Panel :
Value Proposition for Big Data Analytics

Topic Overview:

Mladen Kezunovic, Coordinator/Moderator
Texas A&M University

March 13, 2019

U.S. DEPARTMENT OF ENERGY
EAC
SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

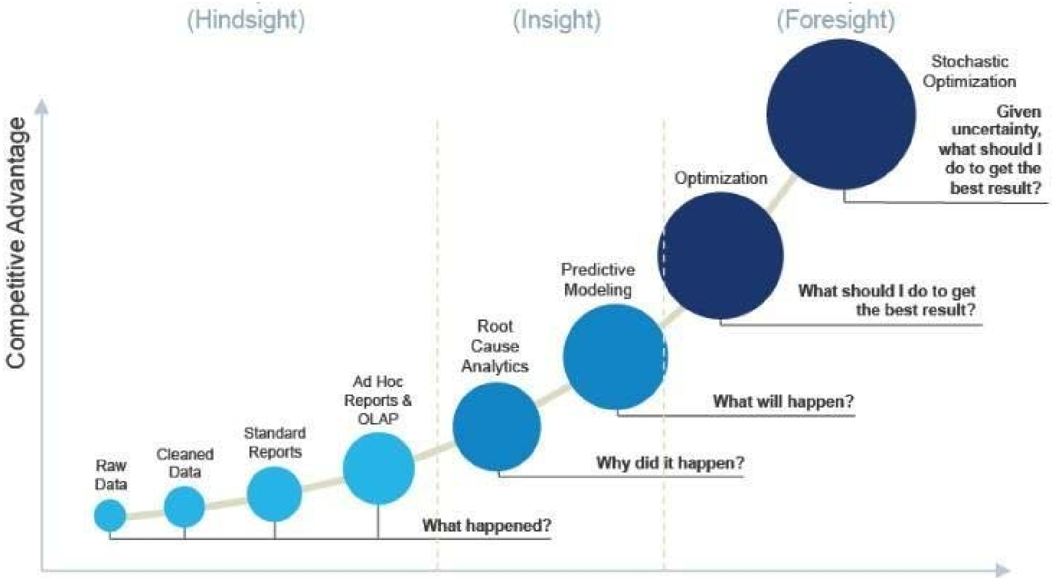
Panel : Value Proposition for Big Data Analytics

- ***Chair, Introduction and overview***: *Mladen Kezunovic*, Director, Smart Grid Center, Texas A&M University
- ***Utility Use Cases***: Doug Dorr, Program Manager, EPRI
- ***Future and visualization***: Mark Johnson, Managing Director, Utility Analytics Institute
- ***Vendors' perspective***: Mahesh Sudhakaran, Chief Digital Officer, IBM Energy and Utility business
- ***Regulatory ,Legal Issues & Consumer Advocate views***: Chris Ayers (substitute for David Colata)

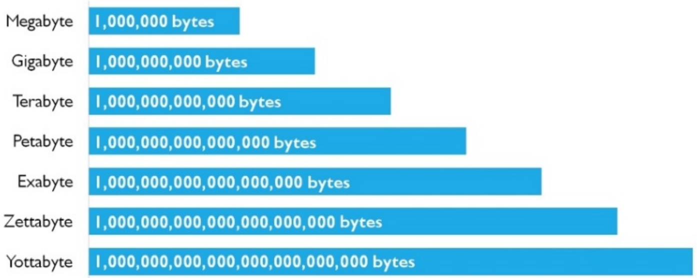
Outline

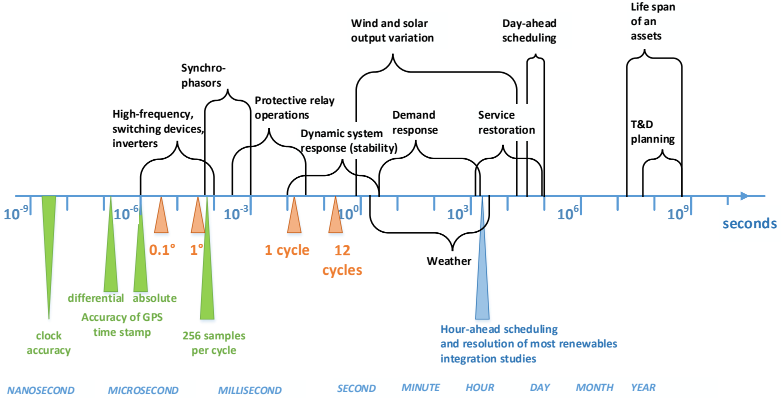
Expectations
Big Data vs. Big Data Analytics
Big Data Properties
Data Science
Big Data Processing Infrastructure
Example: Predicting outages
Panel Introduction:
- Panelist
- Questions

