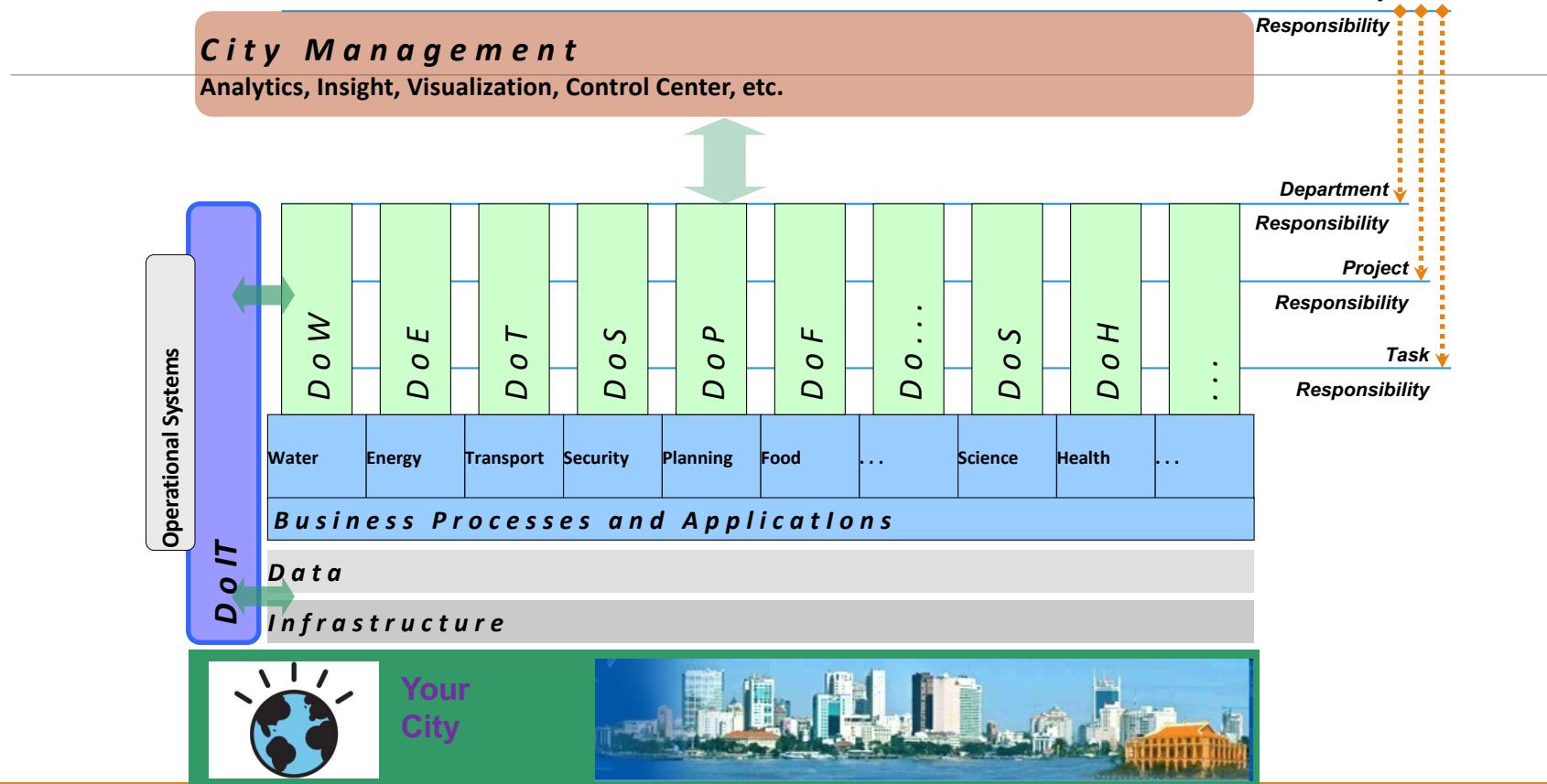
See other FAQs at: <https://sites.google.com/site/biplavsrivastava/research-1/intelligent-systems/scfaqs>

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

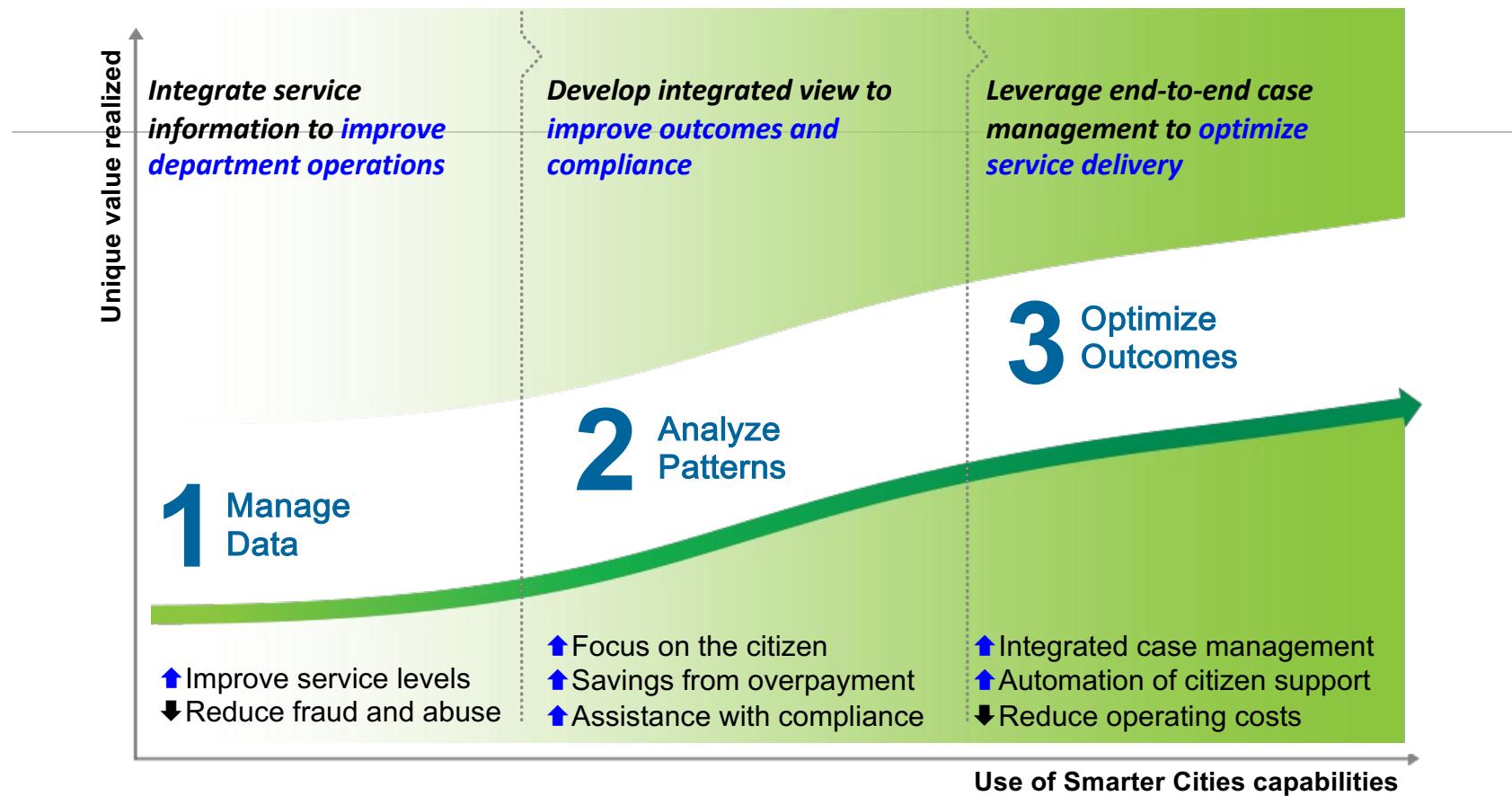
Cities are traditionally built and governed by independent departments operating as domains of functions



