Sandeep Sandha, Dr. Sukanya Randhawa, Dr. Biplav Srivastava
IBM Research – India

Blue Water –

Putting Water Quality Data in India to Productive Use by Integrating Historical and Real-time Sensing Data

Talk at CSE Workshop on **Mainstreaming Citywide Sanitation** at New Delhi, India
4-5 April, 2016


Acknowledgements: Our colleagues at IBM Research and collaborators at various agencies.

Acknowledgements / Partners

S. No.	Area	People, Organization
1	Core Technology	Supratik Guha, Theodore G van Kessel, Hendrik Hamann, Bharat Kumar, Jaikrishnan Hari, Sachin Gupta, Karthik Visweswariah, Anupam Saronwala, IBM Research Worldwide
2	Hindon exploration, Agriculture Use-case	2030 Water Group and their partners; Dr. V. Rajagopalan
3	Yamuna exploration	Delhi Jal Board
4	Ganga exploration, Khumbh use-case	Prof. V. Raychoudhary and students, IIT Roorkee
5	Analytics	Ben Ford, Prof. M. Tambe and colleagues, University of Southern California, USA

What Our Team Can (And Cannot) Do

- We are not water quality experts

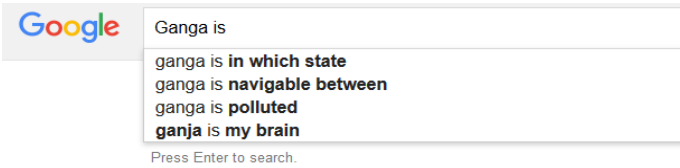
- Expertise in helping make decisions via analytics and machine learning
- Expertise in cloud based data management and apps (web, mobile)
- WW expertise in physics and chemistry—measurement technique development
- Expertise in designing robust sensor network systems

- IBM Research active in India since 1998, winner of a National Award for developing country focused innovation
  - Have collaborated with local faculty via Faculty Awards, PhD Fellowships and internships
  - Taken many “made-in-India” innovations to the world

Main Messages

- We want common citizens to make better decisions around water
- We are building tools that others can use: GangaWatch, Neer Bandhu powered by BlueWater Architecture
- We are measuring water quality with a novel, multi-sensor approach combining traditional lab tests, real-time sensors and mobile apps
  - We use a novel real-time sensing approach of using mobile platform to collect data at fine spatial and temporal granularity
  - We have done actual measurements on Yamuna, Hindon and Ganga
- We are looking for partners and business models to help scale and make real impact in a timely manner

As of April 3, 2016



An indication of queries about Ganga

An indication of possible approaches

Better Information Flow is Critical for Better Water Flow

The nature of water management must rapidly evolve

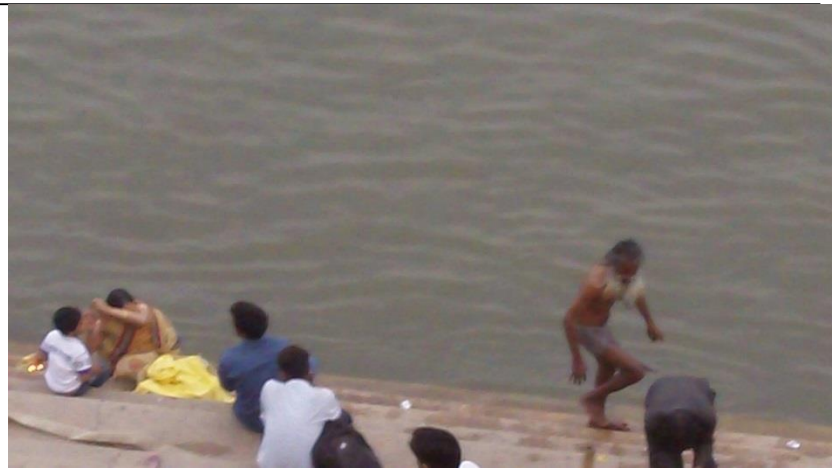
From	*To*
Manual Data Collection	Automated Sensing
Managing in Isolation	Managing Collaboratively
Intermittent Measurement	Real-Time Measurement
Multiple Data Sets	Data Integration
"Guesstimation" Tools	Modeled Decision Support
Commodity Pricing	Value Pricing
Tactical Problem Solving	Strategic Risk Management

"One barrier to better management of water resources is simply lack of data — where the water is, where it's going, how much is being used and for what purposes, how much might be saved by doing things differently. In this way, the water problem is largely an information problem. The information we can assemble has a huge bearing on how we cope with a world at peak water."

Source: **Wired Magazine, "Peak Water: Aquifers and Rivers Are Running Dry. How Three Regions Are Coping", Matthew Power, April 21st, 2008**

[India] Ganga – Local Ground Situation @ Varanasi (Assi/ Tulsi Ghats) + Patna

Assi Ghat post recent cleanup



Bathing on Tulsi Ghat

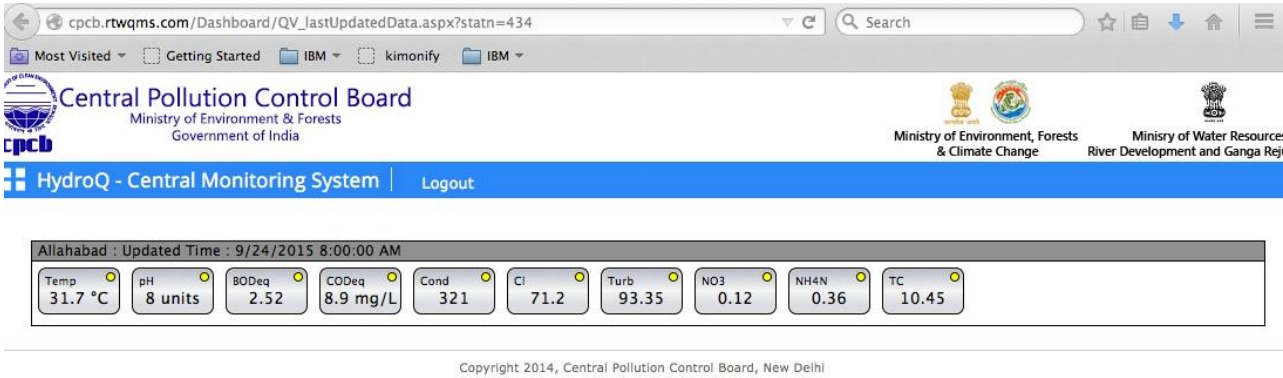
A nullah draining into Ganga

A manual powered boat

Photos at Gandhi Ghat, Patna on 18 March 2016 during 1700-1800 Hrs
Photos of/ at Assi/ Tulsi Ghat, Varanasi on 25 March 2016 during 1700-1800 Hrs

Common scene around Indian water bodies

Decision Example –River Water Pollution



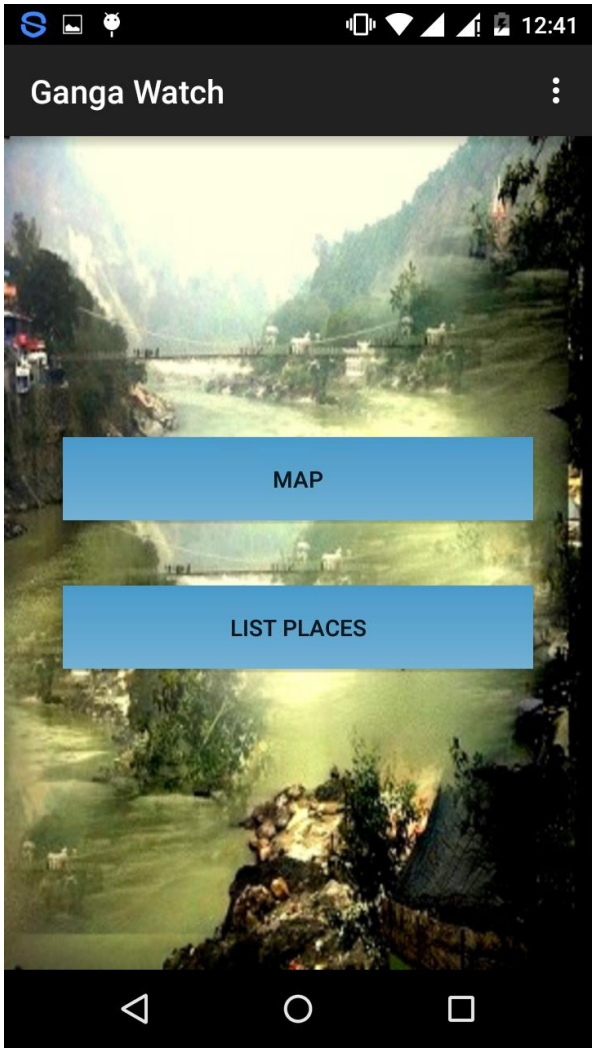
- Value – To individuals, businesses, government institutions
  - Example – Can I take a bath? Will it cause me dysentery?
  - Example – How should govt spend money on sewage treatment for maximum disease reduction?