Expectations

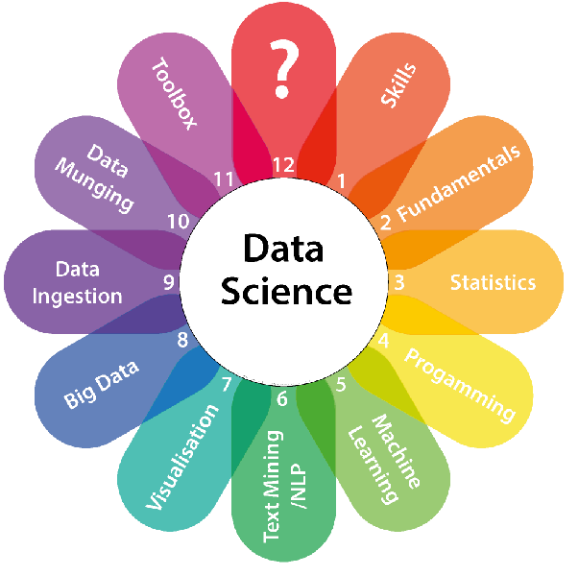


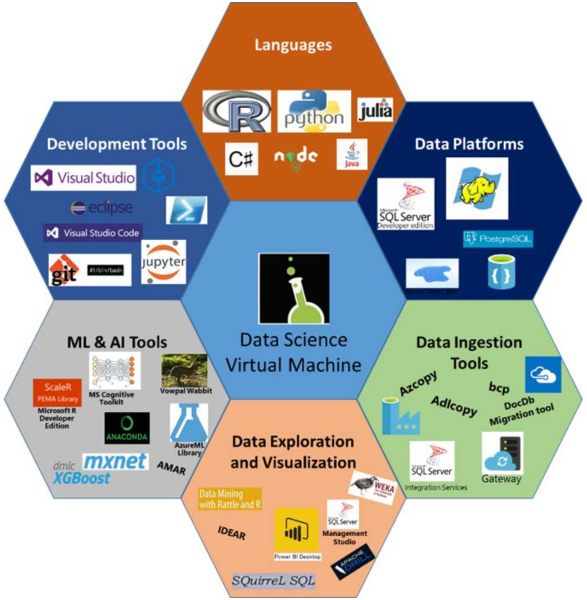
Big Data Properties





Data Science& Processing Infrastructure

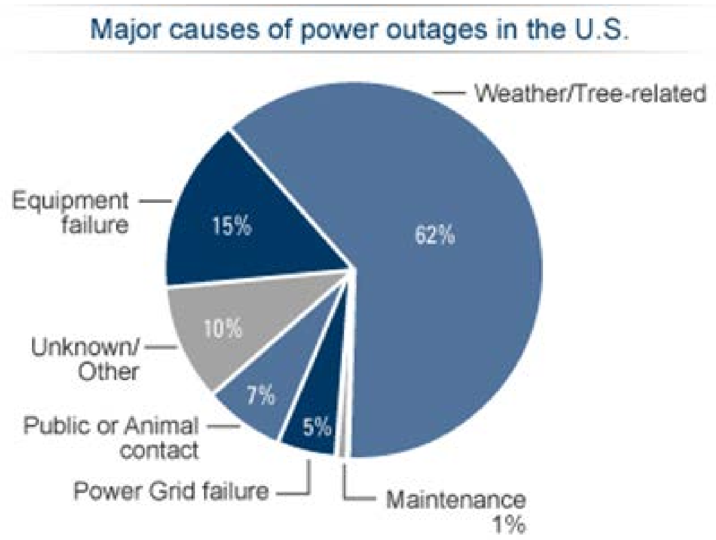


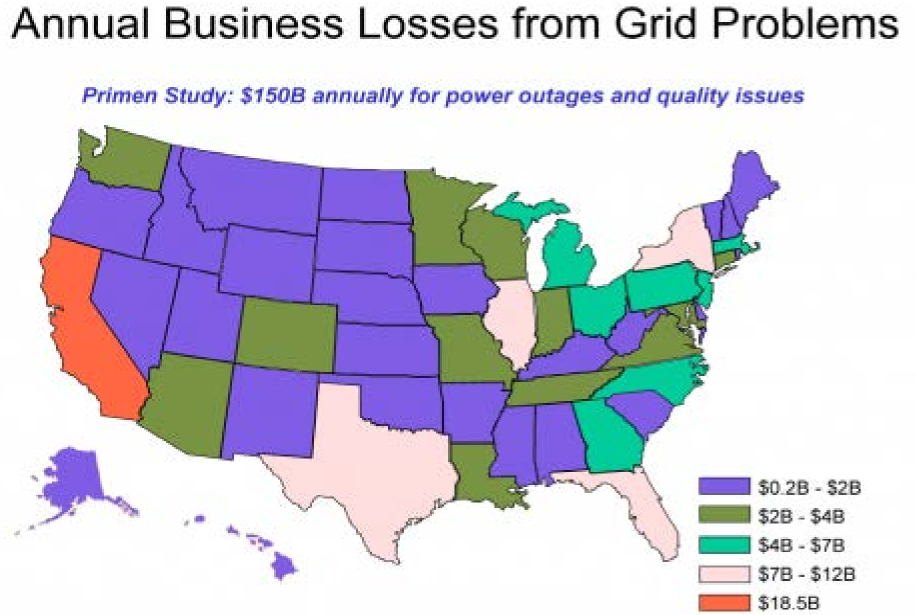


T&D Outage Prediction

Example:
Big Data Analytics

Example: outages

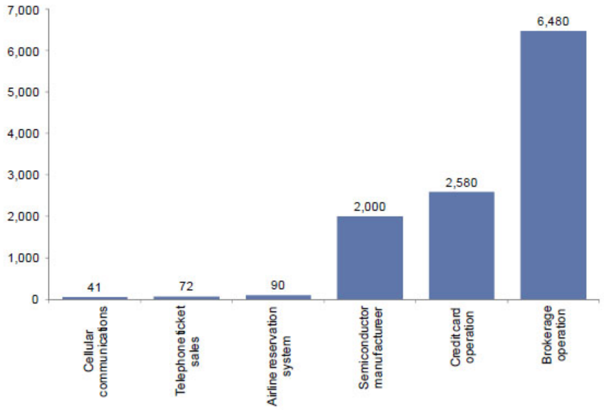




Cost

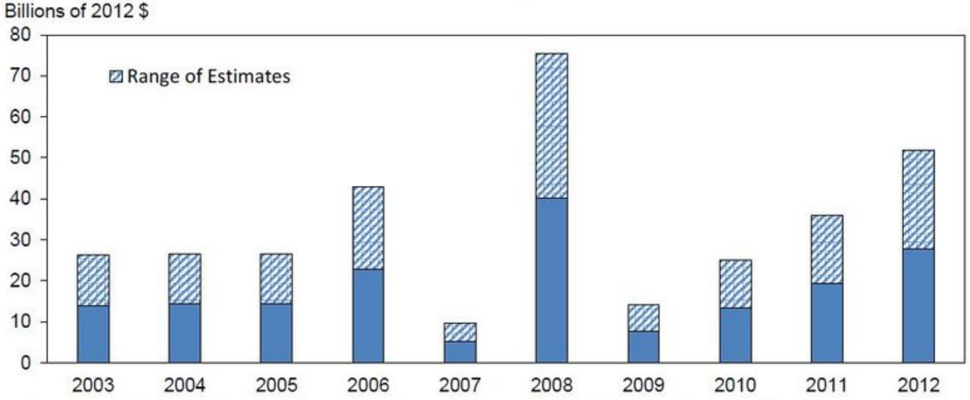
The real victim of power outages are businesses in general

US$'000 (2010); average cost of one hour power interruption in the US per type of customer