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



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) Shahdara (T)
Destination	1. INDIRA PURI LONI BORDER 2. P.S. INDIRA PURI 3. JAWAHAR NAGAR 4. INDIRA PURI XING 5. SHAHDRA.
Bus Stop	

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, Dilshad Gerden,5:30:00,
First Metro Time	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 Djikstra'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:

Search

13 solutions found. Please click on the links to explore

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

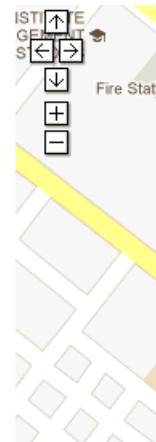
Source: aiims
Destination : rajiv chowk
Route Id : LINE-2
Starting Time : 17:00
Reaching Time : 17:22
Frequency : 005
Mode : metro

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

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

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



IRL - Transit - Search service details page

Plan trip understanding the best means of transportation

Get public transit directions

City:

From:

To:

Start: End:

Mode:

No. of choices:

No. of hops:

Search

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

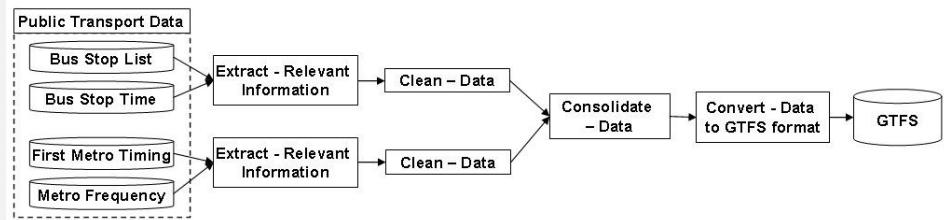
Source: majestic (kbs) Bgl
Destination : jn. of nice rd Bgl
Route Id : G-7
Mode : bus

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

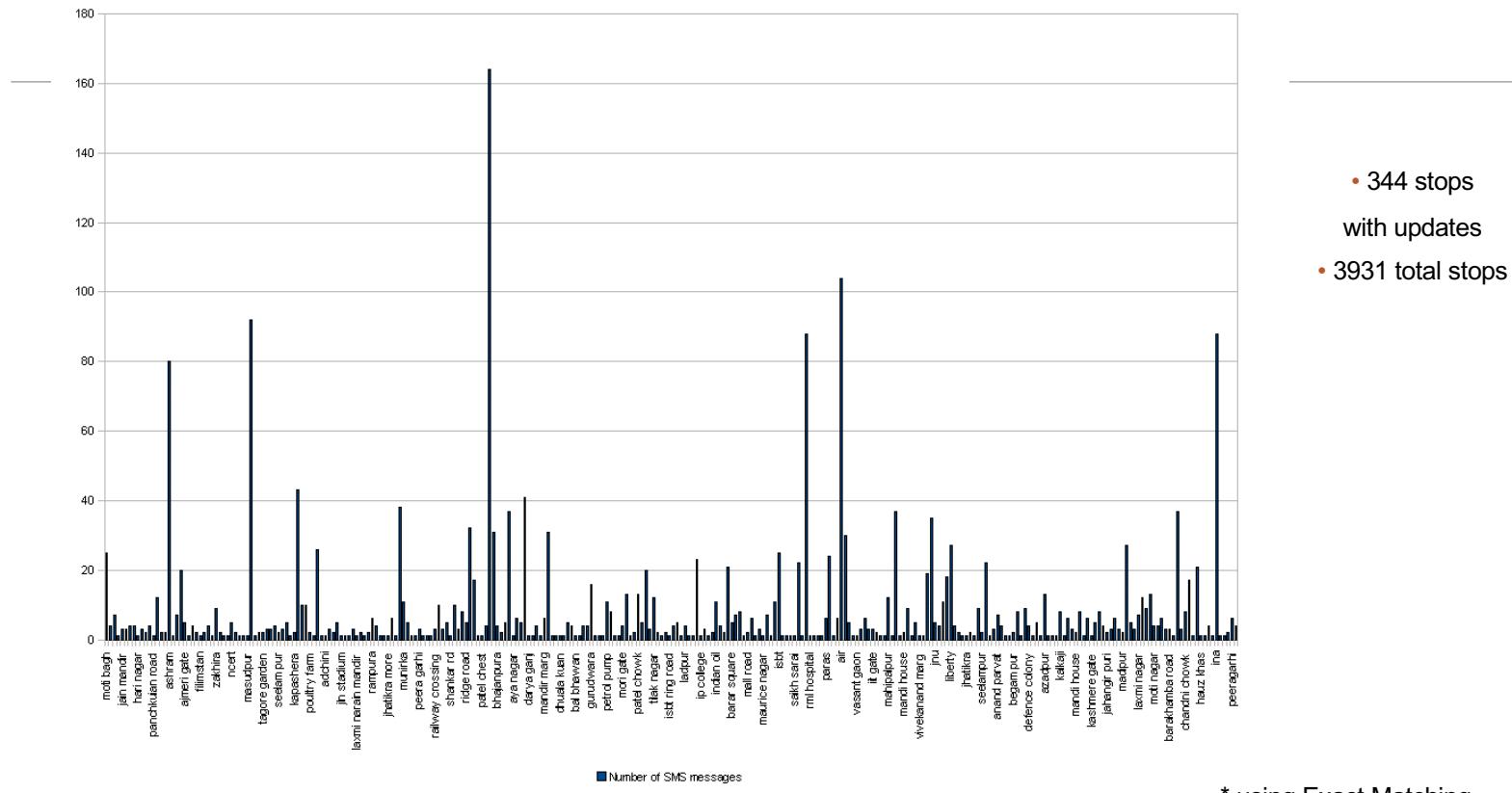
- Output

- 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, Biplay 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: Delhi