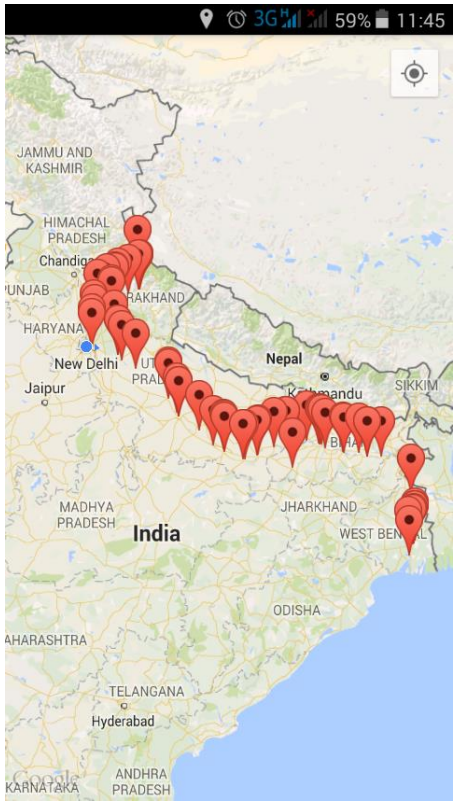
- Data – Quantitative as well as qualitative
  - Dissolved oxygen,
  - pH,
  - … 30+ measurable quantities of interest

- Access –
  - Today, little, and that too in water technical jargon
  - In pdf documents, website

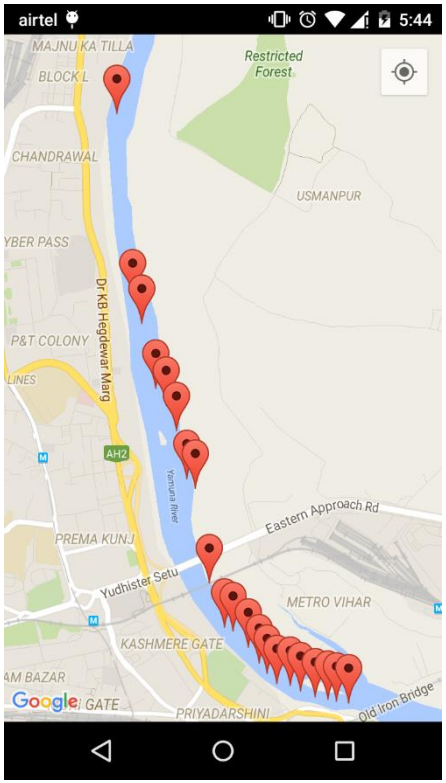
Key Idea: Can we make insights available when needed and help people make better decisions?

Demo: GangaWatch





Data Covering Ganga Basin



Fine-grained
Geo-tagged
Data from a
Real Time
Run on Yamuna

Art of Possible

Tannery Example: Kanpur, India

Background of Leather Tanning Problem

- > 700 tanneries in Kanpur
  - Employing > 100,000 people
  - Bringing > USD 1B revenue
- Discharge water after leather processing to river or Sewage treatment plants (STPs)
  - Requirement
    - Must have their own treatment facility
    - Or, have at least chrome recovery unit
  - But don't implement due to costs which is a burden to main operations
    - Installation
    - Operations : electricity, manpower, technology upgrade, …
  - State pollution board is supposed to do inspections to enforce but doesn't perform effectively
- Government's STPs do not process chrome, the main pollutant
- Knee-jerk reaction: 98 tanneries banned in Feb 2016 by National Green Tribunal; more threatened

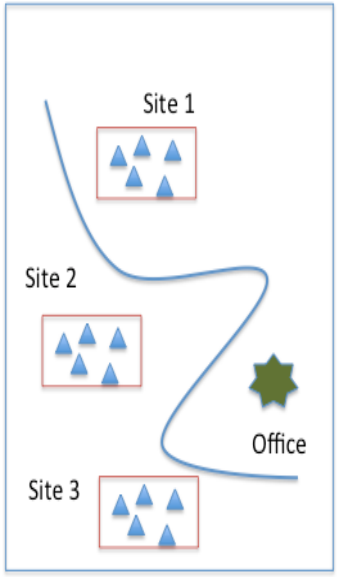
NECTAR: Nirikshana for Enforcing Compliance for Toxic wastewater Abatement and Reduction

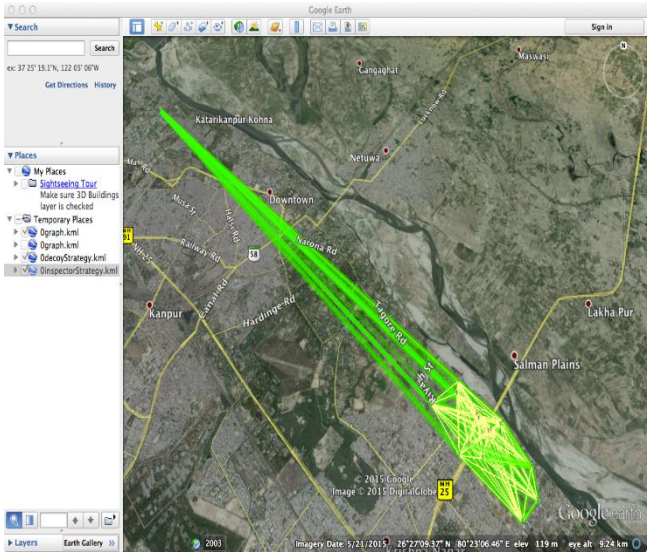
Protecting the Nectar of the Ganga River through Game-Theoretic Factory Inspections, B. Ford, A. Yadav, A. Singh, M. Brown, A. Sinha, B. Srivastava, C. Kiekintveld, M. Tambe
14th International Conference on Practical Applications of Agents and Multi-Agent Systems, Sevilla, Spain, June 1-3, 2016.

Setting

- Attackers
  - M sites with N factory units each
  - When inspection at a site happens, all units know
- Defenders
  - Inspectors base office is fixed
  - Inspection team consists of
    - Environment Inspectors
    - Security personnel
    - Transport provider / drivers
  - Inspection team starts and ends at their office
  - Security and transport can vary daily
- Objective
  - Create daily inspection plan which minimizes violation over a time period

Joint work with USC, USA





Main Results

Proposed method achieves compliance faster than existing methods and scale fast.

- Used actual location of 50 tanneries in Kanpur
- With a fixed fine (one fine amount for all sites) and decoys, compliance from all sites (simultaneously) will be achieved faster than existing methods.
- With a variable fine (based on number of factories at the site), proposed method performs better than existing methods, with or without decoys.
- Can improve further with monitored pollution data

"Very promising approach. Use of decoys and data-driven random were not known in the inspection community where it was known that random could help. Surprise elements of decoys and variable fines provide new factors for compliance. The data from drone monitoring can help improve the plans significantly as future work."

Dr. Venkatraman Rajagopalan, IAS
Ex-Secretary, Ministry of Environment, Forests and Climate Change, and
Ex-Chairman, Central Pollution Control Board, India

Outline

- Background – Challenges, Trends, Motivation
- Illustrative Case Study – Tanneries at Kanpur
- Pollution Sensing, Analytics Platform
  - What's new
  - Yamuna @Delhi [Dec 2015]
  - Ganga @Haridwar [Mar 2016 -]
- Discussion

Water Pollution Sensing

- Method 1: Sample collection and lab-testing
  - Accurate when done well
  - Time-consuming, costly and for a few places at a time
  - Only quantitative
  - **Science**: lab tests, sample collection
- Method 2: Real-time sensing
  - Timely, inexpensive
  - Some parameters are NOT feasible
  - Only quantitative
  - **Science**: how to deploy sensors and analyze data
- Method 3: Crowd-sourcing
  - Timely, inexpensive
  - Only qualitative assessment
  - Practical for India with people and mobiles
  - **Science**: Combining qualitative and quantitative data

Quantitative Sensing Scope

Dimension	{Yamuna | Hindon| Ganga}
Scenario focus	General, Agriculture
Real-time measurement	DO, pH, conductivity, turbidity
Lab / samples	BOD, COD, FCC
Sensing	COTS sensors, Machine learning, In-lab test
Data ingestion	Bluemix cloud, Cloudant database