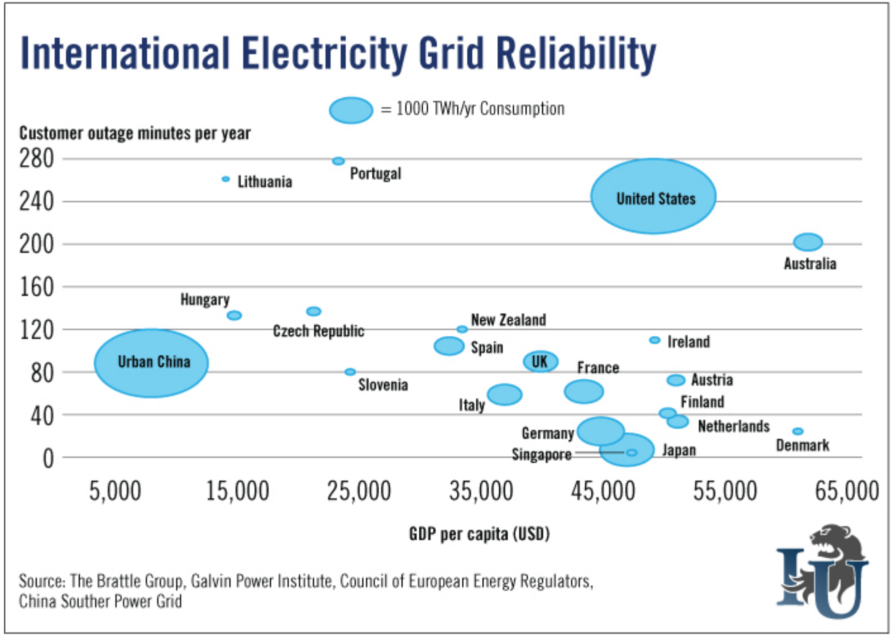
Source: US Department of Energy.

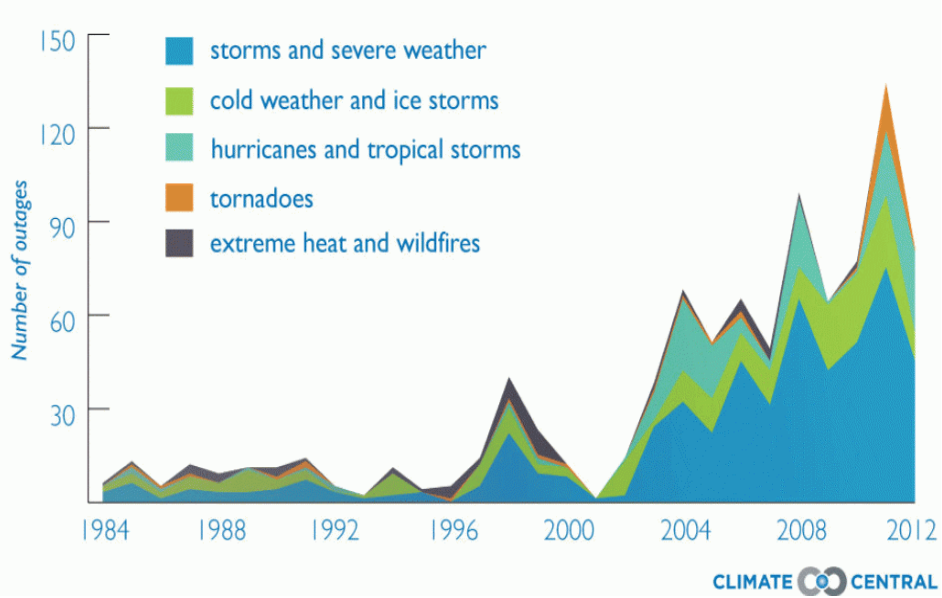
Estimated Costs of Weather-Related Power Outages



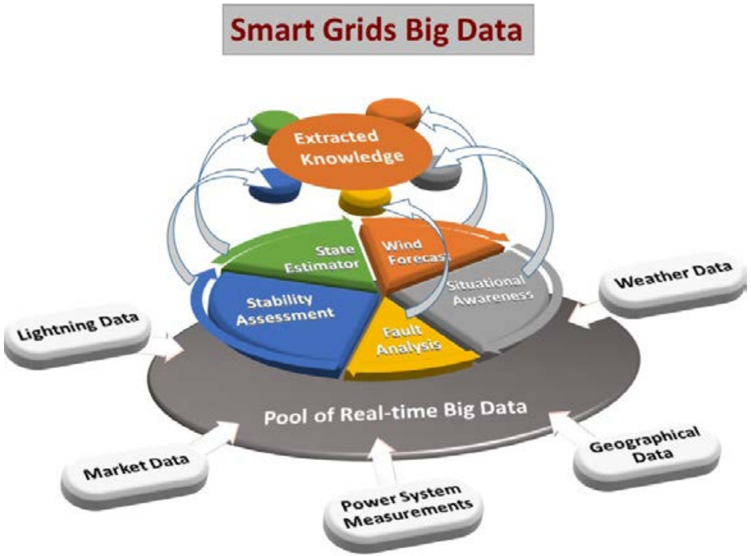
Source: CEA estimates using data from Census Bureau, Department of Energy , Energy I nformation Administration, Sullivan et al 2009.