Search From: aiim

aiims

Search To: ina

ina

Start: ANY

Mode: ANY

City Update: ANY

No. of choices: ANY

No. of hops: ANY

Search

Clear

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

↳ Solution 1[# Hops = 0; Modes used = Bus]

↳ Solution 2[# Hops = 0; Modes used = Bus]

↳ Solution 3[# Hops = 0; Modes used = Bus]

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

↳ Solution 4[# 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 :

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. Message 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

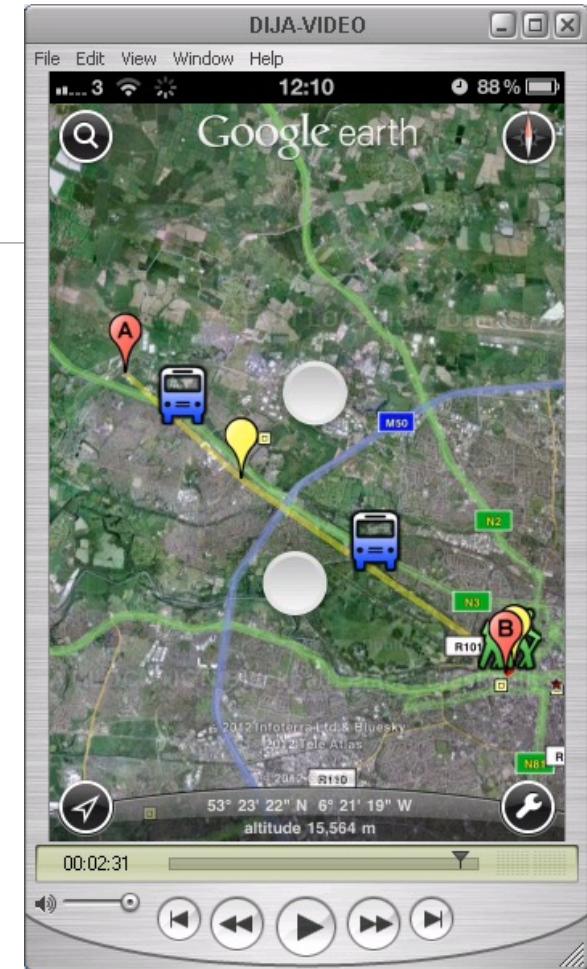
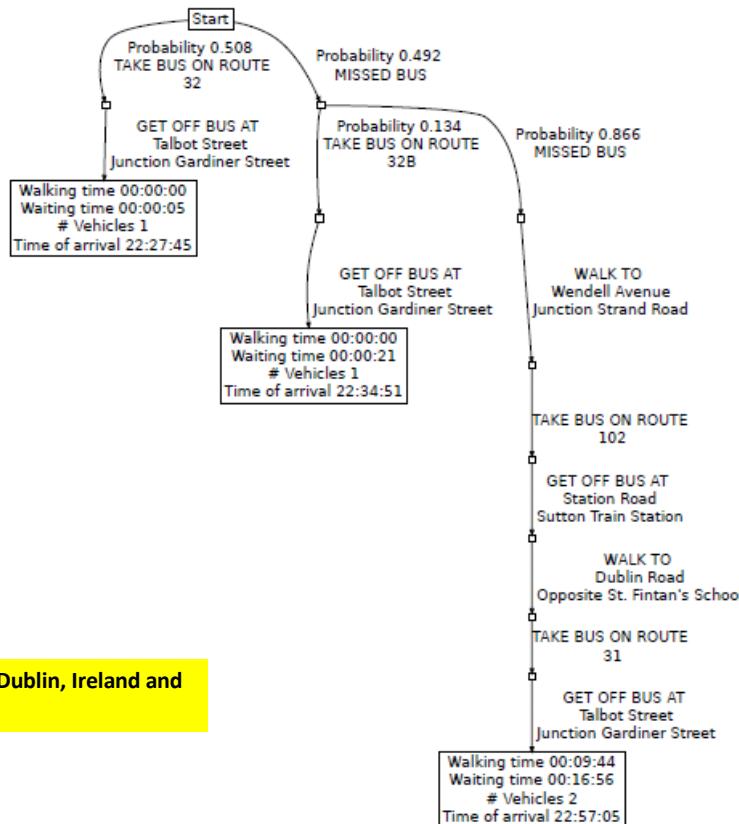
Increase Accessibility and Availability of Bus Information to Passengers (2013-end)

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

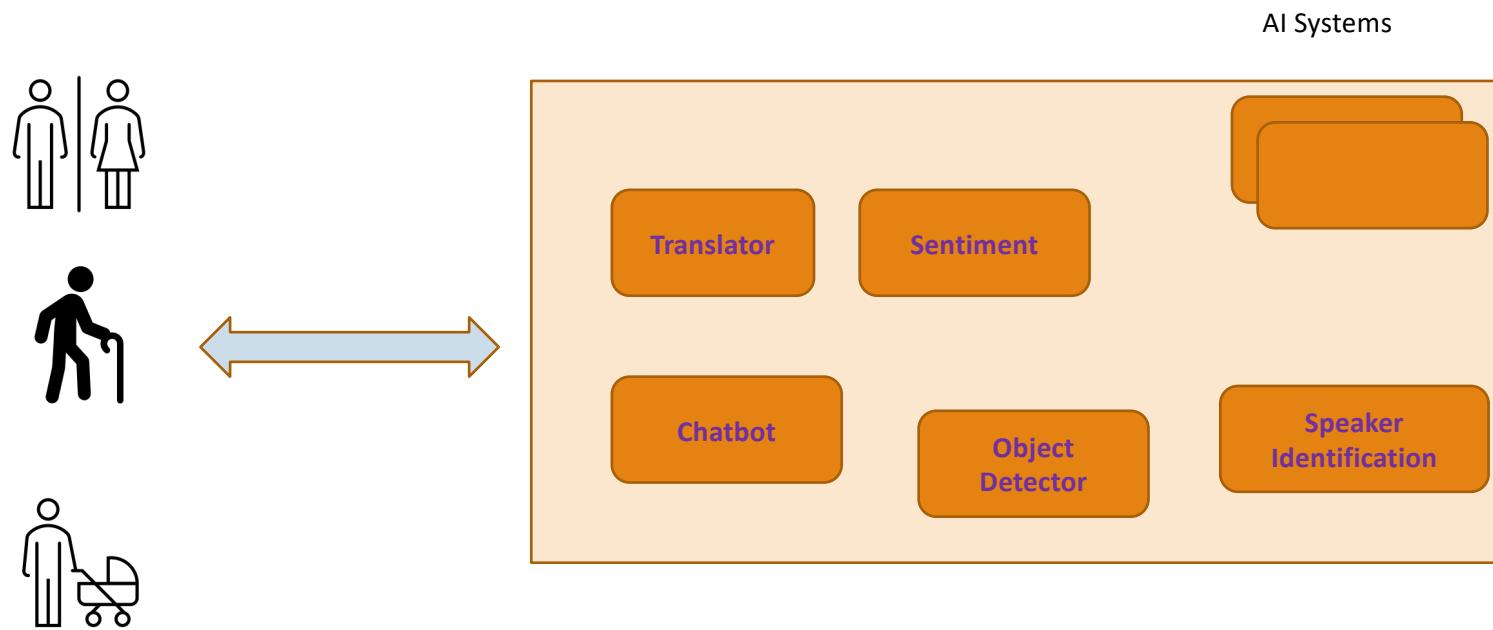
End Vision: Information to Commuters to Reach Destination in All Eventuality