Sensor Measures

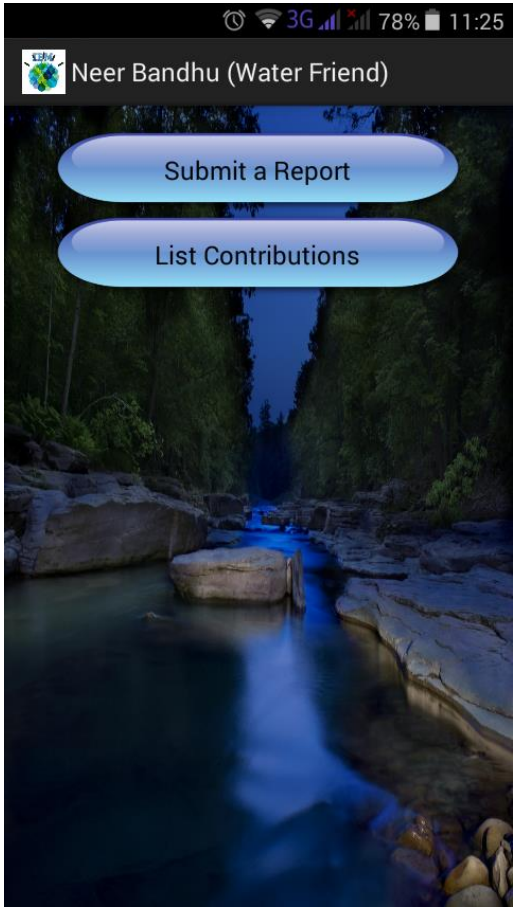
Primary

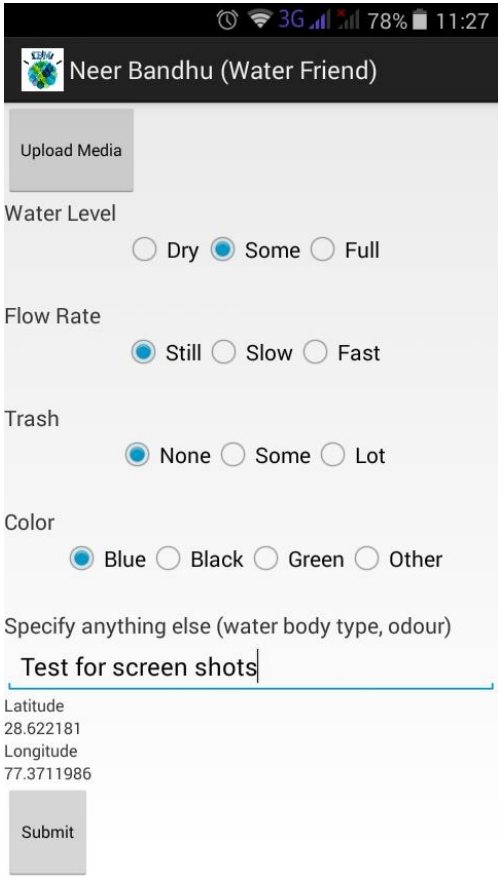
- Temp
- ORP
- D.O
- EC
- Turbidity
- Pressure
- Nitrate
- GPS Lat
- GPS Long

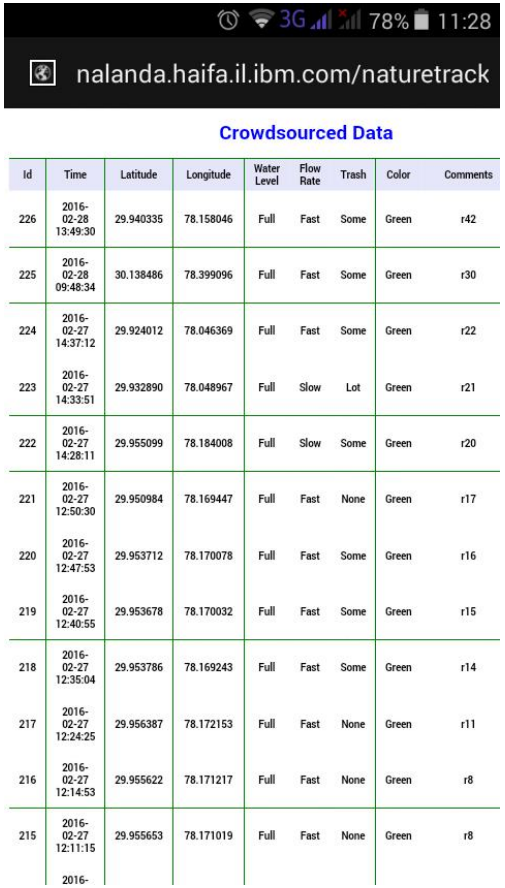
Secondary

- Resistivity
- TDS
- Salinity
- SeaWater Sigma

Water Qualitative Data Via Crowdsourcing – NeerBandhu App







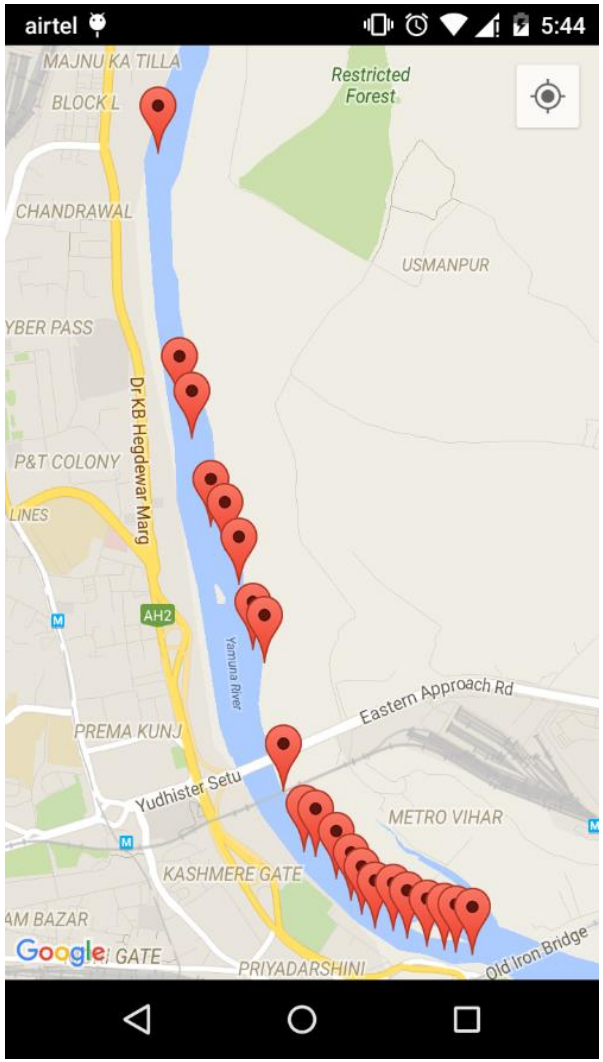
Id	Time	Latitude	Longitude	Water Level	Flow Rate	Trash	Color	Comments
226	2016-02-28 13:49:30	29.940335	78.158046	Full	Fast	Some	Green	r42
225	2016-02-28 09:48:34	30.138486	78.399096	Full	Fast	Some	Green	r30
224	2016-02-27 14:37:12	29.924012	78.046369	Full	Fast	Some	Green	r22
223	2016-02-27 14:33:51	29.932890	78.048967	Full	Slow	Lot	Green	r21
222	2016-02-27 14:28:11	29.955099	78.184008	Full	Slow	Some	Green	r20
221	2016-02-27 12:50:30	29.950984	78.169447	Full	Fast	None	Green	r17
220	2016-02-27 12:47:53	29.953712	78.170078	Full	Fast	Some	Green	r16
219	2016-02-27 12:40:55	29.953678	78.170032	Full	Fast	Some	Green	r15
218	2016-02-27 12:35:04	29.953786	78.169243	Full	Fast	Some	Green	r14
217	2016-02-27 12:24:25	29.956387	78.172153	Full	Fast	None	Green	r11
216	2016-02-27 12:14:53	29.955622	78.171217	Full	Fast	None	Green	r8
215	2016-02-27 12:11:15	29.955653	78.171019	Full	Fast	None	Green	r8
	2016-							

Data at http://nalanda.haifa.il.ibm.com/naturetrack/visualization.php

Gaps Filled by Our Approach

- High spatial and temporal resolution (real-time)
  - Current data are at low resolution of few places and limited time points; limits usage in applications
  - Use floating platform and real-time sensor to collect GPS-enabled data
  - Use location to re-create water body condition
- New source of data (qualitative; crowd-sourcing)
- Fusion of historic and new real-time data on single platform with safety levels and purpose
- Future: contextualize quantitative data with qualitative inputs for data validation and stakeholders buy-in

Sensing on Yamuna



Real-Time Sensor Deployment

Day 1 - multiple anchoring approaches for real-time sensor on another day (16 Dec) in 2-3 km stretch

	16-Dec-15						
	Location Name	**Description**	**Sample - collected**	**Sample - testing**	**Sensor @site**	**Realtime (Stretch)**	**Neer Bandhu**
1	Point 1 [A]	Nigambodh, in water	Y	Y (ph, DO, Temp, Turb, Cond, **BOD, FCC**)	Y		Y
2	Point 2 [B]				Y		Y
3	Point 3 [C]	ITO bridge	Y		Y		Y
4	Point 4 [D]				Y		Y
5	Pointe 5 [E]		Y	Y (ph, DO, Temp, Turb, Cond)	Y		
6	Point 6	Moving (7-8 Kmph)				Y	
7	Point 7	Moving (10 Kmph)				Y	Y
8	Point 8	Drain	Y	Y (ph, DO, Temp, Turb, Cond)		Y	Y
9	Point 9	With Ted buoy				Y	