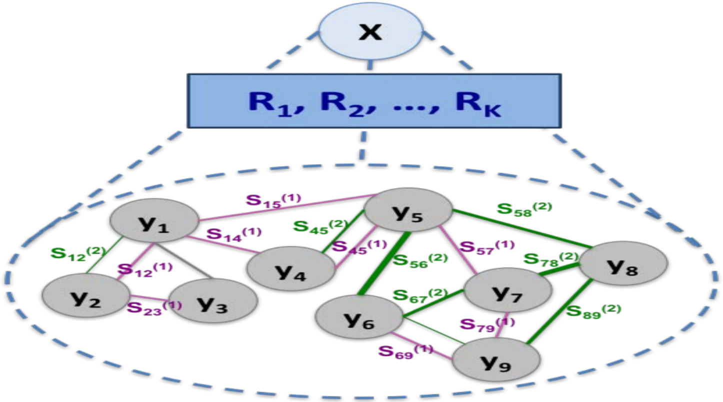
Impact

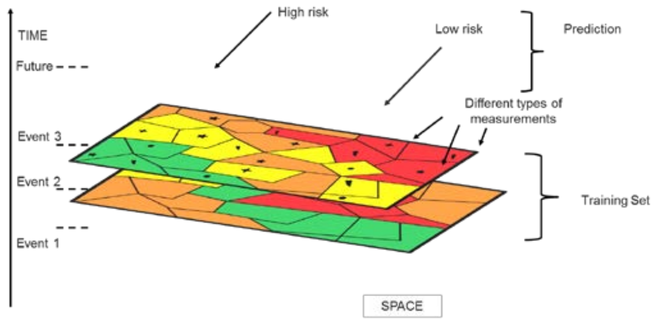

International Electricity Grid Reliability
= 1000 TWh/yr Consumption
Customer outage minutes per year
280
240
200
160
120
80
40
0
Lithuania
Portugal
United States
Australia
Hungary
Czech Republic
New Zealand
Ireland
Urban China
Spain
UK
France
Slovenia
Austria
Italy
Finland
Netherlands
Germany
Singapore
Japan
Denmark
5,000
15,000
25,000
35,000
45,000
55,000
65,000
GDP per capita (USD)
Source: The Brattle Group, Galvin Power Institute, Council of European Energy Regulators, China Souther Power Grid


150
120
90
60
30
Number of outages
storms and severe weather
cold weather and ice storms
hurricanes and tropical storms
tornadoes
extreme heat and wildfires
1984
1988
1992
1996
2000
2004
2008
2012
CLIMATE CENTRAL