A Flexible Journey Plan



Docit: An Integrated System for Risk-Averse Multi-Modal Journey Advising, Adi Botea, Michele Berlingario, 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

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 [html](#)

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

EBOLA DATA JAM

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 35 datasets

World Health Organisation
Ebola data and news maintained by the World Health Organisation.

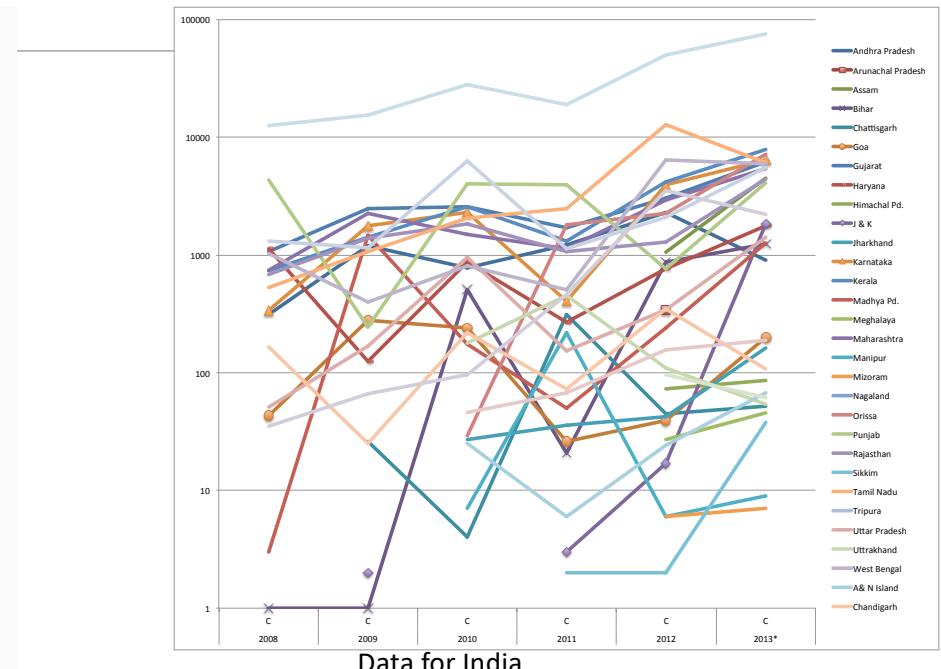
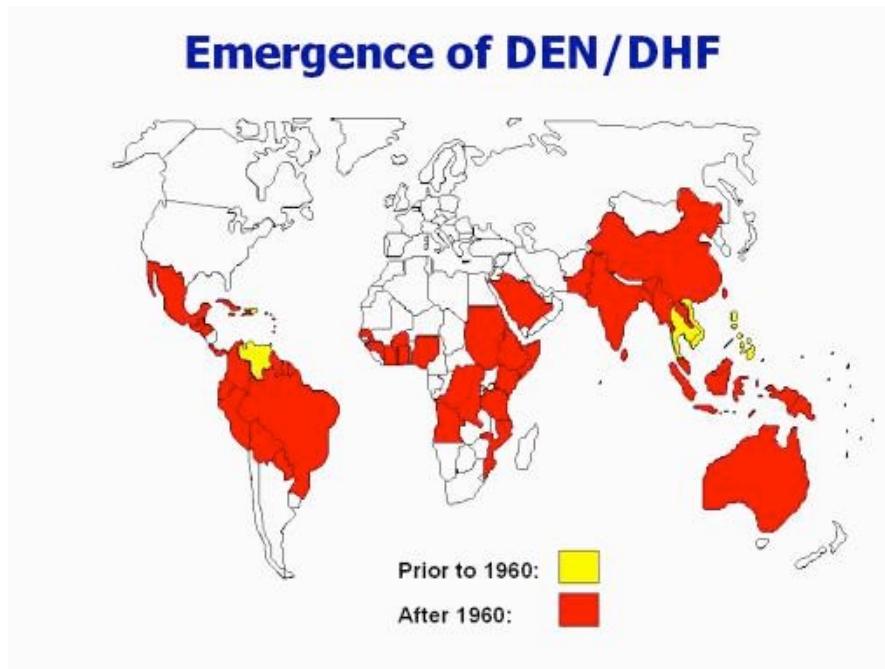
4x [html](#)

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**



Source: <http://nvbdco.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

Ap pro ach	Method s used (Mi)	Nature (Region, Population, area, year)	Expens e 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

Better Governance

Context: Elections



<https://founders.archives.gov/documents/Adams/99-02-02-6371>

"I do not say that democracy has been more pernicious on the whole, and in the long run, than monarchy or aristocracy. **Democracy has never been and never can be so durable as aristocracy or monarchy; but while it lasts, it is more bloody than either. ... Remember, democracy never lasts long. It soon wastes, exhausts, and murders itself. There never was a democracy yet that did not commit suicide.** It is in vain to say that democracy is less vain, less proud, less selfish, less ambitious, or less avaricious than aristocracy or monarchy. It is not true, in fact, and nowhere appears in history. Those passions are the same in all men, under all forms of simple government, and when unchecked, produce the same effects of fraud, violence, and cruelty. When clear prospects are opened before vanity, pride, avarice, or ambition, for their easy gratification, it is hard for the most considerate philosophers and the most conscientious moralists to resist the temptation. Individuals have conquered themselves. Nations and large bodies of men, never."