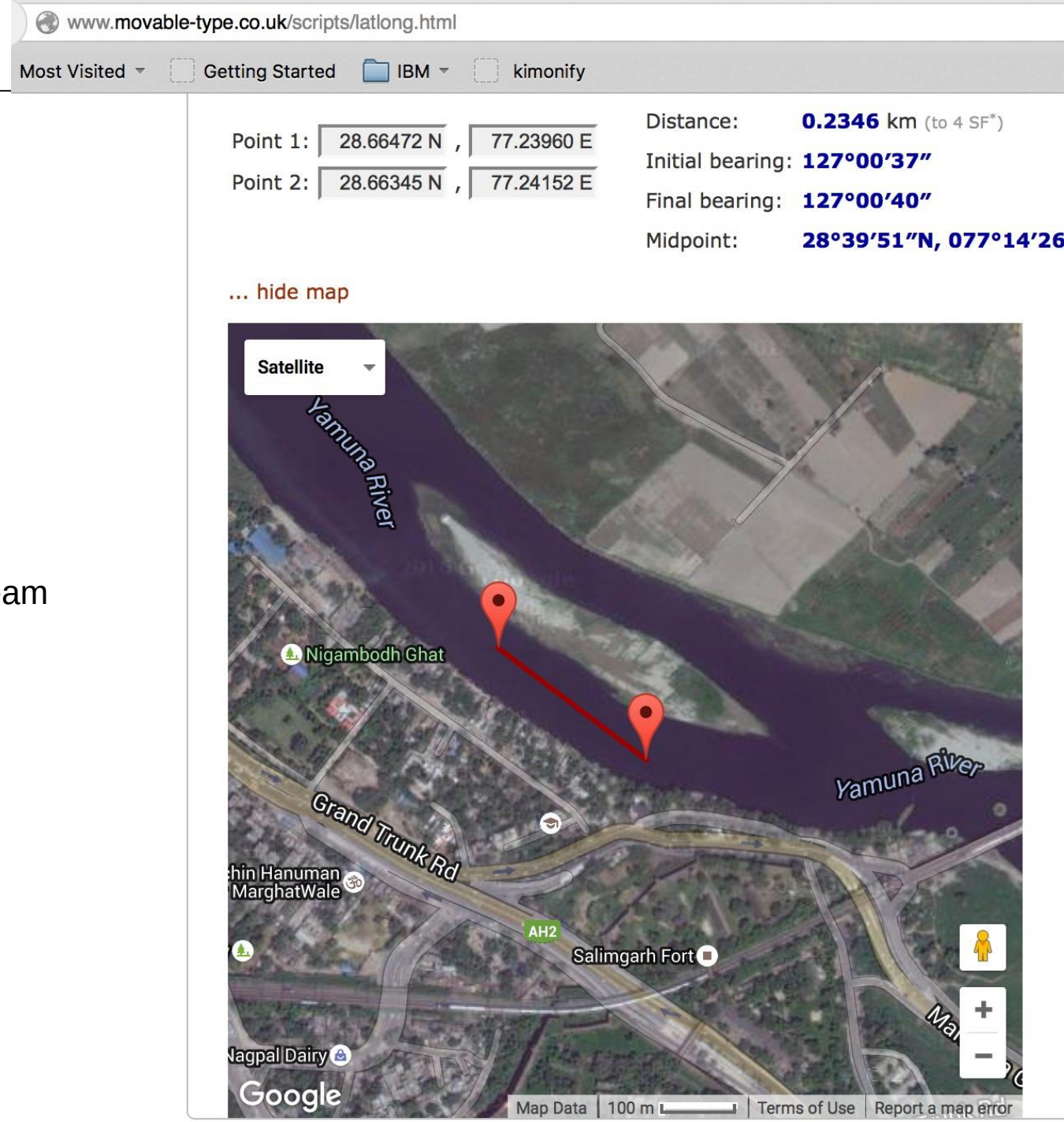
Dec 16

Example Run

16/12/15 13:46:50

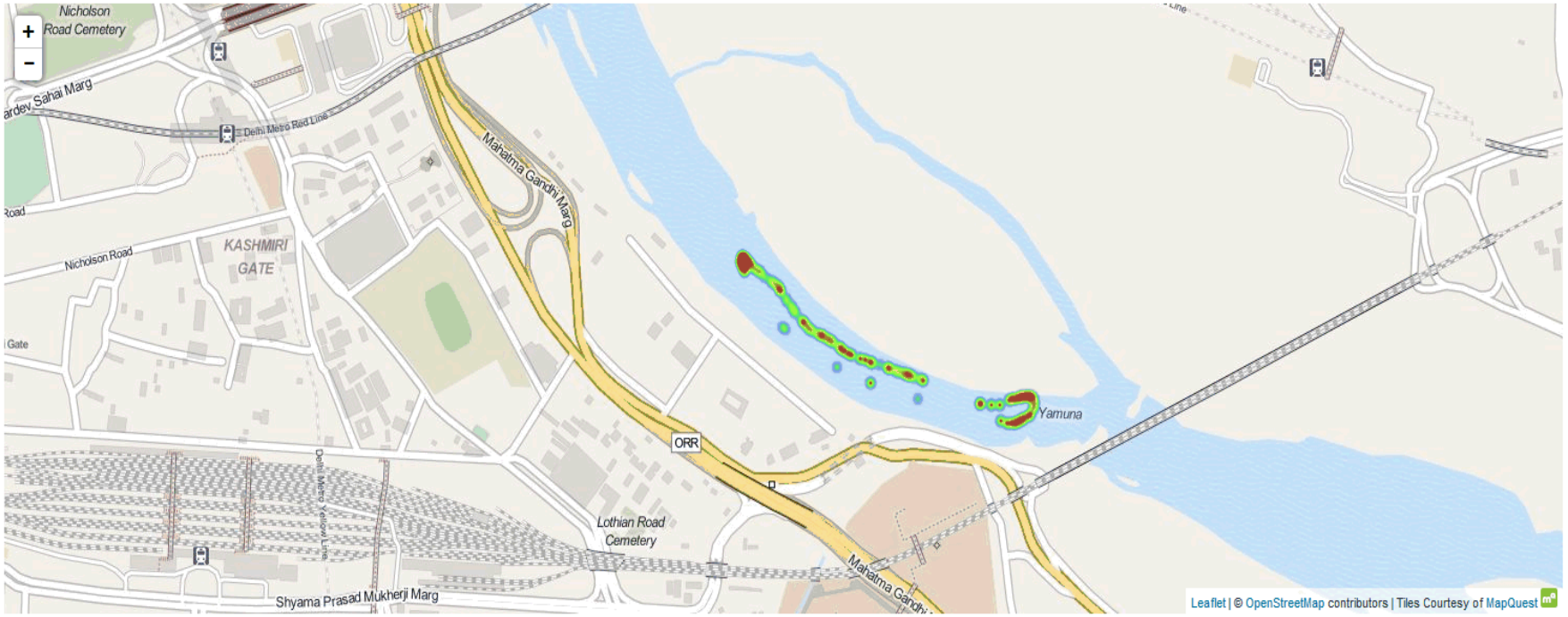
- ~12 minute downstream travel
- 765 data points

16/12/15 13:59:34



Turbidity in Yamuna – measured on 16th Dec, 2015
Data min: 56.7
Data max: 138
Gradient: *Default*

Express



Welcome to Express

Day 2 - Covered ~7-8 km one-way on one of the days(18 Dec) roughly covering 33 % of the navigable stretch of Yamuna in Delhi (22 km one-way).

	18-Dec-15						
	Location Name	**Description**	**Sample - collected**	**Sample - testing**	**Sensor @site**	**Realtime (Stretch)**	**Neer Bandhu**
1	Point 21 [AA]	Nigambodh, in water	Y	Y (ph, DO, Temp, Turb, Cond, **BOD, FCC**)		Y	Y
2	Point 22 [AB]	Past rope (ISBT)				Y	Y
3	Point 23 [AC]	2nd rope				Y	Y
4	Point 24 [AD]	Drain	Y	Y (ph, DO, Temp, Turb, Cond)	Y	Y	Y
5	Pointe 25 [AE]	Drain				Y	Y
6	Point26 [AF]	Drain, gurudwara				Y	Y
7	Point 27 [AG]	Wazirabad bridge	Y	Y (ph, DO, Temp, Turb, Cond)	Y	Y	Y
8	Point 28 [AH]	Majnu ka tila, greenery				Y	Y
9	Point 29 [AI]	1st rope, ISBT				Y	