Predictive Data Analytics

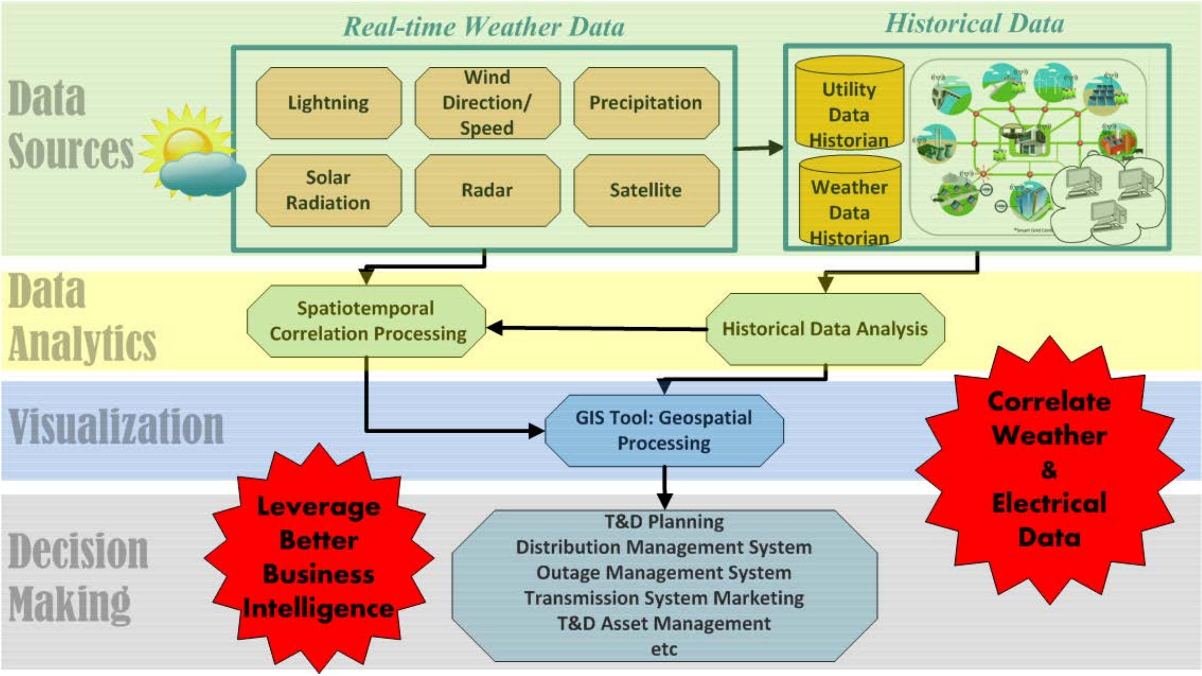




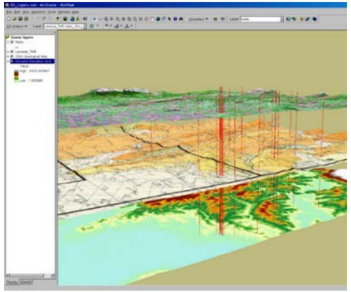


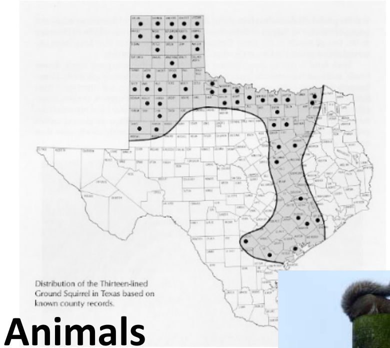
M. Kezunovic, Z. Obradovic, T. Dokic, B. Zhang, J. Stojanovic, P. Dehghanian, and P. -C. Chen, "Predicating Spatiotemporal Impacts of Weather on Power Systems using Big Data Science," Springer Verlag, Data Science and Big Data: An Environment of Computational Intelligence, Pedrycz, Witold, Chen, Shyi-Ming (Eds.), ISBN 978-3-319-53474-9, 2017.