John Adams, one of the Founding Fathers of the United States

The Semantics We Will Use

- What is Democracy ?
 - People electing people govern people, independently

- What is voting v/s electing?
 - Voting is exercising right to make a choice (including not voting for available candidates)
 - Electing is selecting a candidate from the panel available (≥ 2 people)
- Party v/s people
 - Democracy cannot be complete when restriction is put on people - e.g., voters or candidates
 - Democracy can be present without any party
 - But if there are parties, they should be two or more; if there is one party, independent candidates have to be allowed - otherwise, there is restriction on candidates
 - Democracy cannot be present without empowered voters
 - Voters should be able to make independent decisions
- Why is the topic (“Credible Elections”) important?
 - Governance model that respects individual liberty
 - Best model known to ensure peaceful transition of governance over time

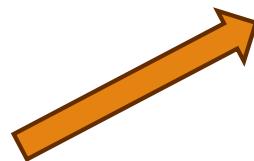
The Election's Challenge

- The voting turnout rate (62.8%, 2020) in the U.S. is very low in comparison with other countries [1].
- Promoting voter participation using traditional approaches is often costly, time-consuming, and yielding little to no results.

[1] D. Desilver, [Turnout in U.S. has soared in recent elections but by some measures still trails that of many other countries](https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/),
<https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/>, PEW Research, 2022

Election Ecosystem and PROMISE Team

- Technology (AI)
 - Data (creation, access, provenance)
 - Improve competence (performance, robustness)
 - Transparency and explainability
- People
 - Needs and wants, behavior (individual)
 - Belief, Influence, ... (group dynamics)
 - Diversity – culture, ethics (across the world)
- Human-AI collaboration
 - Human centered design
 - Improve access (for diverse backgrounds)
- Frameworks, Standards, Funding
 - Benchmarks and reference standards for tackling technical challenges without politicizing
 - NIST AI Safety Consortium



Stakeholders considered

- **Voters (V):** *make choices*
- **Candidates (C):** *available to make choices*
- **Election Commission (EC):** *conduct election*

Others, not considered

- Media
- Equipment vendors
(Technology, material,...)

A New Approach

The Case for Official Information to Tackle Information Disorders

Biplav Srivastava, A Vision for Reinventing Credible Elections with Artificial Intelligence,
Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI-25), Philadelphia, USA, Feb
2025, [AI, Elections]

Terminology

- **Information disorder:** misinformation, disinformation, malinformation
- **Official information:** providers required by law to give right and timely information – e.g., authoritative agencies (elected officials, election officials); and if they do not, they can be held accountable via legal means

The Election's Challenge and Solution Value

Challenges

- The voting turnout rate (62.8%, 2020) in the U.S. is very low in comparison with other countries [1].
- Promoting voter participation using traditional approaches is often costly, time-consuming, and yielding little to no results.

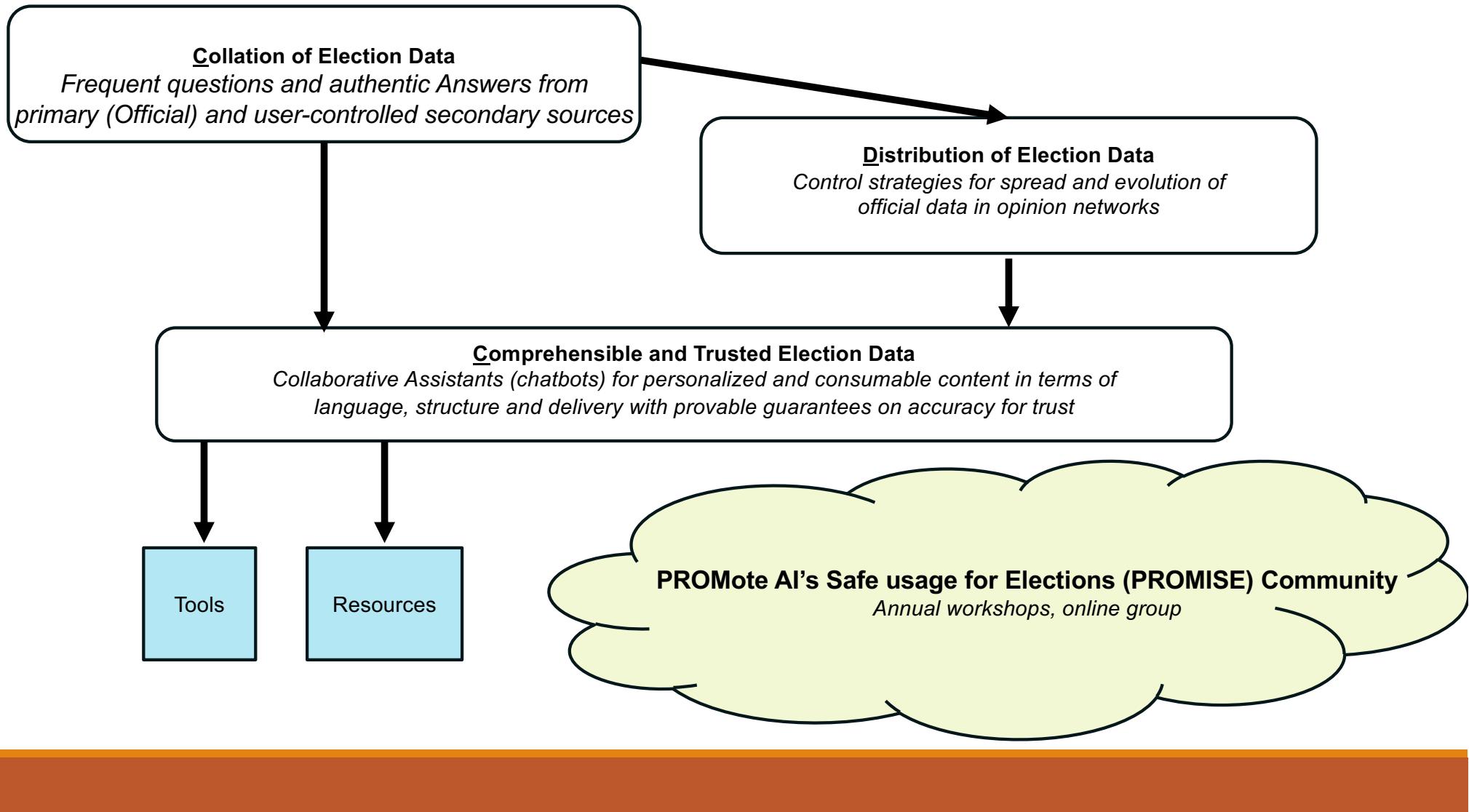
Hypothesis: Using technology for voter engagement, especially among seniors and youths, will increase participation [2].

CDC approach

- **Collate:** Collate good/ official information
- **Distribute:** Safely distribute in opinion and social networks
- **Comprehension:** Make content easy to understand

[1] D. Desilver, [Turnout in U.S. has soared in recent elections but by some measures still trails that of many other countries](https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/),
<https://www.pewresearch.org/short-reads/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/>, PEW Research, 2022

[2] Bergan, D. E., Carnahan, D., Lajevardi, N., Medeiros, M., Reckhow, S., & Thorson, K. (2022). Promoting the youth vote: the role of informational cues and social pressure. *Political Behavior*, 44(4), 2027-2047.