Dec 18

Example Run

2015/12/18,12:51:37

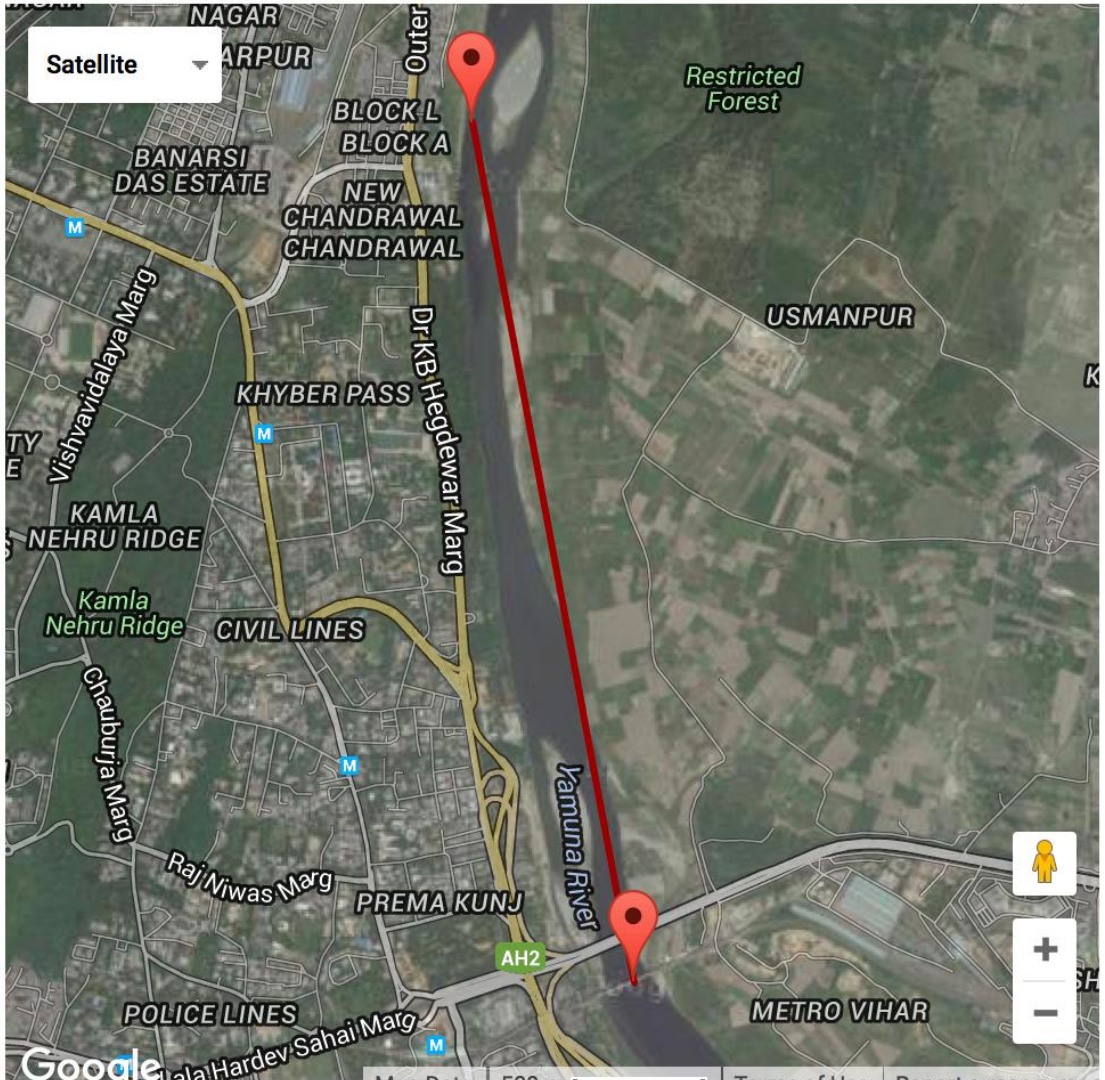
- ~38 minute upstream travel
- 2273 data points

2015/12/18,12:13:45

Point 1:	28.66983N	,	77.23566E
Point 2:	28.69821N	,	77.22956E

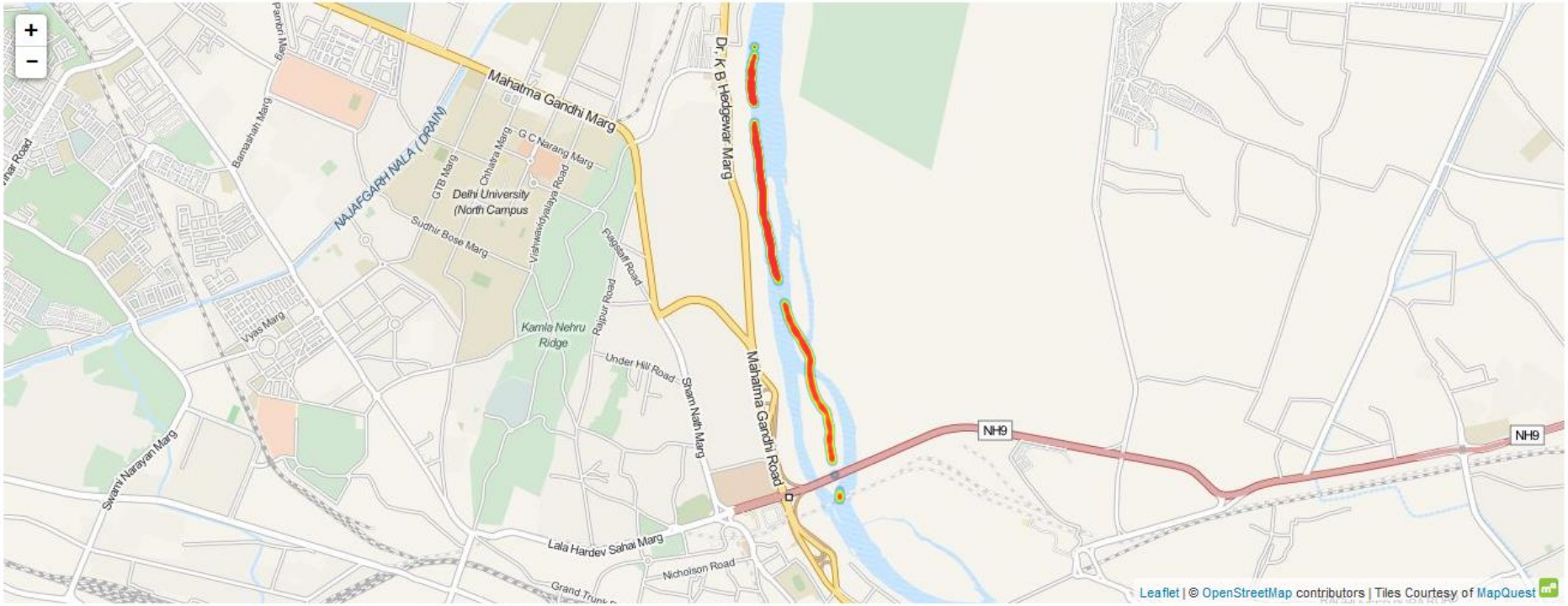
Distance: **3.211** km (to 4 SF*)
Initial bearing: **349°19′23″**
Final bearing: **349°19′12″**
Midpoint: **28°41′02″N, 077°13**

... hide map



Turbidity in Yamuna – measured on 18th Dec, 2015
Data min: 37.5
Data max: 144.4
Gradient: *Default*

Express



Welcome to Express

Lab Samples and Traditional Testing

Newcon Consultants & Laboratories

An ISO 9001 : 2008, ISO 14001 : 2004,OHSAS 18001 : 2007 Certified Laboratory NABL ISO/IEC 17025 : 2005 (Chemical Testing, Cert. No. T-1761, Biological Testing, Cert. No. T-3285), Accreditated Laboratory, Recognised with MOEF & U.P. Pollution Control Board

Website : www.newconlab.com

ISO 9001/14001/OHSAS 18001 Reg. No. R191/7381

TEST CERTIFICATE

WASTE WATER SAMPLE ANALYSIS REPORT

Page 1 Of 1

TEST REPORT NO : NCL/DBSND/E-158/12/2015	DATE OF REPORT : 21/12/2015
Name And Address Of Customer	DR.BIPLAV SRIVASTAVA IBM RESEARCH INDIA , 4, BLOCK-C ,ISID CAMPUS, , ,VASANT KUNJ,NEW DELHI, INDIA

SAMPLING DETAILS

Analysis Start Date	17/12/2015	Analysis End Date	21-12-2015
Date of Sampling	16/12/2015	Sampling ID No.	316/12
Time of Sampling	12:05		
Sampling Done By	NCL		
Sampling Location	RIVER YAMUNA		
Sampling Description	SAMPLE MARKED AS YAM-1		
Sampling Protocol	IS:3025(Part-I)	Sampling Quantity	ONE Lt
Packing Condition	Sealed	Packed In	PVC BOTTLE

TEST RESULT

S.No.	Parameter	Unit	Protocol	Result
1	pH	--	APHA-4500(H+B)	7.28
2	Bio Chemical Oxygen Demand (3 days at 27°C)	mg/L	APHA-5210 (B)	46
3	Dissolved Oxygen (D.O.)	mg/L	APHA-4500- O (D)	<1.0
4	Turbidity	NTU	APHA-2130-B	34
5	Conductivity	µs/cm	APHA-2510	1393
6	Temperature	° C	APHA-2550 (B)	18
7	Faecal coliform	No./100 ml	APHA-9221	430

**** End Of Report****

FOR NEWCON CONSULTANTS & LABORATORIES

TECHNICAL MANAGER

CHECKED BY

Dr. Ramesh C. Tripathi
M.Sc. Ph.D. Environmental Science

AUTHORIZED SIGNATORY

Format no NCL/QSP-28/TC-WWT/FMT-04 Rev.No.1 Date 18.07.2011

NOTE : 1. The Results reported above pertains to the Tested parameters only. Endorsement of the same is neither inferred nor implied. 2. All disputes subject to **GHAZIABAD JURISDICTION.** 3. The Report shall not be reproduced except in full without the permission of **CHIEF ANALYST.** 4. Our liability is limited to invoiced value only.

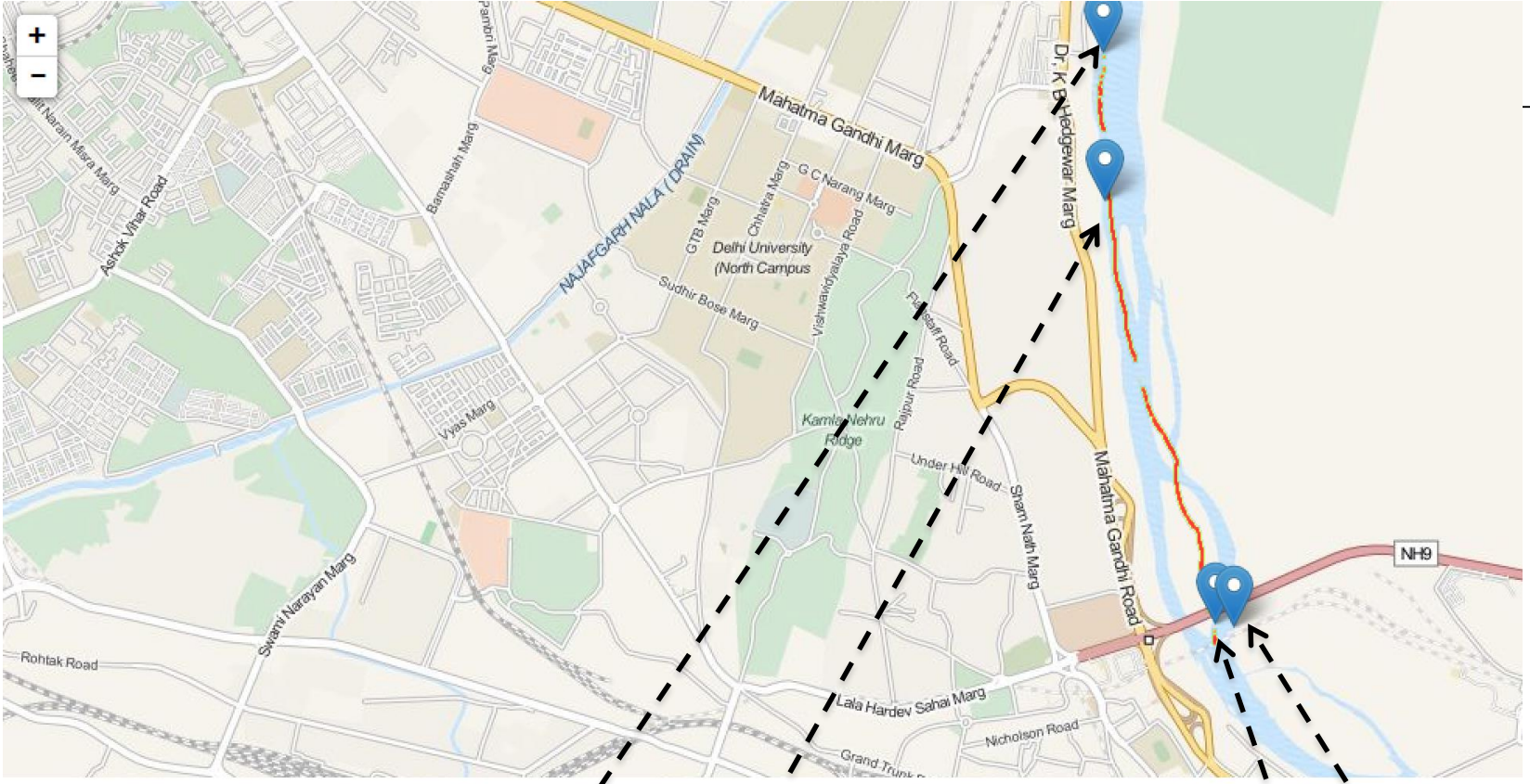
Laboratory : 8th K.M. Stone, Delhi Meerut Road, Morta (Opp. Manan Dham Mandir) GHAZIABAD - 201 003 (U.P.) Telefax : (0120) 2675225, Mobile : 9810430345

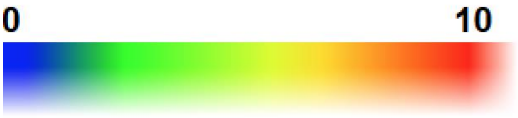
E-mail : info@newconlab.com, newconlab@gmail.com

Change in parameters measured for two different days

More water released into river

		16/12/2016	18/12/2016
Sensor	Temp(°C)	15.93	15.34
	pH	7.82	7.81
	ORP(mV)	-182	-86.4
	D.O(mg/L)	3.76	3.53
	EC (µS/cm)	1604	1279
	Turbidity (F.N.U)	84.25	66.9
Lab	BOD (mg/L)	46	28.2
	Fecal Coliform (No./100 mL)	430	210





Correlating RT Sensor and Crowd Data to Get Verifiable Data!

nalanda.haifa.il.ibm.com/naturetrack/visualization.php

Search

Most Visited | Getting Started | IBM | kimonify

203	2015-12-18 09:20:3	28.698308	77.229535	Some	Slow	Some	Black	yamuna, point 27, wazirabad br	photo
202	2015-12-18 09:15:00	28.670732	77.235700	Some	Fast	None	Black	yamuna, point 26, drain+gurudw	photo
201	2015-12-18 09:13:02	28.670732	77.235700	Some	Fast	Some	Black	yamuna, point 25, anoth drain	photo
200	2015-12-18 09:02:52	28.670732	77.235700	Some	Fast	Some	Black	yamuna, point 24,drain	photo
199	2015-12-18 08:53:54	28.670732	77.235700	Some	Slow	Some	Black	yamuna, point 23 ,2 rope	photo
198	2015-12-18 08:41:39	28.670732	77.235700	Some	Slow	Some	Black	yamuna, point 22, past rope,	photo
197	2015-12-18 08:19:59	28.670546	77.236743	Some	Slow	Some	Black	yamuna, point 21, nigambodh in	photo
196	2015-12-18 07:12:19	28.625278	77.308005	Full	Slow	Lot	Black	yamuna, nigambodh, near bank	photo

Sensing on Ganga

Joint work with Prof. Vaskar Raychoudhury and students at IIT Roorkee

Use-Case: Understand Impact of a Large-Scale Religious cum Tourism Event

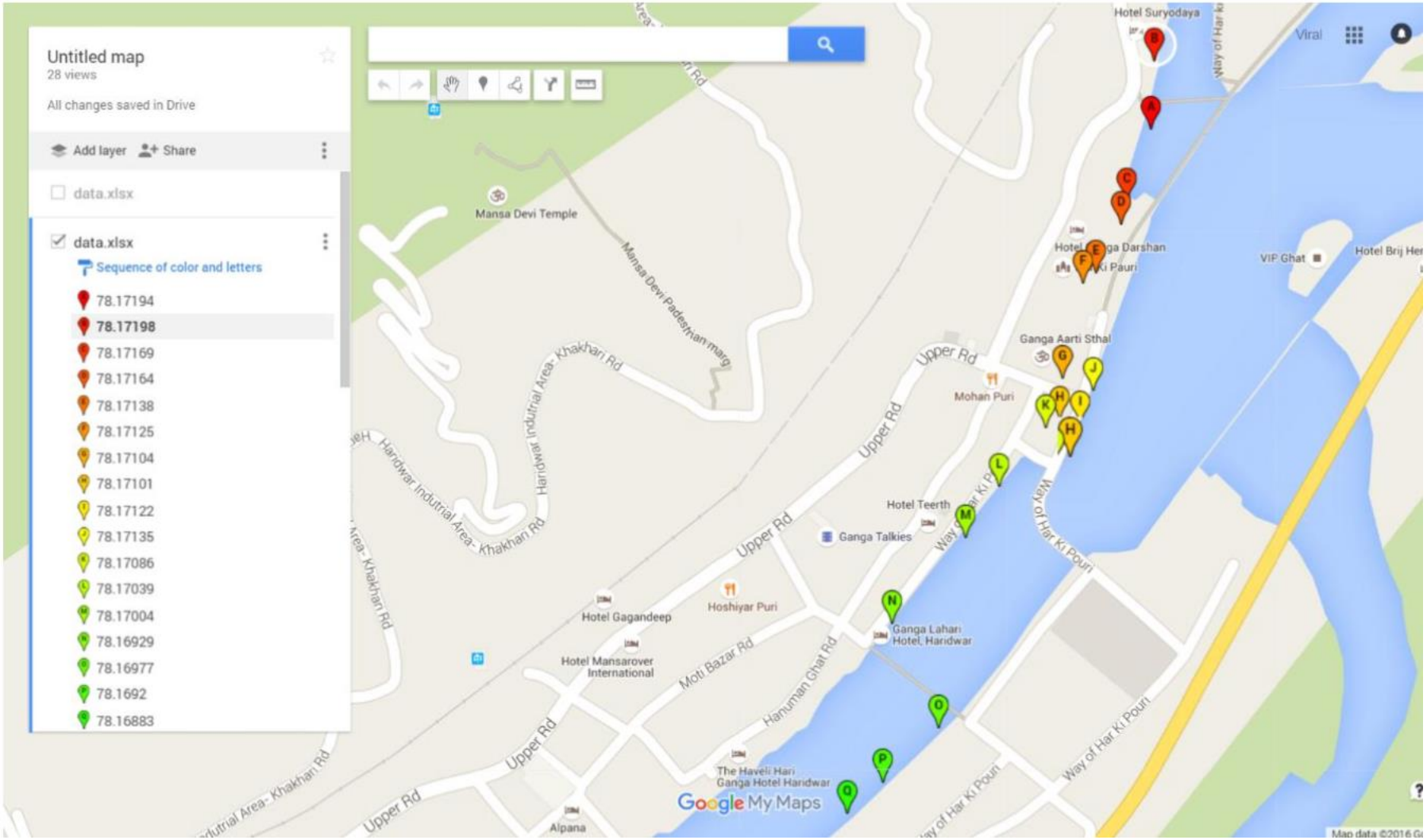
- Haridwar Ardh Khumbh Mela 2016
  - January 1, 2016 to April 30, 2016
  - Millions are expected to attend; Many will take a dip in river
  - Major bath sub-events during the period have high burst of visitors

- Question
  - How much does human activity impact river?
  - Where is the impact highest? Of what kind?

Data Collection Points around Har-ki-pauri, Haridwar

Feb 27-28, 2016

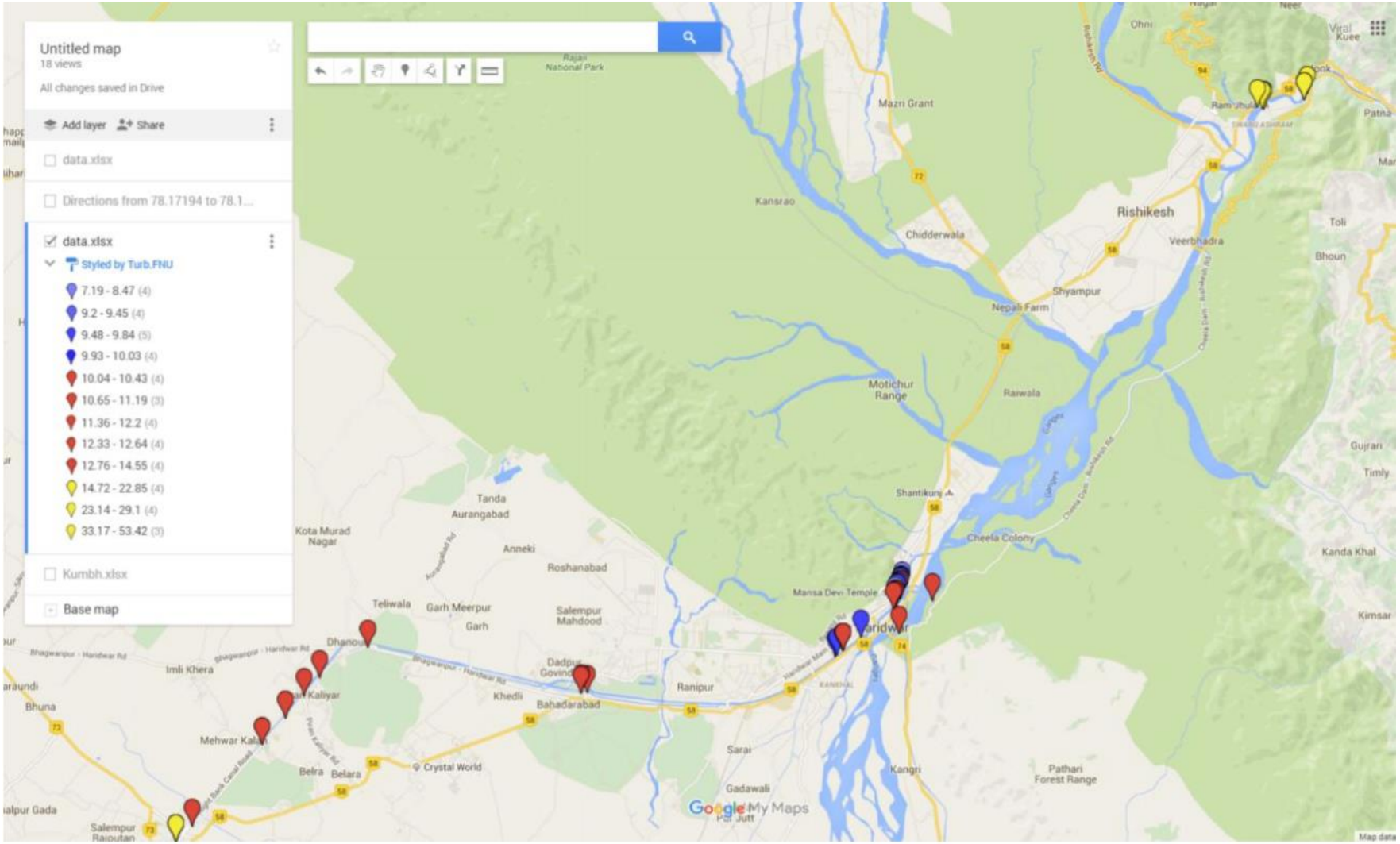
45+ places from Rishikesh to Ganga Canal (Roorkee) (75+ KM)



Carrying sensor on a buoy for long stretch was not possible due to water speed.

Turbidity Variations

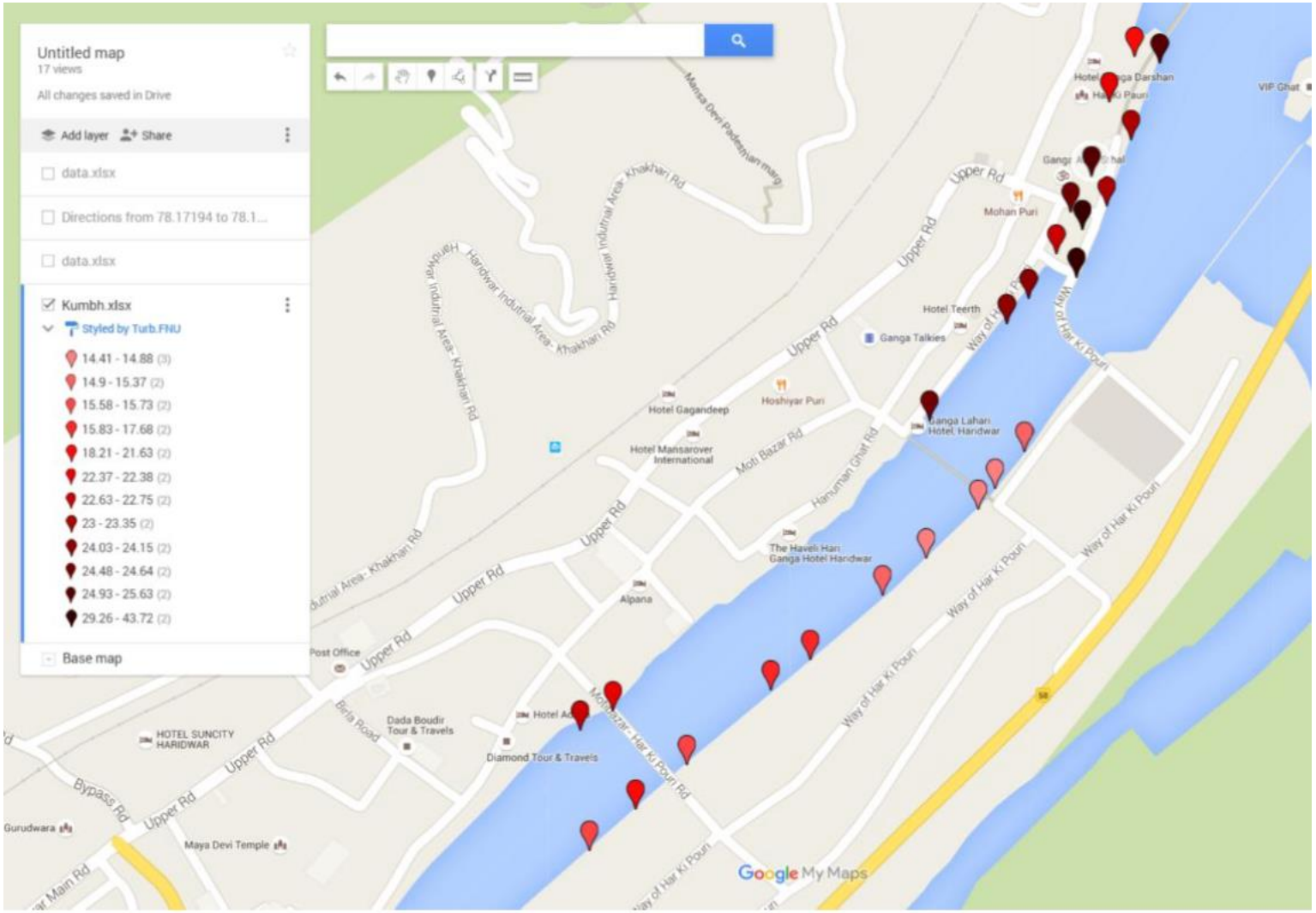
Feb 27-28, 2016



Turbidity values at different places (places marked red have turbidity value above the drinking range, places marked blues ha turbidity value in range of drinking water)

Pollution on Major Bath Day around Har-ki-pauri, Haridwar

March 7, 2016



Turbidity values at different places (places marked red have turbidity value above the drinking range, places marked blues ha turbidity value in range of drinking water)

NB Qualitative Data

http://nalanda.haifa.il.ibm.com/naturetrack/visualization.php

Crowdsourced Data

Id	Time	Latitude	Longitude	Water Level	Flow Rate	Trash	Color	Comments	Photo	Submitter	Upvote	Downvote	verified
204	2015-12-18 09:32:09	28.691198	77.229582	Some	Slow	Some	Black	yamuna, point 28, majnu tila g	photo	biplav srivastava	0	0	1
203	2015-12-18 09:20:28	28.698308	77.229535	Some	Slow	Some	Black	yamuna, point 27, wazirabad br	photo	biplav srivastava	0	0	1
202	2015-12-18 09:15:00	28.670732	77.235700	Some	Fast	None	Black	yamuna, point 26, drain+gurudw	photo	biplav srivastava	0	0	1
201	2015-12-18 09:13:02	28.670732	77.235700	Some	Fast	Some	Black	yamuna, point 25, anoth drain	photo	biplav srivastava	0	0	1
200	2015-12-18 09:02:52	28.670732	77.235700	Some	Fast	Some	Black	yamuna, point 24,drain	photo	biplav srivastava	0	0	1
199	2015-12-18 08:53:54	28.670732	77.235700	Some	Slow	Some	Black	yamuna, point 23 ,2nd rope	photo	biplav srivastava	0	0	1
198	2015-12-18 08:41:39	28.670732	77.235700	Some	Slow	Some	Black	yamuna, point 22, past rope, i	photo	biplav srivastava	0	0	1
197	2015-12-18 08:19:59	28.670546	77.236743	Some	Slow	Some	Black	yamuna, point 21, nigambodh in	photo	biplav srivastava	0	0	1
196	2015-12-18 07:12:19	28.625278	77.308005	Full	Slow	Lot	Black	yamuna, nigambodh, near bank	photo	biplav srivastava	0	0	1
195	2015-12-16 09:51:23	28.670679	77.236888	Some	Slow	Some	Black	yamuna, point 8/i, isbt, drain	photo	biplav srivastava	0	0	1
194	2015-12-16 09:32:31	28.672593	77.232038	Some	Still	Lot	Black	yamuna, point 7, red fort	photo	biplav srivastava	0	0	1
193	2015-12-16 08:57:38	28.670679	77.236888	Some	Slow	Some	Black	yamuna, point d/4, djobi	photo	biplav srivastava	0	0	1
192	2015-12-16 08:48:38	28.670679	77.236888	Some	Slow	None	Black	yamuna, point c/3, near isbt b	photo	biplav srivastava	0	0	1
191	2015-12-16 08:40:07	28.670679	77.236888	Some	Slow	Some	Black	yamuna, point b/2, in water	photo	biplav srivastava	0	0	1
190	2015-12-16 08:30:36	28.670679	77.236888	Some	Slow	Some	Black	yamuna, @nigambodh water, poin	photo	biplav srivastava	0	0	1
189	2015-12-16 07:45:28	28.670679	77.236888	Some	Still	Some	Black	yamuna, ghat 28 at nigambodh	photo	biplav srivastava	0	0	1

Data Usage – Partial Differential Equation

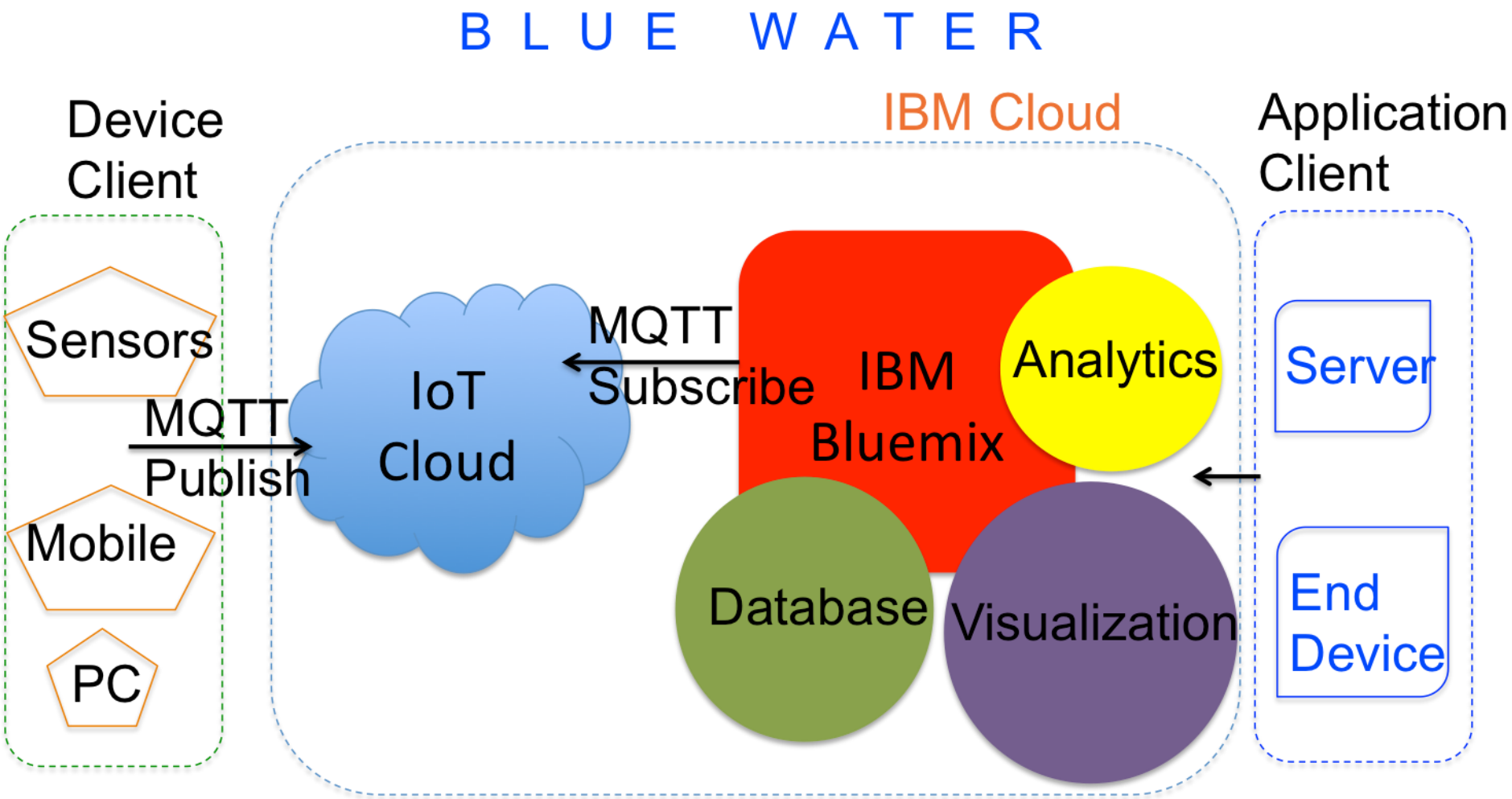
$$\frac{dc}{dt} + \mathrm{D} * \frac{d^2c}{dx^2} + \mathrm{U} * \frac{dc}{dx} + R * c + Q = 0$$

- D = diffusion term, effect of diffusion is considered using the 2nd order PDE term.
- U = velocity, convection term, effect of convection is considered using 1st order PDE term.
- R = Substance decay rate, rate of decay is considered linear, which means the more the substance is, the more it decays. (several other types of functions can also be considered depending on pollutants.)
- Q = rate of change of concentration of substance due to source. (sources can be considered implicitly like here or explicitly, in the form of boundary conditions.)

Outline

- Background – Challenges, Trends, Motivation
- Illustrative Case Study – Tanneries at Kanpur
- Pollution Sensing, Analytics Platform
  - What's new
  - Hindon @Meerut and upstream [Sep 2015]
  - Yamuna @Delhi [Dec 2015]
  - Ganga @Haridwar [Mar 2016 -]
- Discussion

Blue Water Architecture



Research Issues

- Sensing
  - How to sense cost-effectively? (Quantitative sensing)
    - Install sensors
    - Ensure sensor up-keep, inspections
  - How to involve people-as-sensors? (Qualitative sensing)
    - Use people as inspectors (increase resources for defense)
    - Mobilization when needed on short notice
    - Devising incentives for contribution

Research Issues

- Interconnection
  - Within water: quantitative and qualitative; relation between fresh and sewage water
  - Across domains: energy implications on water management, physical safety, waste water treatment

- Analytics
  - Deliver overall-value from invested assets
  - Pricing to incentivize water conservation and behavioral change

Call for Action

- Join environment community under Indian open data, http://data.gov.in
- User NeerBandhu to contribute data, use them
- Use GangaWatch app to use available data
- Focus on a water use-case and look at how you can formulate a basic problem; solve them