BD for Risk Assessment

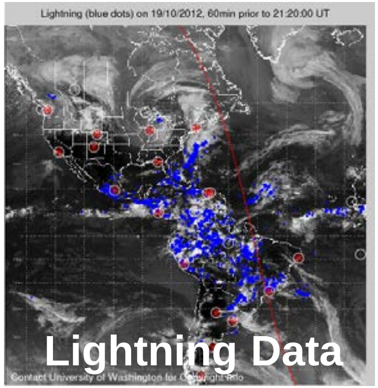


BD Data Aggregation



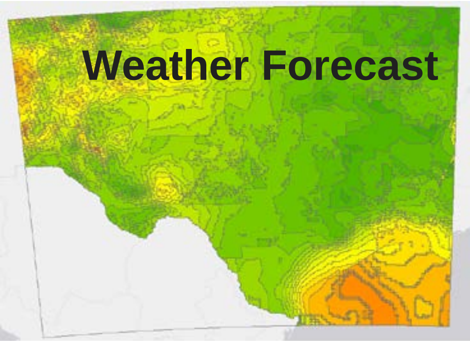


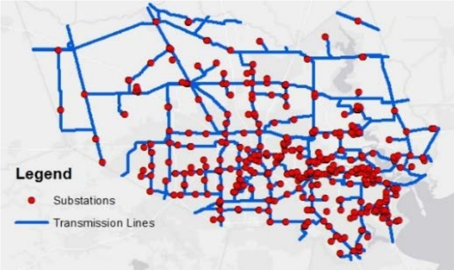
Animals Data



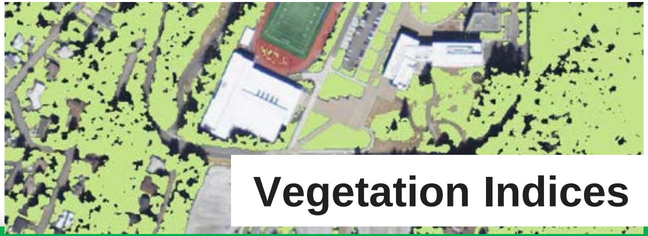
UAS

GIS





Network Assets Data



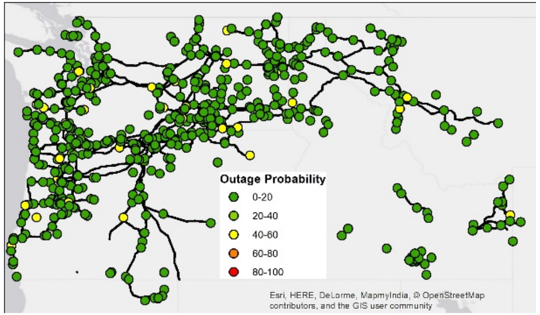
BD Data Properties

			VELOCITY	VOLUME	
	Source	Data Type	Temporal Resolution	Spatial Resolution	Measurements
V A R I E T Y	Automated Surface Observing System	Land-Based	1 min	900 stations	Air Temperature, Dew Point, Relative Humidity, Wind Direction, Speed and Gust, Altimeter, Sea Level Pressure, Precipitation, Visibility…
	Level-2 Next Generation Weather Radar	Radar Data	5 min	160 high-resolution Doppler radar sites	Precipitation and Atmospheric Movement
	NOAA Satellite Database	Satellite Data	Hourly, daily, monthly	4 km	cloud coverage, hydrological observations (precipitation, cloud liquid water, total precipitable water, snow cover…), pollution monitoring…
	Vaisala U.S. National Lightning Detection Network	Lightning Data	Instantaneous	Median Location Accuracy 200-500m	Date and Time, Latitude and Longitude, Peak amplitude, Polarity, Type of event: Cloud or Cloud to Ground
	National Digital Forecast Database	Weather Forecast Data	3 hours	5 km	Wind Speed, Direction, and Gust, Relative Humidity, Convective Hazard Outlook, Tornado Probability, Probability of Thunderstorms…
	Texas Parks & Wildlife Department	Texas Ecological Mapping Systems Data	static	10 m	Distribution of different tree spices
	Texas Natural Resources Information System	NAIP	year	50 cm – 1 m	High Resolution Imagery
	National Aeronautics and Space Administration	3D Global Vegetation Map	static	1 km	Canopy height data
	National Cooperative Soil Survey	gSSURGO	static	10 m	Soil type
	Utility	Historical Outage Data	instantaneous	Feeder section	Location, start and end time and date, number of customers affected, cause code
		Tree Trimming Data	day	Feeder	Feeder location, date, trimming period, number of customers affected, cost of trimming
		Network GIS data	static	Infinity (shapefile)	Poles: location, material/class, height Feeders: location; conductor size, count, and material; nominal voltage
		Historical Maintenance Data	day	Tower location	Start and end date and time, location, type (maintenance, replacement), cost, number of customers affected
		Insulator asset data	static	Infinity (shapefile)	Surge Impedances of Towers and Ground Wires, Footing Resistance, Component BIL
		In-field measurements	instantaneous	Tower location	Leakage Current Magnitude, Flashover Voltage, Electric Field Distribution, Corona Discharge Detection, Infrared Reflection Thermography, Visual Inspection Reports

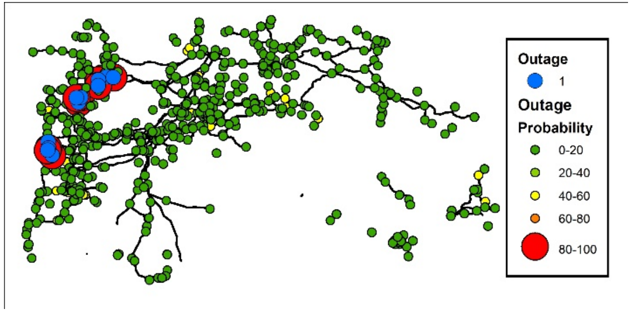
	Data Class	Data Source (Measurements)	VOLUME (Data file size)	VELOCITY (Rate of use)	VERACITY (Accuracy)
V A R I E T Y	Utility measurements	SM	120GB per day/ device	Every 5-15 min	error <2.5%
		PMU	30GB per day/device	240 samples/sec	error <1%
		ICM	5GB per day/device	250 samples/sec	error <1%
		DFR	10MB per fault/device	1600 samples/sec	error <0.2%
	Weather data	Radar [27]	612 MB/day per radar	Every 4-10 min	1-2 dB; $m\ s^{-1}$
		Satellite [28]	At least 10 GB per day	Every 1-15 min	VIS<2%; IR<1-2K
		ASOS [29]	10 MB/day per station	Every 1 min	T-1.8°F, P<1%, Wind speed - 5%, RR - 4%
		NLDN [30]	40 MB/day	During lightning	SE < 200m, PCE <15%
		NDFD [31]	5-10 GB/day per model	1 - 12 hours	Varies by parameter
	Vegetation and Topography	TPWD EMST [32]	2.7 GB for Texas	static	SE < 10 m
		TNRIS [33]	300 GB for Texas	static	SE < 1 m
		LIDAR [34]	7 GB for Harris Co.	static	HE < 1m, VE < 150 cm