The Case for Official Information

- Information disorder is prevalent
 - Why? – (**hypothesis**): Due to lack of understanding and effort to promote official information.
- Official information: Anyone, other than official sources, giving information are open to doubt.
 - Analogy: When the state executes a person, it is called justice; when a person executes another, it is called a murder.
- Continuous efforts are needed to improve official information ecosystem: better collection, preservation and dissemination.
- Need to provide official data easily to AI tools that in turn can help stakeholders make better decisions.

Number of Queries and Sources

Country	No. of Queries	Sources
India (IND)	11	https://www.vice.com/en/article/gy4x3m/10-most-googled-questions-about-indian-election
United States of America (USA)	39	https://www.usa.gov/how-to-vote ; https://www.cnn.com/interactive/2020/politics/voting-questions-answers/
Indonesia (IDN)	6	https://www.vice.com/en/article/neawv8/most-googled-questions-about-indonesian-election-2019
Ireland (IRE)	11	https://www.citizensinformation.ie/en/government-in-ireland/elections-and-referenda/types-of-elections-and-referendums/
United Kingdom (UK)	7	https://fullfact.org/election-2019/election-questions-answered/
Estonia (EST)	50	https://www.valimised.ee/en/internet-voting/frequently-asked-questions/
South Africa (SA)	47	https://www.elections.org.za/content/for-voters/FAQ-For-Voters/
Canada (CA)	11	https://www.tvo.org/article/your-federal-election-questions-answered

Table 1: The number of queries collected from each country with the corresponding country name and source.

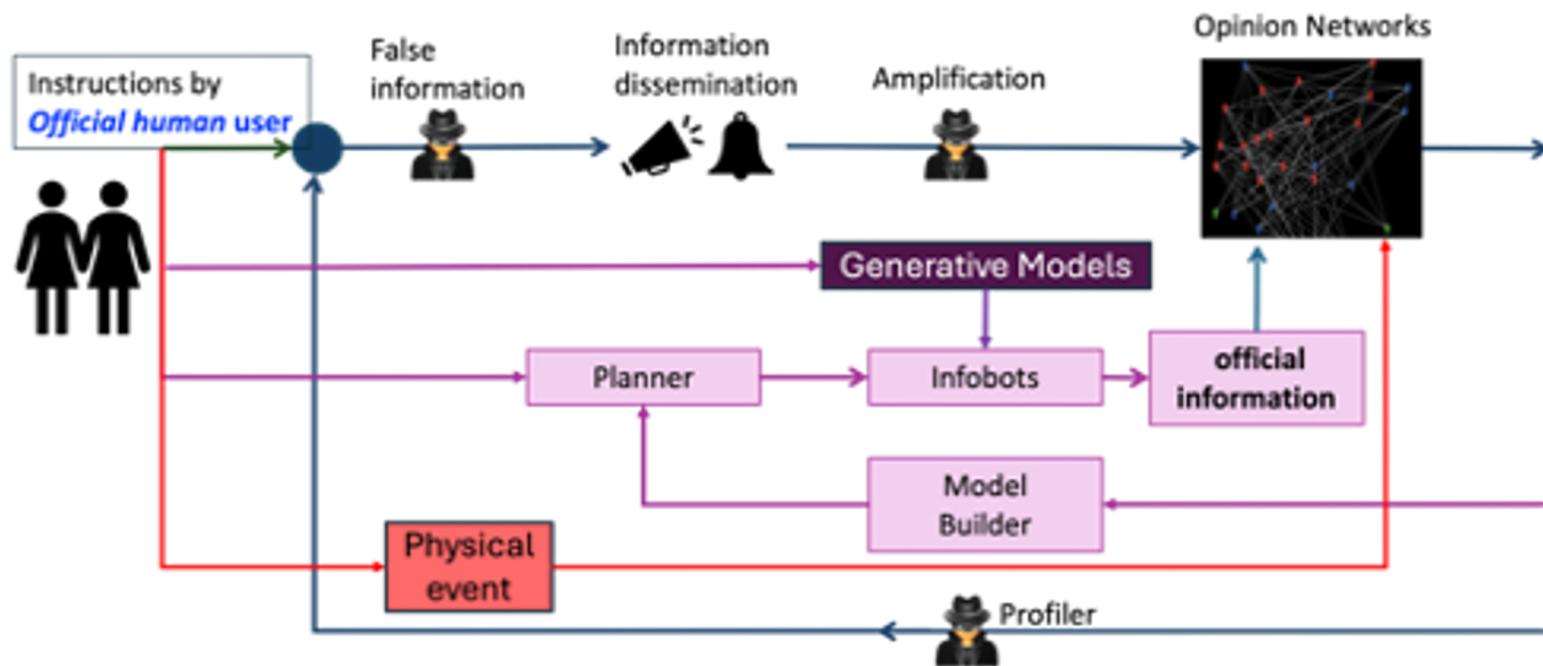
Data Sample

- **Query categories based on the stakeholders involved:** voters, candidates, and Election administration / Journalists.
 - **Sub-categories:** ‘What’, ‘How’, ‘When’, and ‘Who’ questions.

Stakeholders	Sub-category	Question - general	Parameter(s)	Region(s) Applicable	Reference to country-specific tab
Voter	What	What counts as a "fair" election?	None	Any	IDN-Q6
	How	How would the voters of advanced age vote as actively as the young?	None	Any	EST-Q36
	When	When can the <region> begin to count absentee & mail-in ballots?	<region>	Any	USA-Q35
	Who	Who can run for <Head-of-state> of <region>?	<Head-of-state>, <region>	<Head-of-state>, <region> = King Queen, Canada, President, USA, India, South Africa, Indonesia	IRE-Q2
Candidate	What	Do <poll officers> have any guidelines about conditions when a candidate can request a recount in a close result?	<poll officers>	Any	UK-Q3
	How	How is it ensured that each voter votes themselves (independently) (in <region>)?	<region>	Any	EST-Q20, EST-Q43
Election Administration / Journalist questions	What	What goes into the cost of an election in <region>?	<region>	Any	CA-Q7
	How	How does <region> handle voter abstention?	<region>	Any	IDN-Q3
	Who	Who are Millennial voters?	None	Any	IDN-Q1

Table 2: Sample generalized queries extracted from different sources with corresponding stakeholders, sub-categories, parameters, regions applicable, and reference to the country-specific tab. The country codes for each country are shown in Table 1.

A Simulator for modeling and controlling spread of information in the presence of other data sources and actors



Part 2: Distribution of Data

Simulation Tool

Expressive and Flexible Simulation of Information Spread Strategies in Social Networks Using Planning

Bharath Mupparani, Vignesh Narayanan, Biplav Srivastava,
Michael N. Huhns

Proceedings of the Thirty-Eighth Annual Conference of Association for Advancement of Artificial Intelligence - Demonstrations Track (AAAI-24 Demo), 2024

<https://www.youtube.com/watch?v=Sn0FFTsTqXM>



Interaction With Official Data

On safe and usable chatbots for promoting voter participation
Bharath Muppasani, Vishal Pallagani, Kausik Lakkaraju, Shuge Lei, Biplav Srivastava, Brett Robertson, Andrea Hickerson, Vignesh Narayanan

Solution Value: Role of Chatbots

- **Hypothesis:** Using technology for voter engagement, especially among seniors and youths, will increase participation [2].
- CDC approach
 - **Collate:** Collate good/ official information
 - **Distribute:** Safely distribute in opinion and social networks
 - **Comprehension:** Make content easy to understand
- Chatbots can proactively and interactively deliver useful information for voting compared with websites.
 - A safe chatbot architecture is needed that can be verified (like voting systems)
 - The effectiveness of using chatbot to improve voting turnout has not been evaluated.

Promise and Problem with Chatbots

Promise

- A. Allows people to interact with data naturally
- B. Can adapt content to user's cognitive and learning ability (language/ dialect, style, mode – text / graphics/ audio, font, accent)
- C. Make content easily available when needed

Demonstration of SafeChat-Elections-SC:
<http://casy.cse.sc.edu/Election-Chatbot-SC-main/>

Problems

- A. Difficult to link output to input data sources (**provenance**)
- B. Hard to control output (**fragile**)
 - a. May contain abusive language.
 - b. May exhibit bias.
- C. Requires deep AI skills to build (**costly**)
- D. Long development time to build (**costly**)
- E. Lack of testing guidance (**unreliable**)

1. Biplav Srivastava, Kausik Lakkaraju, Tarmo Koppel, Vignesh Narayanan, Ashish Kundu, Sachindra Joshi, Evaluating Chatbots to Promote Users' Trust -- Practices and Open Problems, On Arxiv at: <https://arxiv.org/abs/2309.05680>, 2023 [Chatbot evaluation, AI Trust]
2. Biplav Srivastava, Did Chatbots Miss Their 'Apollo Moment'? A Survey of the Potential, Gaps and Lessons from Using Collaboration Assistants During COVID-19, To Appear Cell PATTERNs, Aug 2021. (Preprint on Arxiv: <https://arxiv.org/abs/2103.05561>, March 2021) [Chatbots, COVID]

SafeChat / Novelty Statement

The unique aspects of SafeChat that can address the various problems with the chatbots are: (**how**)

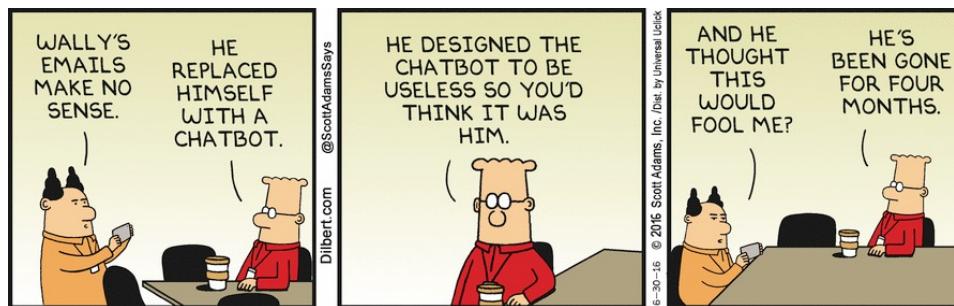
1. A safe design where only responses that are grounded and traceable to an allowed source (e.g., official question/answer) will be answered via system's self-awareness (metacognition) [Problem A]
2. A do-not-respond strategy that can deflect certain user questions which may be harmful if answered. [Problem B]
3. Trust ratings provided to communicate the chatbot's expected behavior [Problem B]
 1. abusive language and
 2. bias exhibited.
4. A CSV-driven chatbot building workflow that does not require deep AI expertise, making it accessible to developers with varying levels of AI knowledge and experience. [Problems C, D]
5. A low-programming design pattern based on the open-source Rasa platform to generate chatbots quickly for any setting (e.g., domain, language, localization) [Problems C, D]. The backend can be extended with CSV-driven web integration.
6. Support for control and treatment group formation and analysis of results, to support RCT testing [Problem E]

Problems (What)

- A. Difficult to link output to input data sources (**provenance**)
- B. Hard to control output (**fragile**)
 - a. May contain abusive language.
 - b. May exhibit bias.
- C. Requires deep AI skills to build (**costly**)
- D. Long development time to build (**costly**)
- E. Lack of testing guidance (**unreliable**)

Our approach works with any rule-based, also called slot-based or event-condition-action based chatbot building platform, like Rasa or DialogFlow. They in turn may use learning for their functionality like intent specification.

Solving Real Business Problems



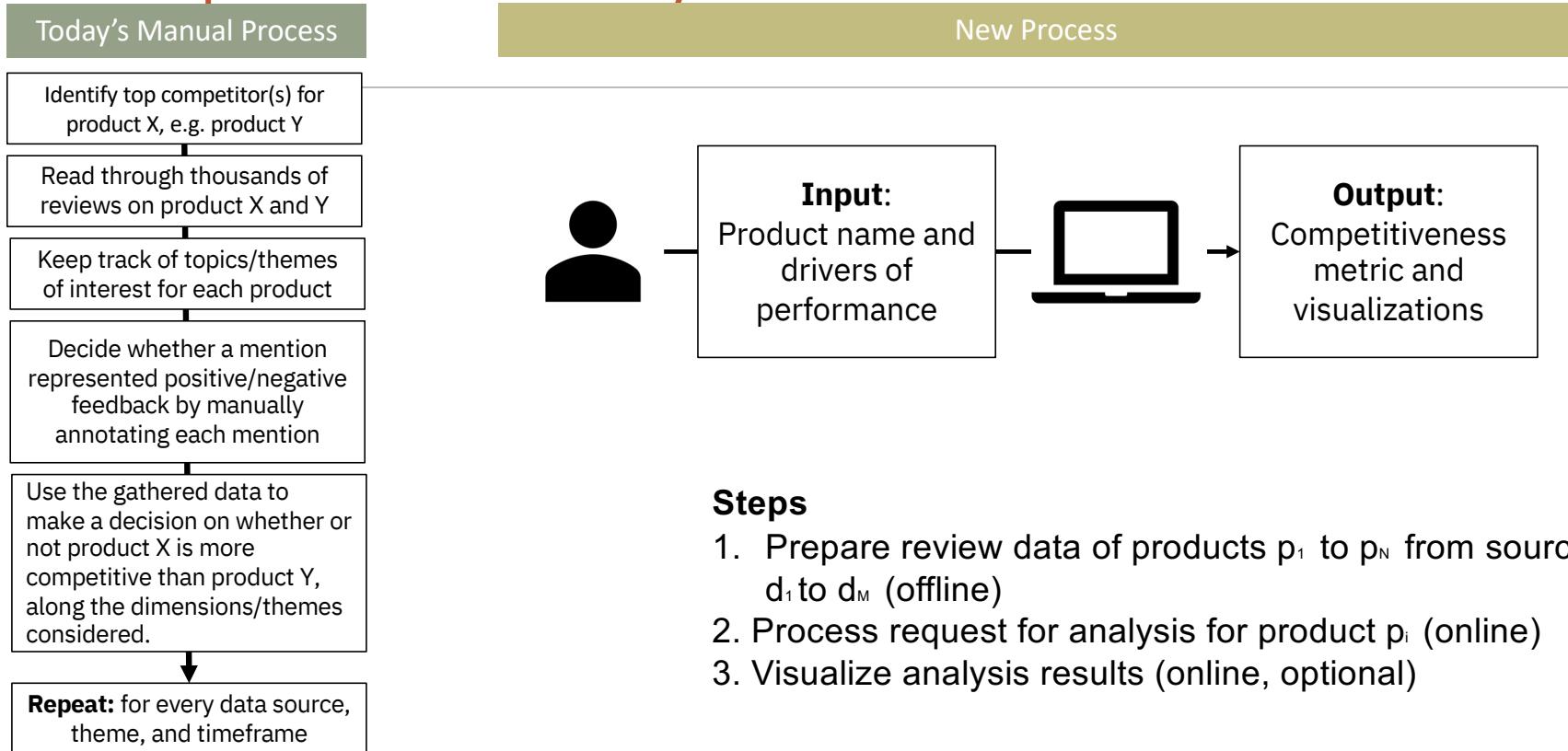
Credit: Dilbert – June 30, 2016

Case Study of Sentiment in Business

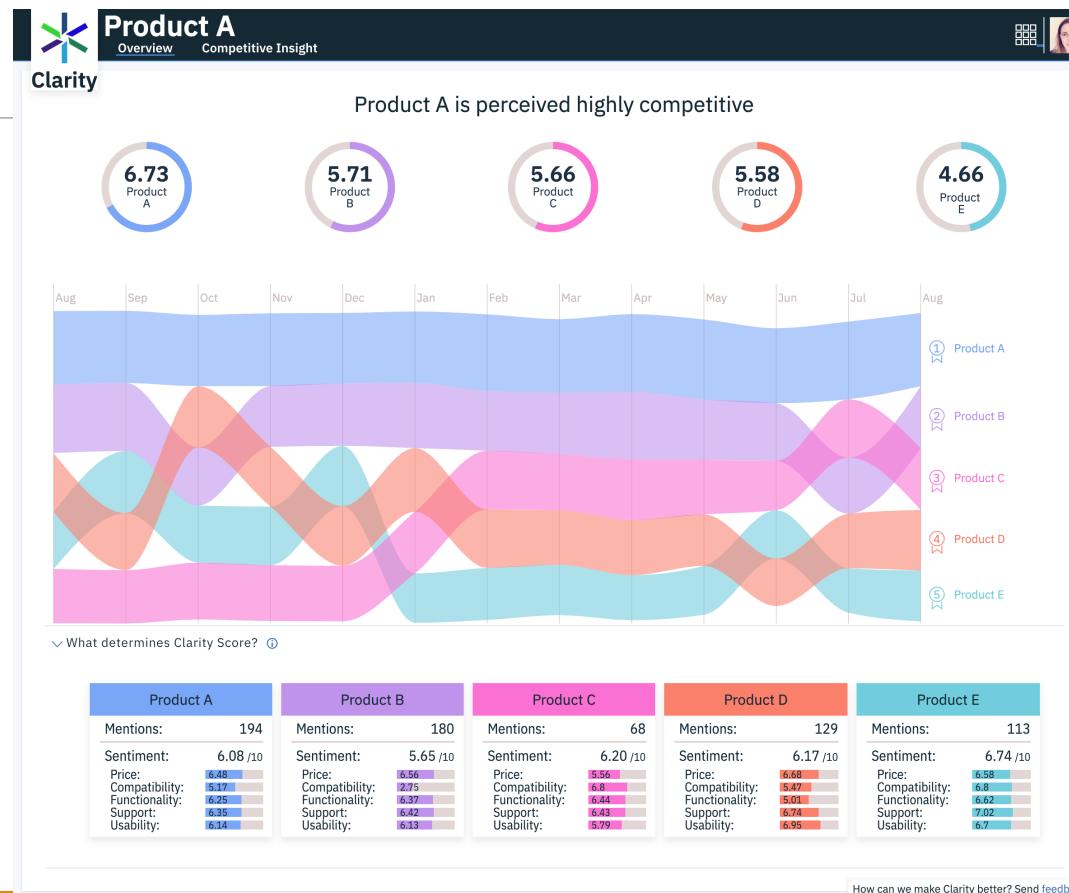
Clarity: Data-Driven Competitive Analysis

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



Illustrative Output



Clarity Score and Trends

Methodology

Data ingestion

Feature Extraction

Sentiment Analysis

Aspect mining

Aggregation

Sentiment Analysis

Keywords



Calculate keyword level sentiment

Separate out the negative and positive parts of review

Run summarization algorithm

Review Summary

"The DB2 LUW performance on analytical workloads was exceptional and the speed improvements using the new queries for both stored procedures and views were phenomenal. The configuration is flexible in allowing for both column and row oriented tables comprising both column and row store tables. DB2 is an excellent RDBMS which supports all the features of column store tables. Very few problems when we migrated to the new version and excellent customer support. The DB2 support team has always been supportive with getting the right resources to insure our company's DB2 databases avoid any performance issues."

"Issues were resolved quickly. Excellent documentation, decent support but since its IBM it's not surprising that quality is high. The DB2 is fast, stable but requires a lot of vendor intervention to troubleshoot when issues arise. It's a great product and I would say all that great for pure up-and-running functionality. It's probably the best RDBMS. Some people like Oracle better but personally I really like how issue resolution was organized and quick. Not as popular as Oracle or MySQL but it's still a good choice because not as popular third party tools (performance measurement tools / etc) are available. The Oracle consumes more RAM than DB2 does. Like I said."

Course Project

Discussion: Projects

- New: two projects
 - Project 1: model assignment
 - Project 2: single problem/ llm based solving / fine-tuning/ presenting result

Lecture 28: 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

Thanks – Optional Next Class ?