BD Analytics Outcomes

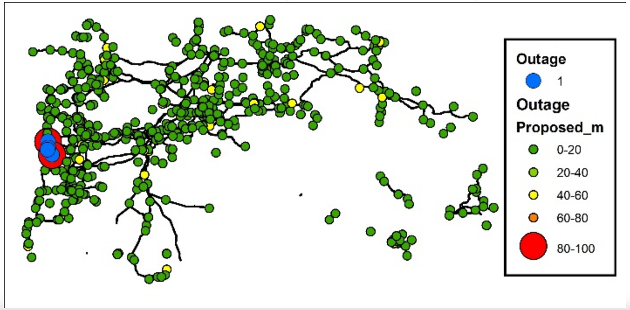
Probabilities of outages for no outage



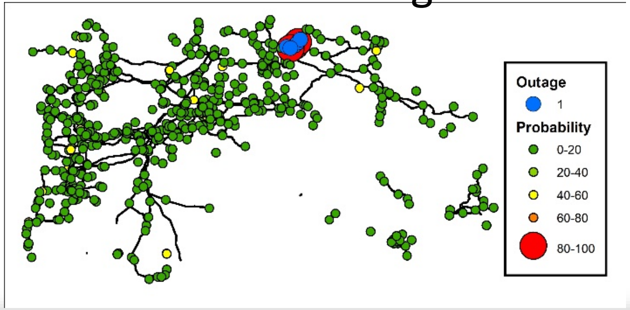
Probabilities of outages for lightning



Probabilities of outages for vegetation



Probabilities of outages for ice



Takeaways

- Extensive research is needed to bring BD Analytics into utility practice:

  - Data analytics has been used in the power system domain for over 50 years, but Big Data Analytics is in its infancy

  - The Big Data Applications require intensive and costly effort to prepare the data (ingestion, cleansing, curation)

  - The gap between the Big Data platforms and utility legacy software (EMS, DMS, MMS) uses is huge, and costly

  - Utility predictive methods do not explore data sciences advances (Deep learning, spatiotemporal scaling, etc.)

- The government contribution may be in the following areas:

  - Make government sources of data readily useable

  - Fund research in new applications of Big Data Analytics

  - Help industry demonstrate the new business opportunities

  - Explore benefits of predictive methods in solving grand challenges

Panel : Value Proposition for Big Data Analytics

- ***Chair, Introduction and overview***: *Mladen Kezunovic*, Director, Smart Grid Center, Texas A&M University
- ***Utility Use Cases***: Doug Dorr, Program Manager, EPRI
- ***Future and visualization***: Mark Johnson, Managing Director, Utility Analytics Institute
- ***Vendors' perspective***: Mahesh Sudhakaran, Chief Digital Officer, IBM Energy and Utility business
- ***Regulatory ,Legal Issues & Consumer Advocate views***: Chris Ayers (substitute for David Colata)

Panel Questions

Use Cases: What is the state of the art in Big Data Analytics products and R&D developments suitable for power industry applications?

Barriers and gaps: What prevents faster development and deployment of the solutions that utilize Big Data Analytics? Is there a compelling business case (value) for the stakeholders (utilities, ISO's, Load Serving Entities, third party aggregators, data providers) to adopt Big-Data Analytics? What is this business case?

Customer role and needs: How to access data related to energy consumption that resides at the customer site or is collected at the points of customers interfacing to the grid? How to distribute data and knowledge to end-use customers in a manner that facilitates decisions?

***Regulatory and legislative framework**:* How does such framework shape crucial issues in data access such as cybersecurity, privacy, data ownership, critical infrastructure confidentiality, data as service?

U.S. DEPARTMENT OF ENERGY

EAC

TEES

SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION