Smart Grid Big Data Analytics: Applications

Mladen Kezunovic, Ph.D., P.E.
Regents Professor
Texas A&M University
June 4, 2019

Outline

Background:

- The evolving Grid
- Resilience

The Grid Edge:

- Application centric view
- Data centric view

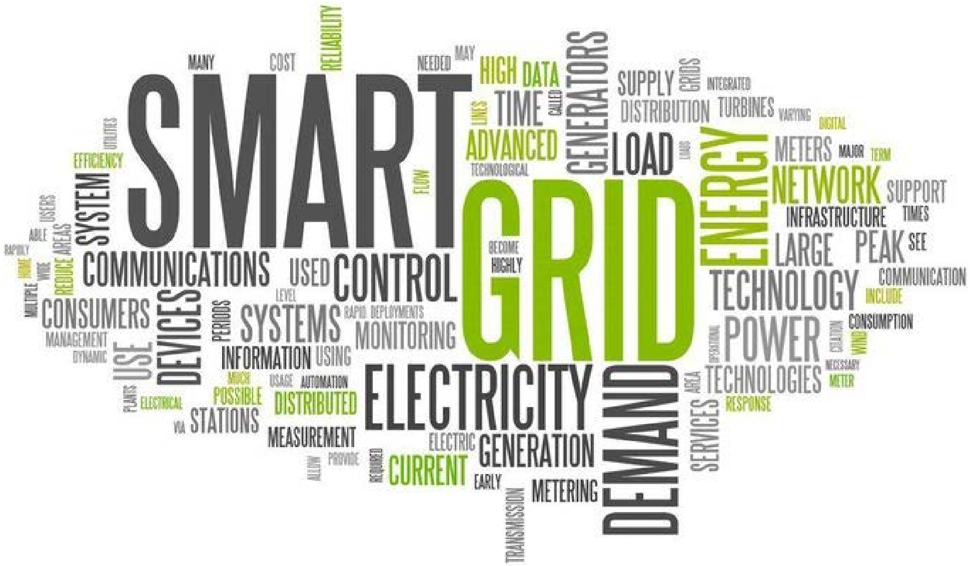
Big Challenges

- Big Data Properties
- Expectations

Example: Predicting outages/failures

- Transmission
- Distribution

Takeaways

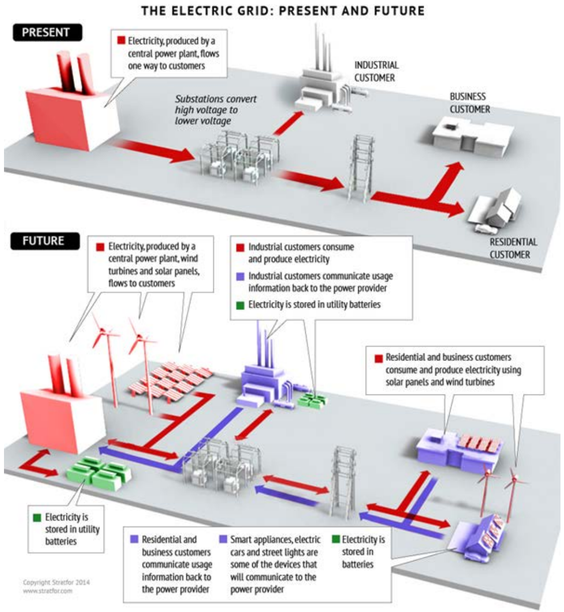


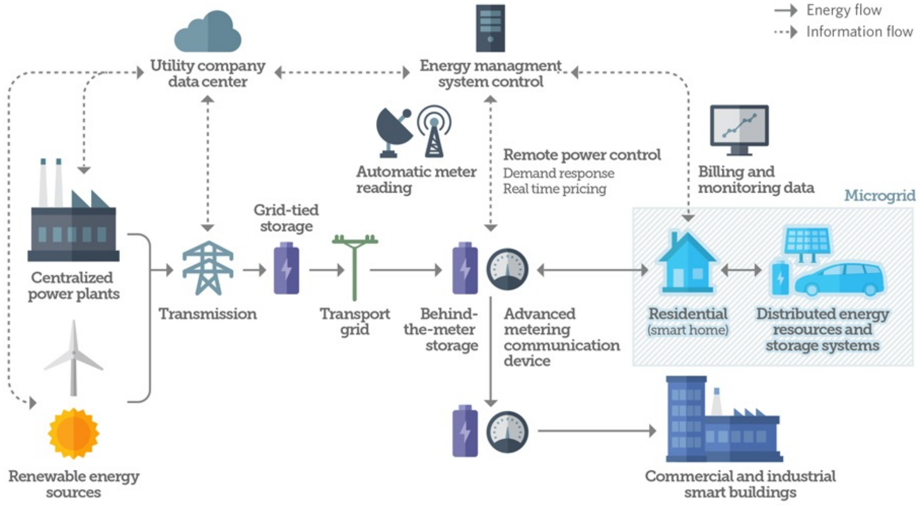
Outline

Background:

- **The evolving Grid**
- **Resilience**

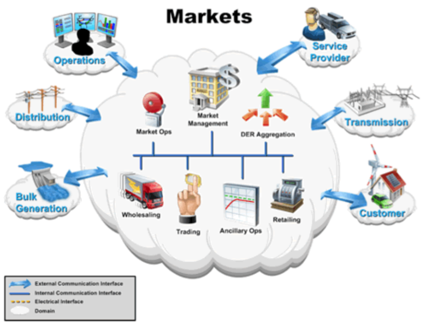
The evolving grid

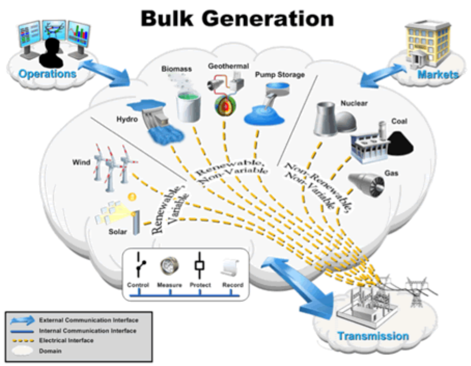


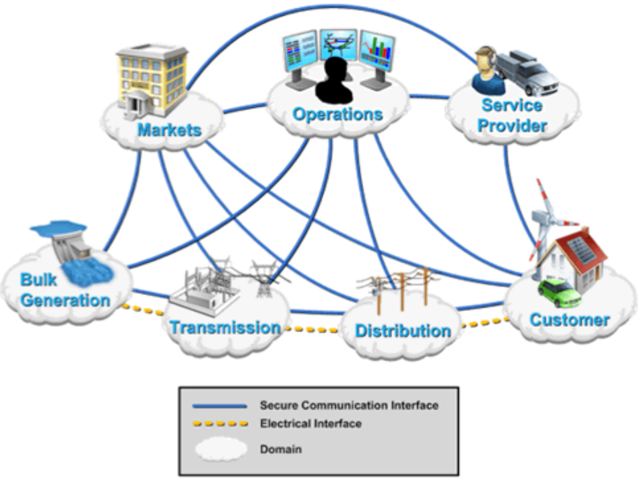


Source: LG CNS

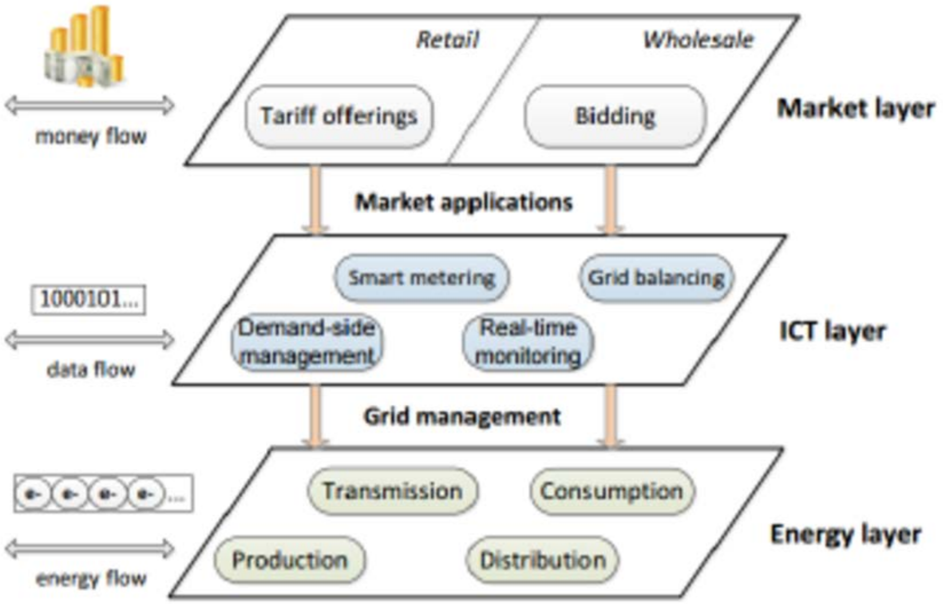
The evolving grid



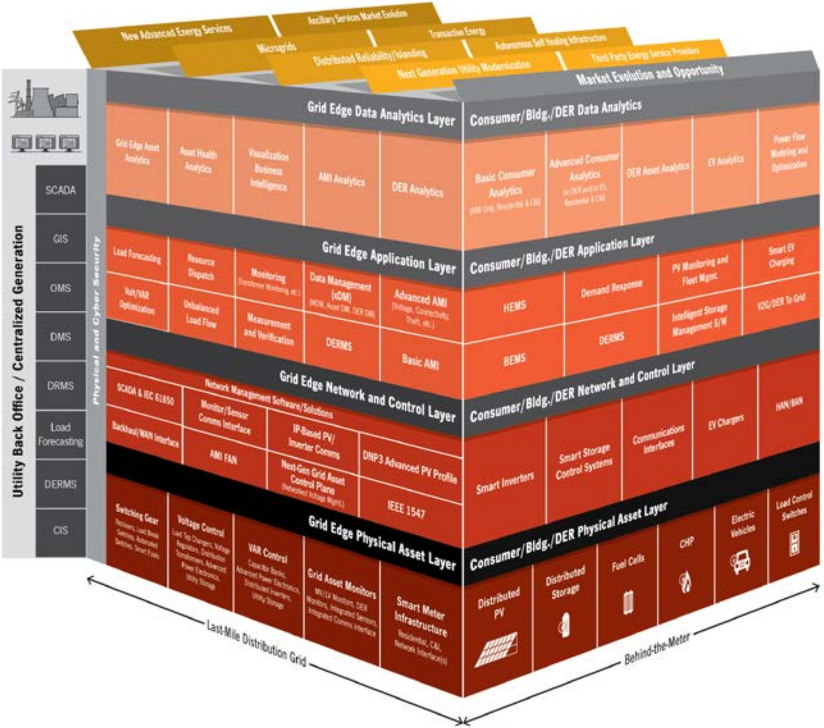




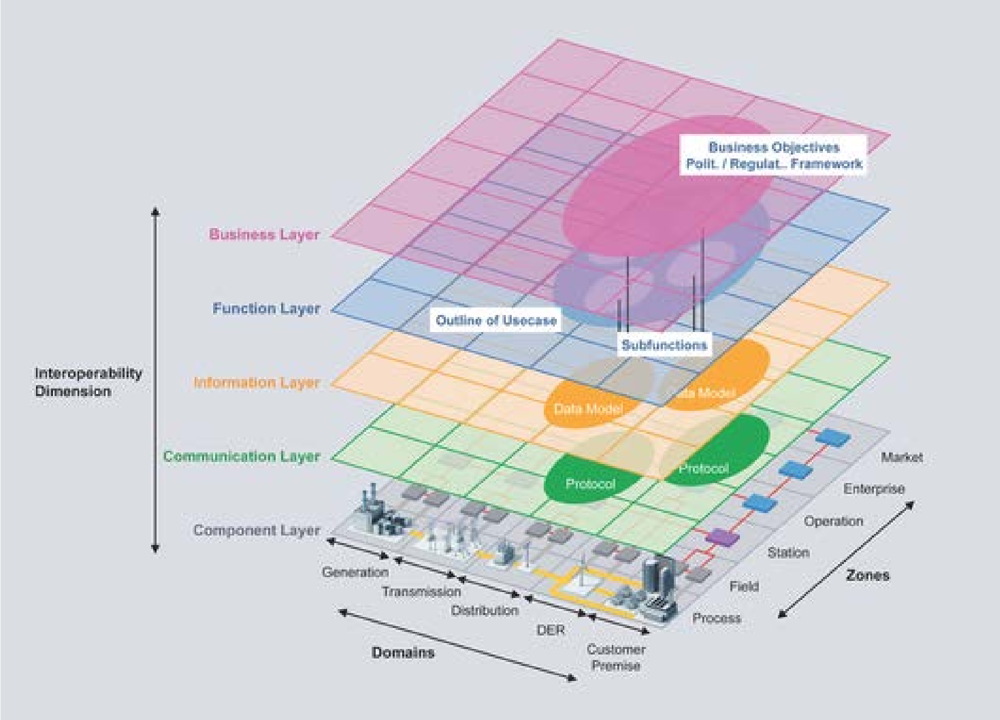
Layered Architecture



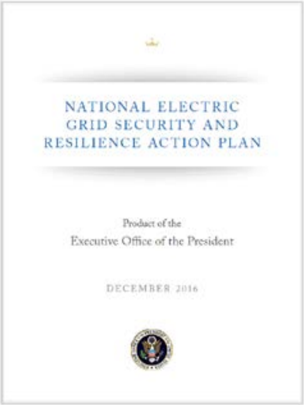
Control/application layer

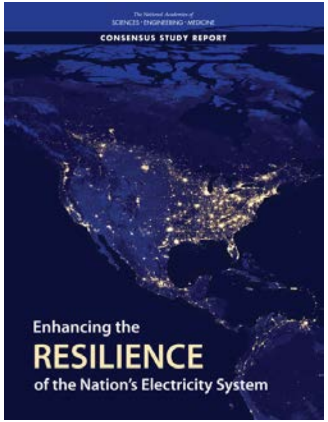


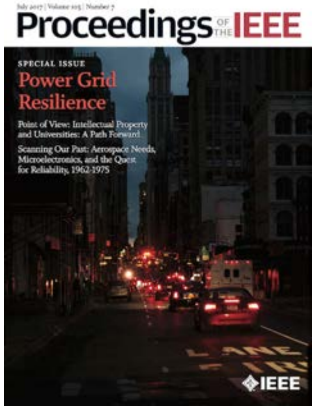
Interoperability layer

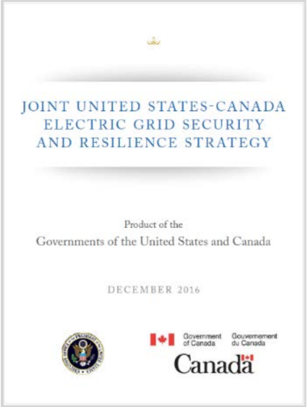


Resilience







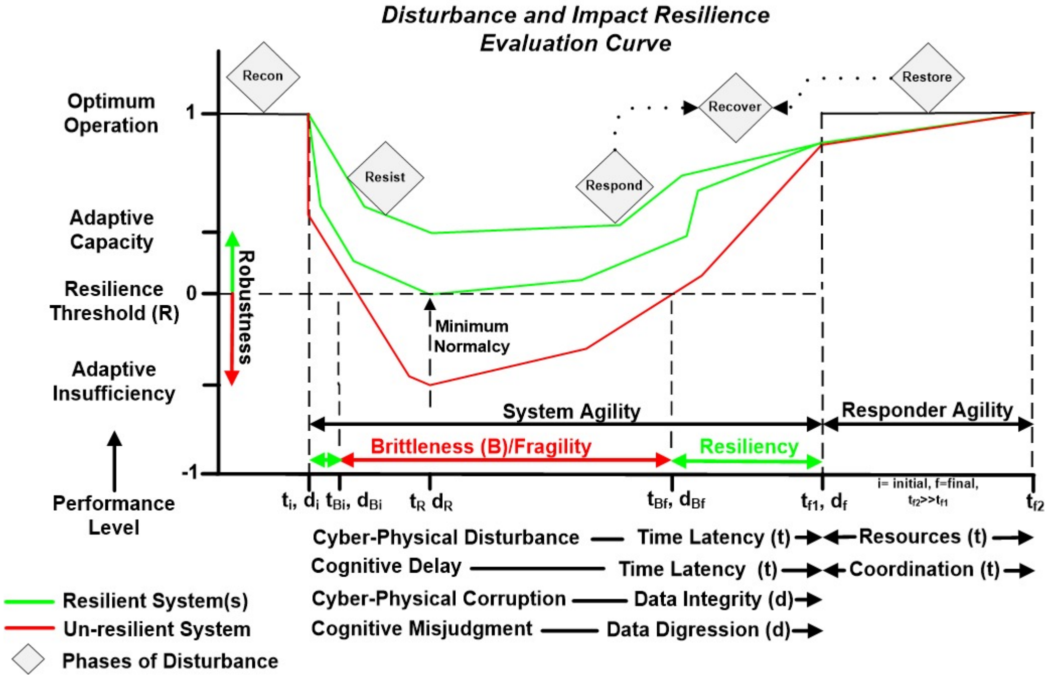


- *degrades gradually, and not abruptly, when it experiences stressed conditions and it is able to restore back into its normal state thereafter*
- *learns from its previous lessons and experiences under major disturbances and uses this knowledge to adapt and fortify itself to prevent or mitigate the consequences of a similar event in the future.*
- *minimizes interruptions of service during an extraordinary and hazardous event*
- *anticipates, absorbs, adapts to and/or rapidly recovers from a disruptive event*
- *plans and prepares for a disruptive event, absorbs it and is able to recover from it*

Resilience

K. Eshghi, B.K. Johnson, C. G Rieger"Power System Protection and Resiliency Metrics." 2015 Resilience Week, Workshop Proceedings, Idaho National Laboratory, August 2015

T. Mc. Junkin, C.G. Rieger, "Electricity distribution system resilient control system metrics." 2017 Resilience Week, Workshop Proceedings, Idaho National Laboratory, September 2017



Design Requirements: 3RAP

- ***Robustness*** (withstand low probability but high consequence events),
- ***Resourcefulness*** (effectively manage a disturbance as it unfolds),
- ***Rapid recovery*** (get things back to normal as fast as possible after the disturbance),
- ***Adaptability*** (absorb new lessons from a catastrophe).
- ***Predictability*** (learns from the past and anticipates future disturbances)

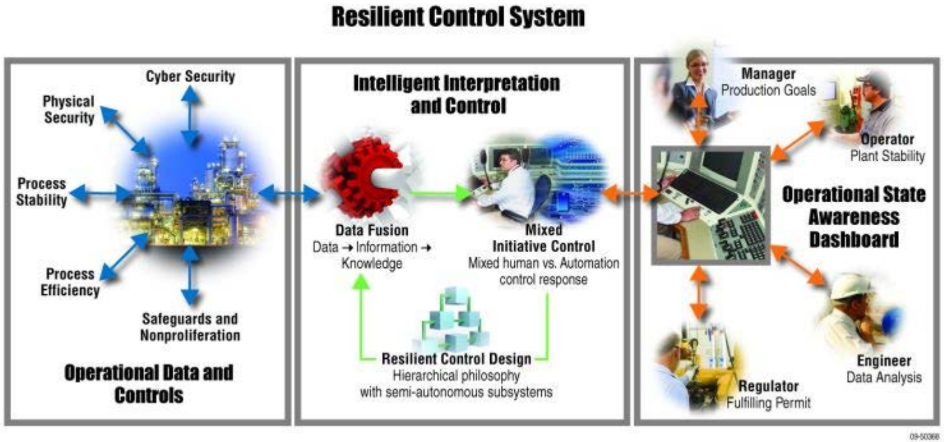
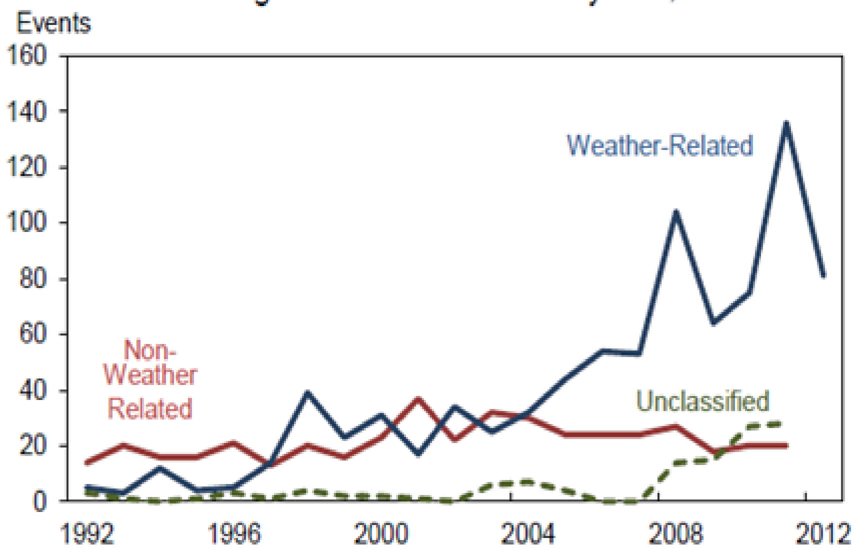


Figure: Resilient Control System Framework (Wikipedia)

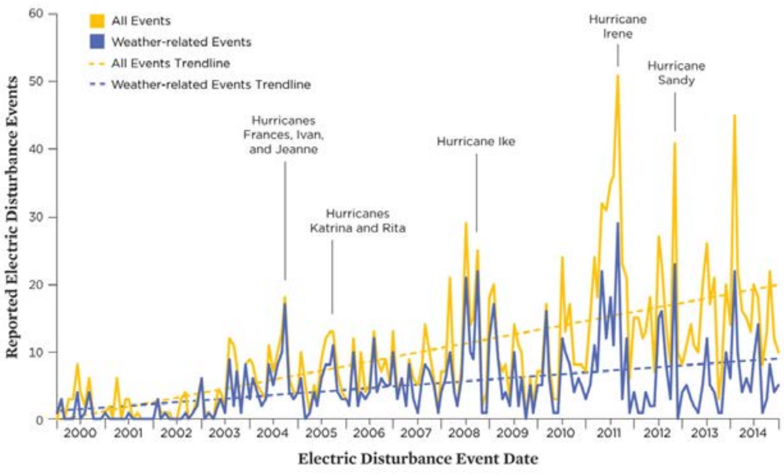
Manifestation

Observed Outages to the Bulk Electric System, 1992-2012



Source: Energy Information Administration

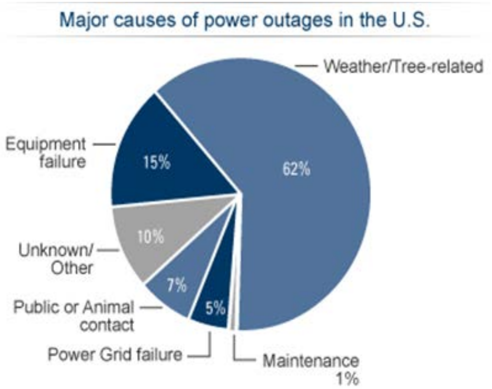
FIGURE 1. U.S. Electric Grid Disruptions

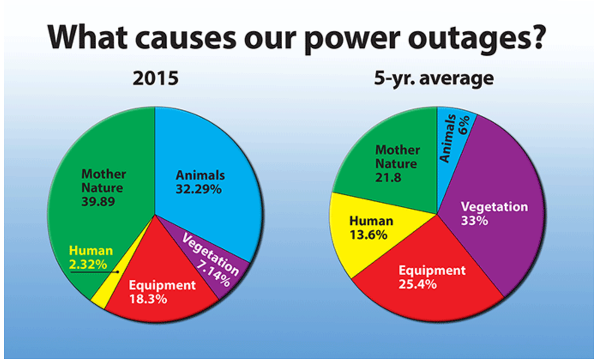


The Department of Energy tracks major electric disturbance events through Form OE-417. Utilities submit information about qualifying incidents, including when they occurred, where they occurred, what triggered them, and how many customers were affected. Notably, while the reported number of non-weather-related events is high, the vast majority of incidents resulting in customer outages occur because of weather.

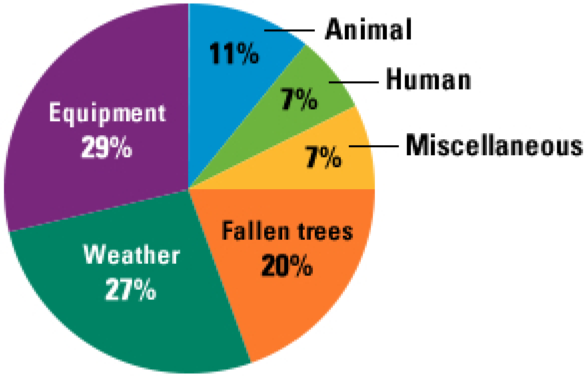
SOURCE: UCS ANALYSIS, BASED ON OE N.D.

Causes

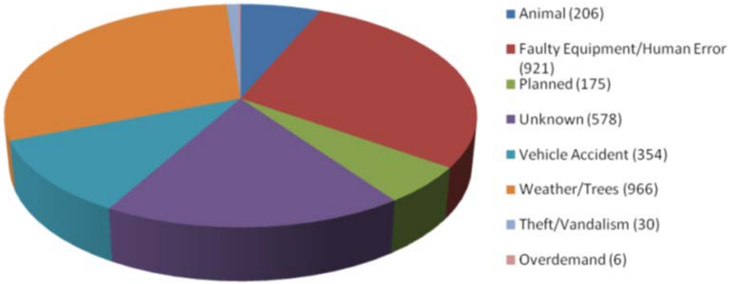




Source: Alaska Electric light and Power Company



Source: We Energies

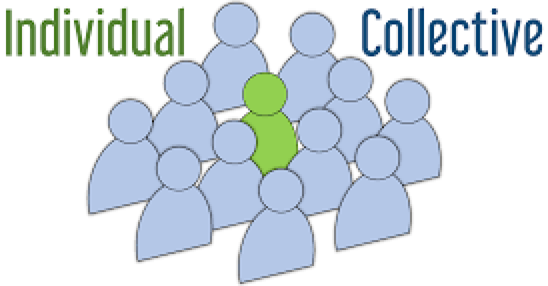


Source: Annual Eaton Investigation 2013

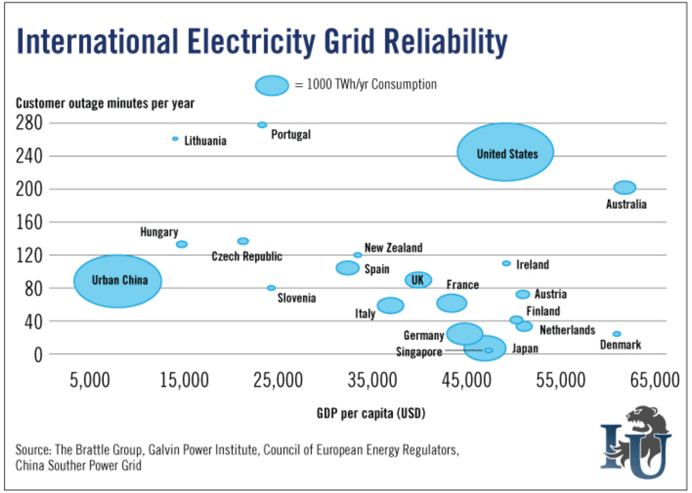
Outline

The Grid Edge:

- Application centric view
- Data centric view



Application centric view



Utility Centric

The "Edge" Centric

The Grid View

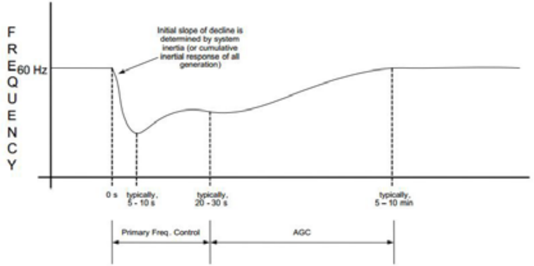
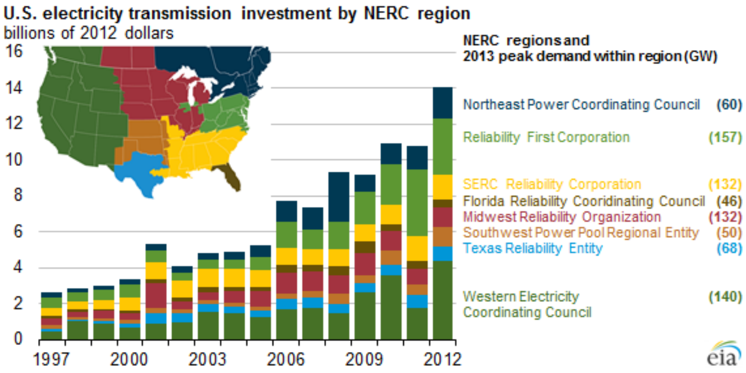
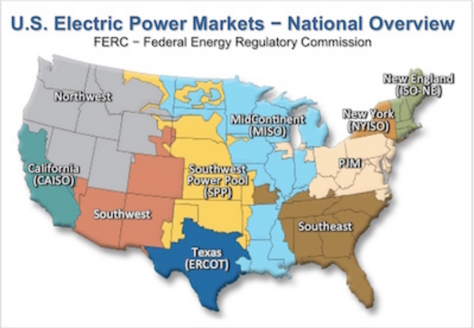
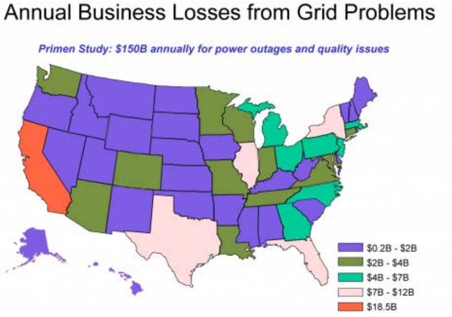
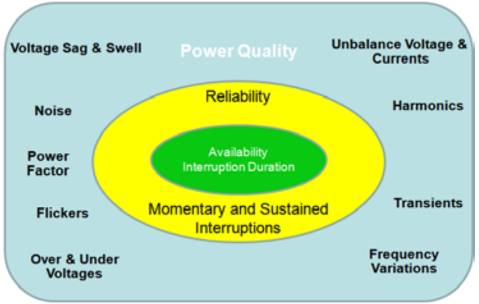


Figure ES-2. Frequency trace following a large contingency event (i.e., loss of a large generating unit). Inertial control, PFC, and AGC (secondary frequency control) each serve a different purpose, and their response timeframes are also at different points of the frequency recovery.

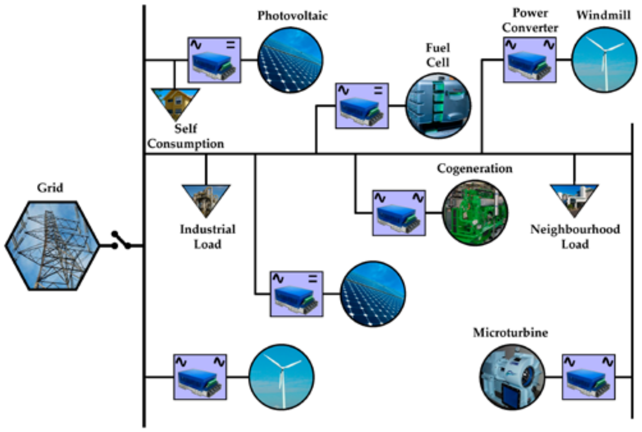


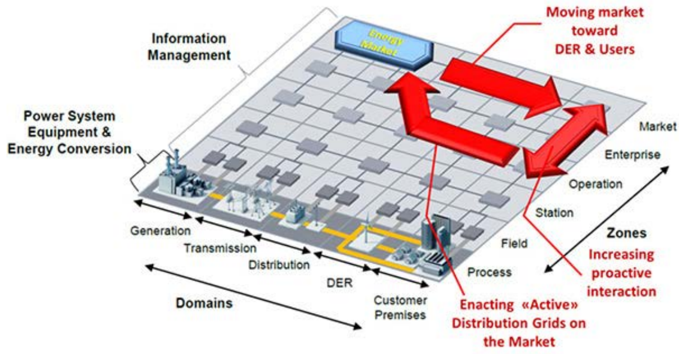


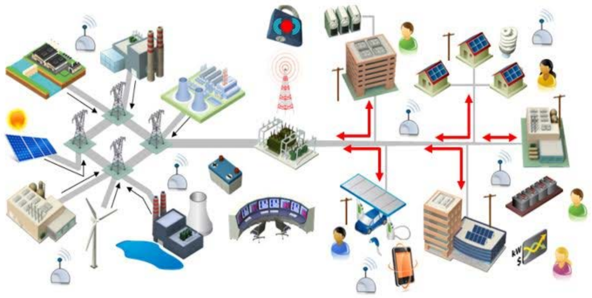




The Edge View







Data Centric view

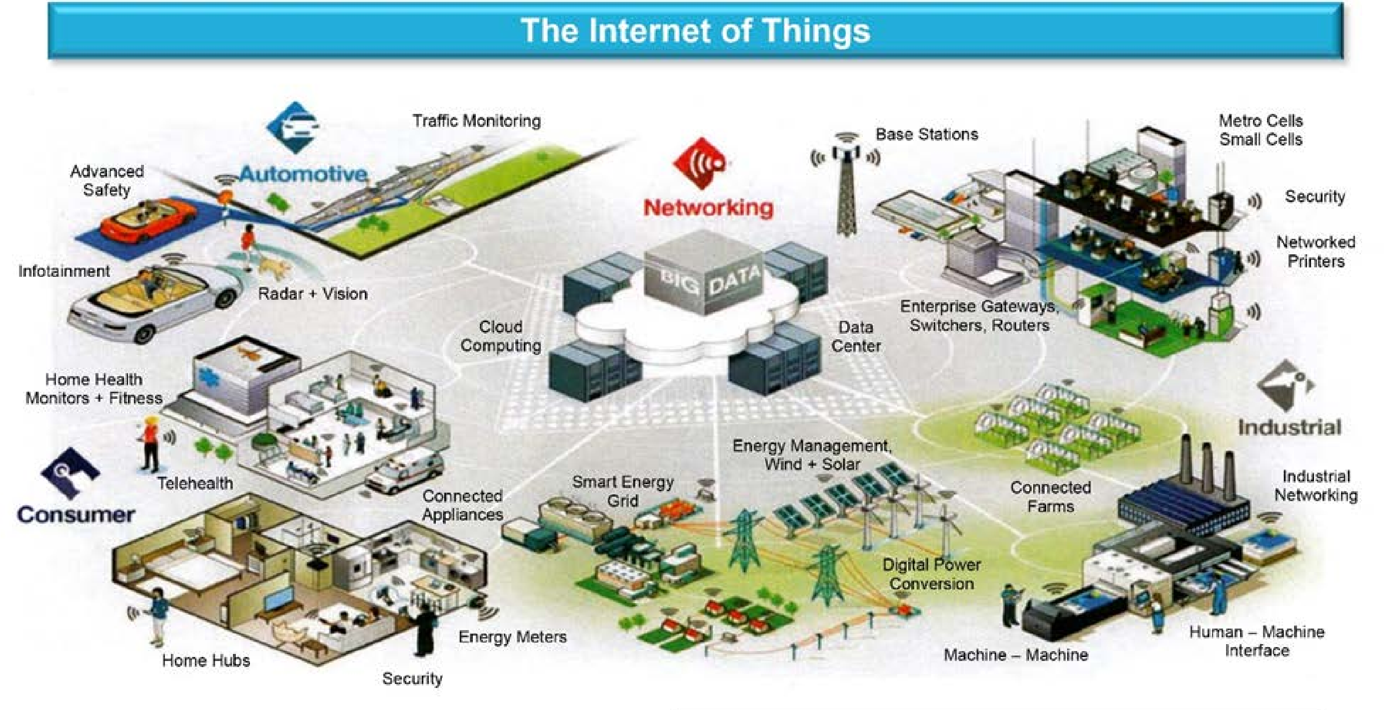
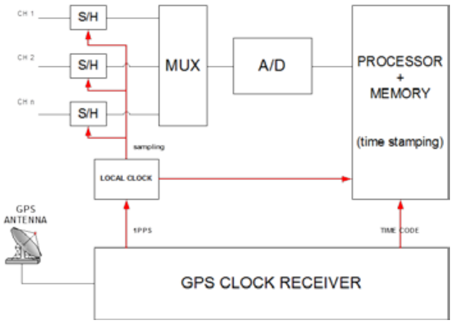
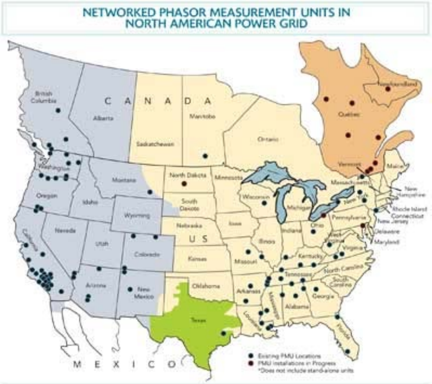
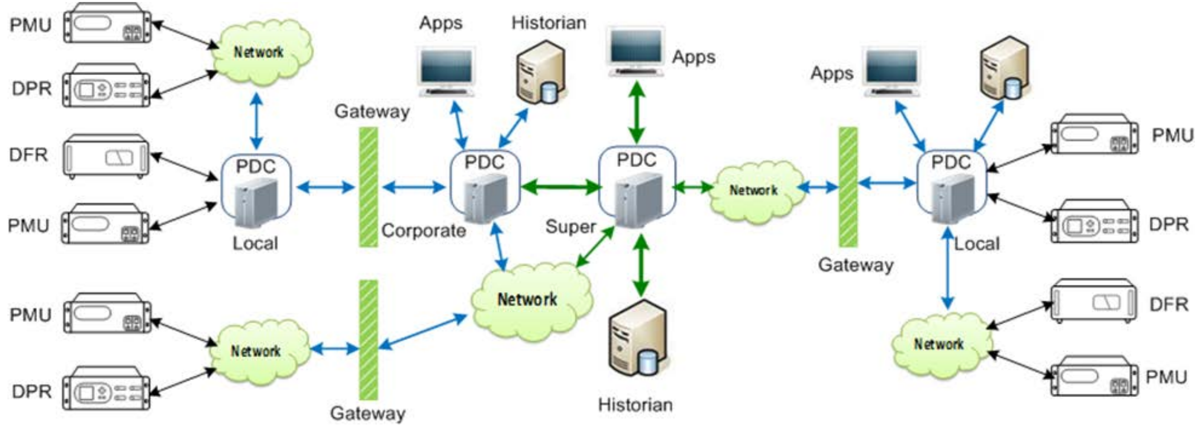


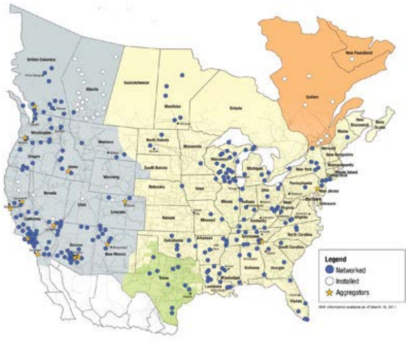
ILLUSTRATION CREDIT: The Register. Website: http://regmedia.co.uk/2014/05/06/freescale_internet_of_things_overview_1.jpg

Synchrophasor Data





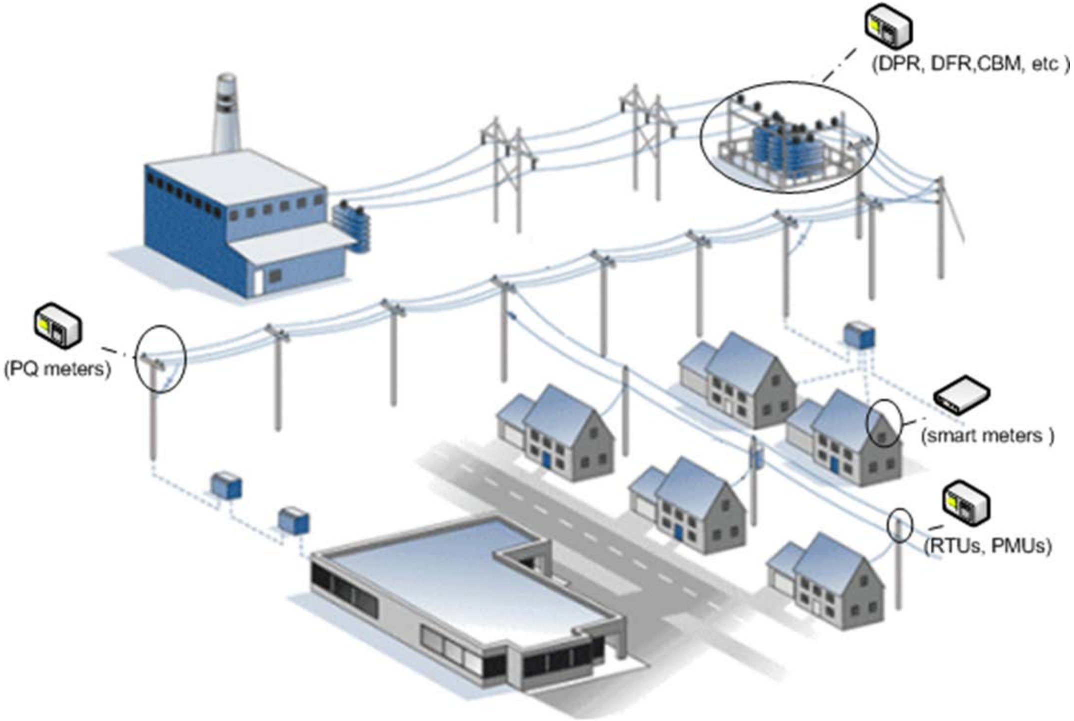




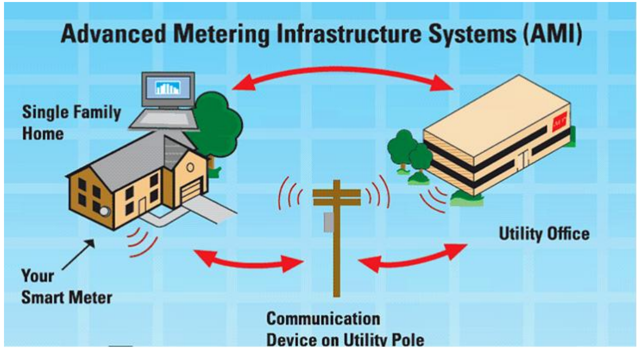
TEXAS A&M UNIVERSITY

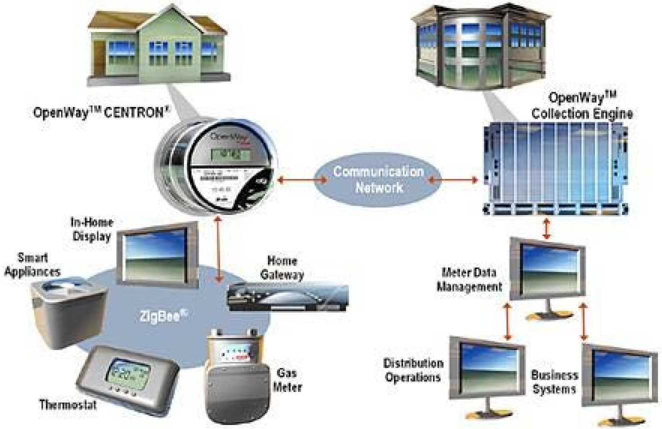
TEES SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

Feeder Data

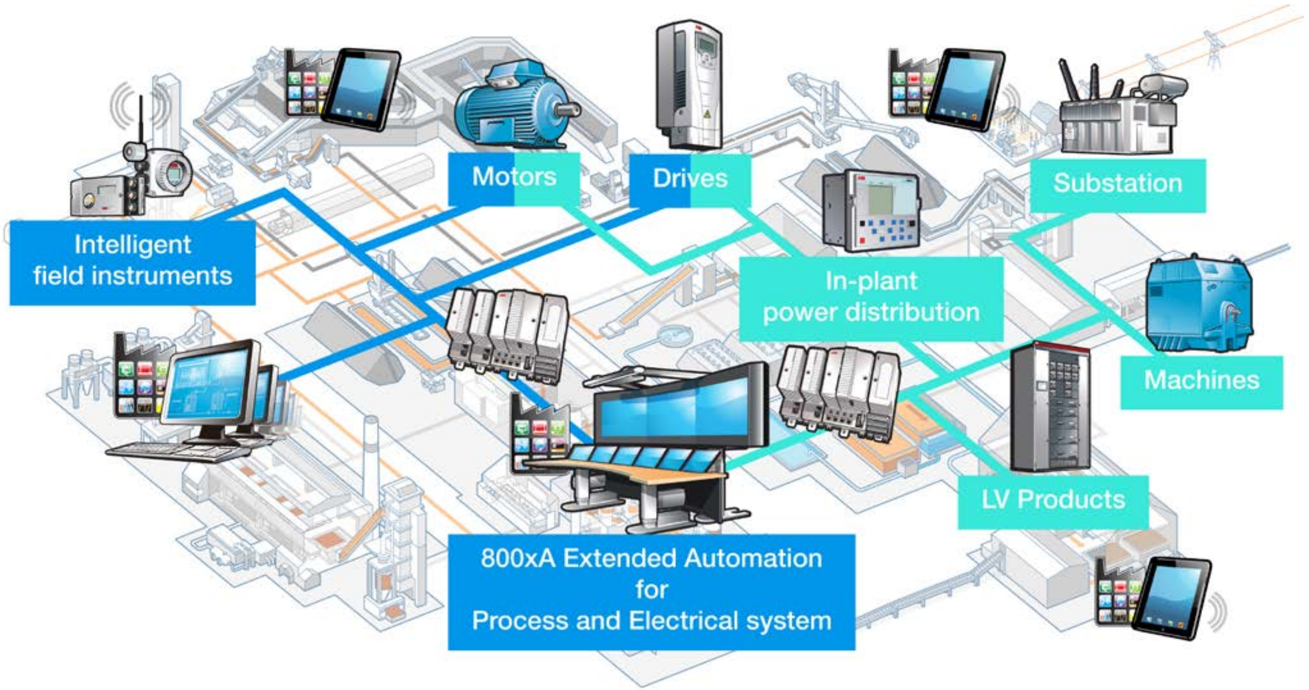


Smart Meter Data



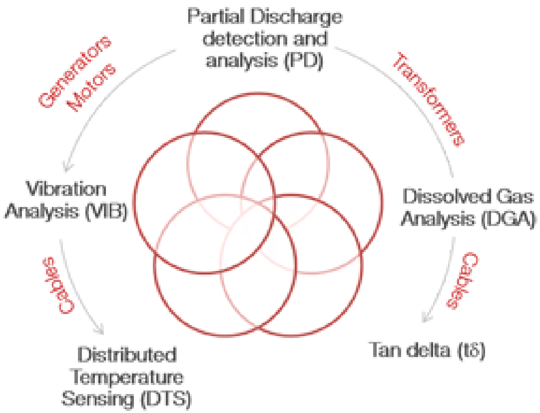


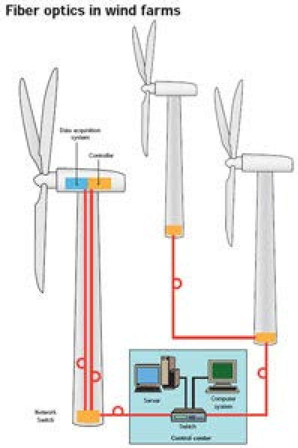
Industrial Plant Data

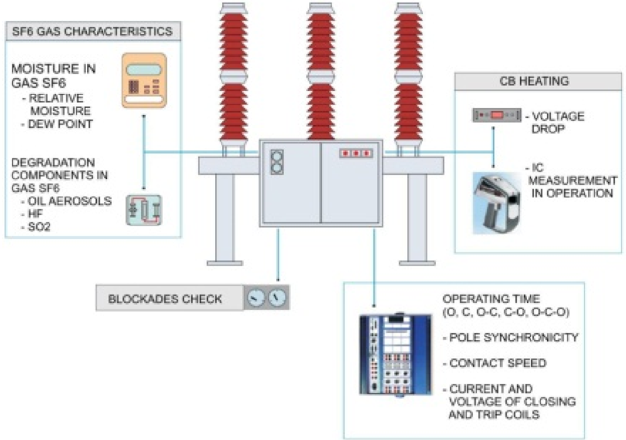


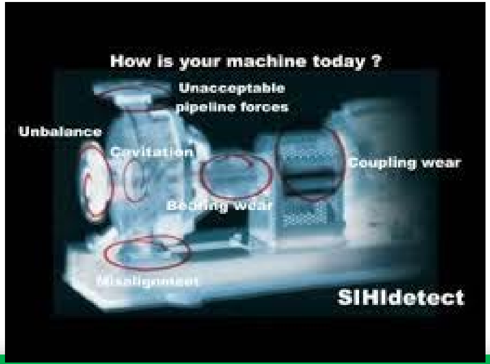
Renewables Data

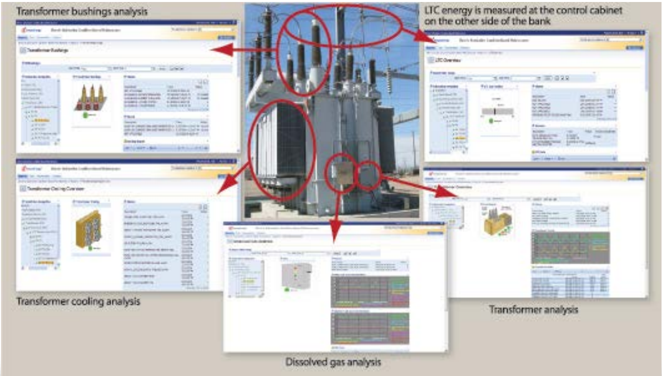
Asset Condition Data



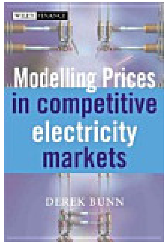




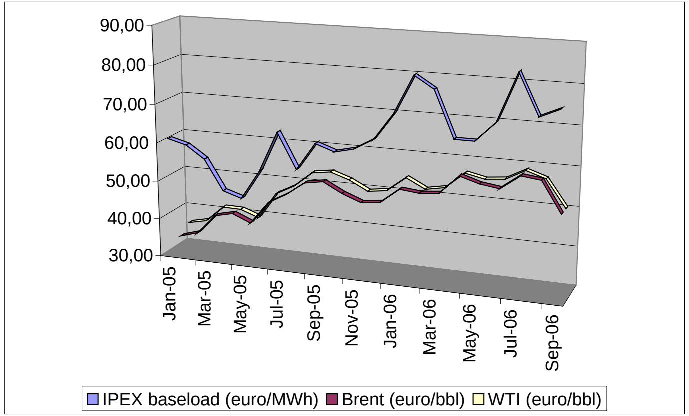




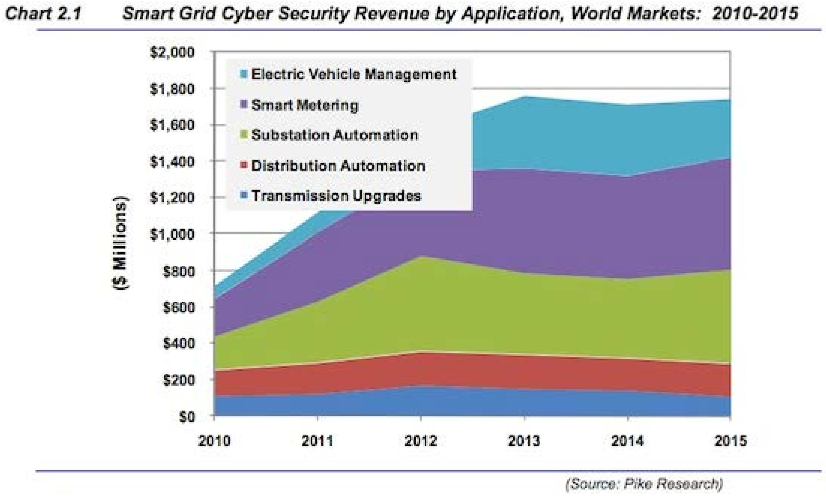
Market Data



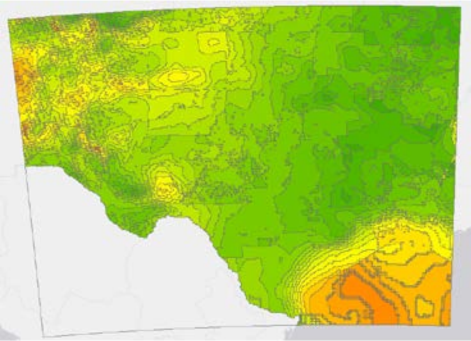
Oil and Electricity Monthly Average Price



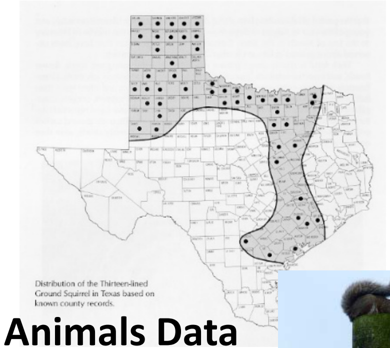
Cyber Security Data



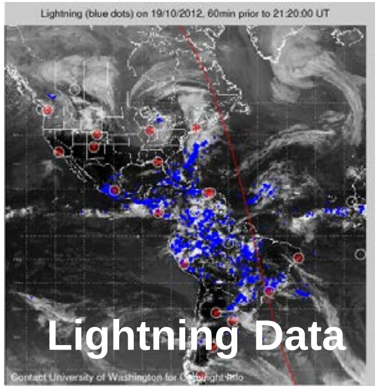
Data from other sources



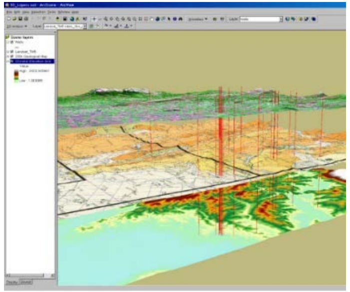
Weather Forecast



Animals Data

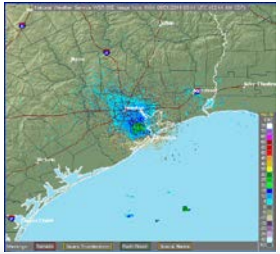


Lightning Data

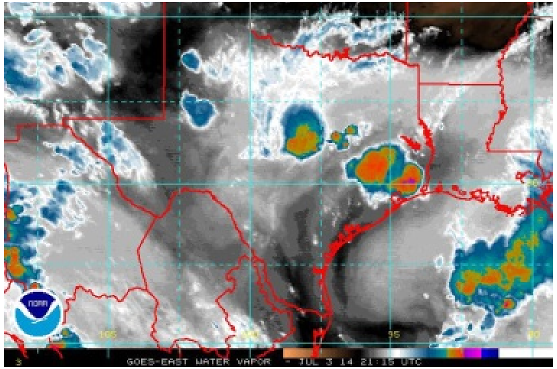


GIS

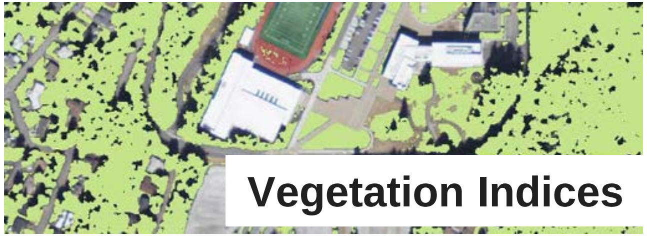
UAS



Radar data



Satellite data



Vegetation Indices

TEXAS A&M UNIVERSITY

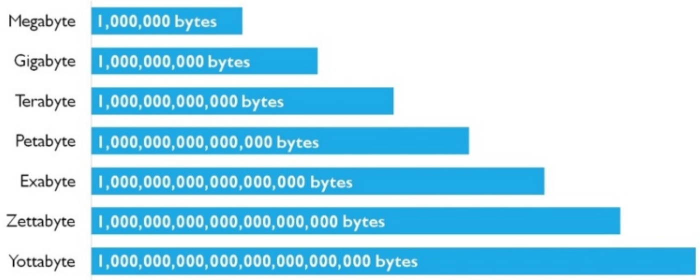
SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

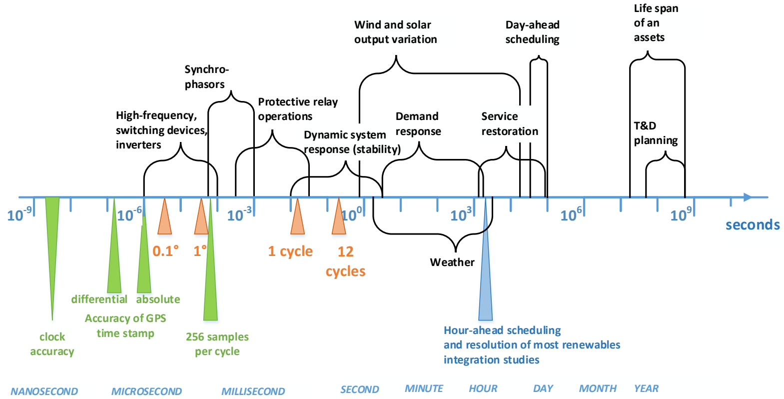
Outline

Big Challenges

- Big Data Properties
- Expectations

Big Data Properties

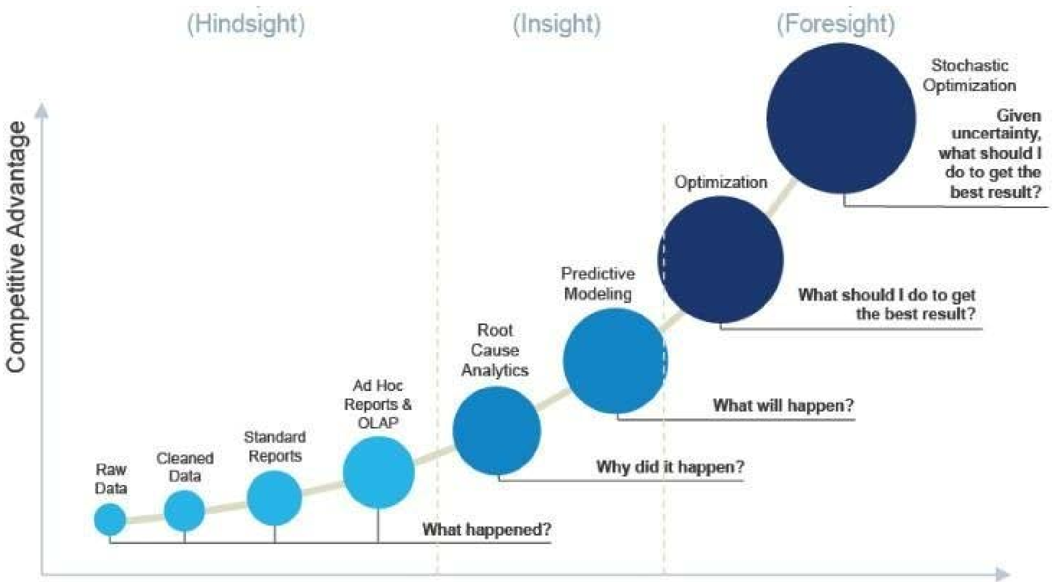




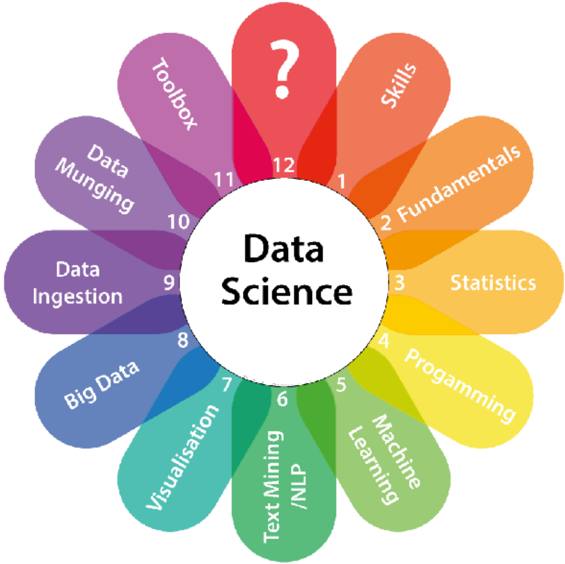
TEXAS A&M UNIVERSITY®

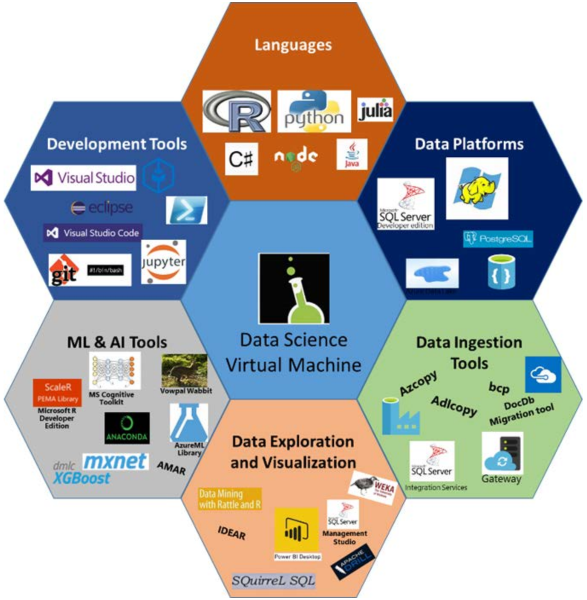
TEES | SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

Expectations



Data Science& Processing Infrastructure

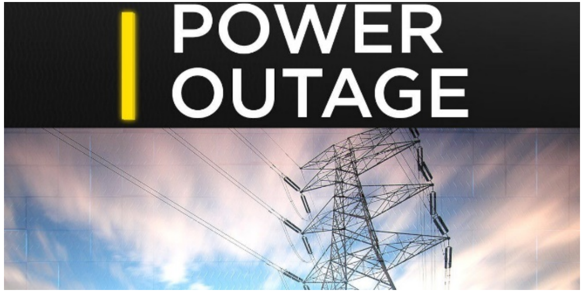




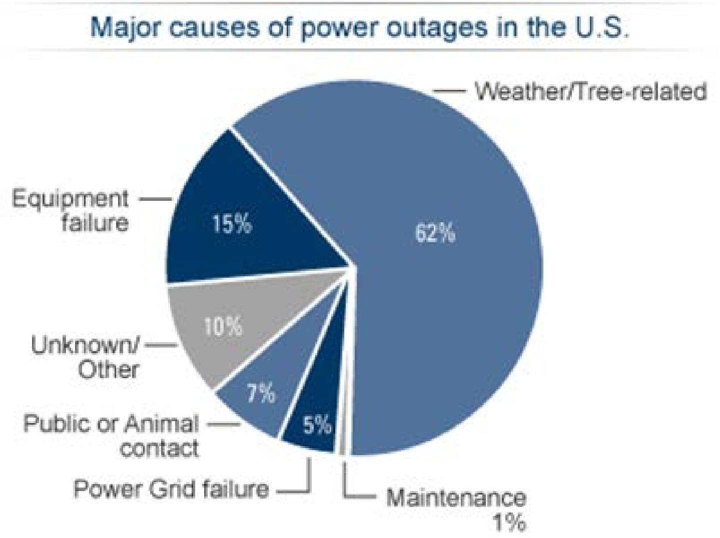
Outline

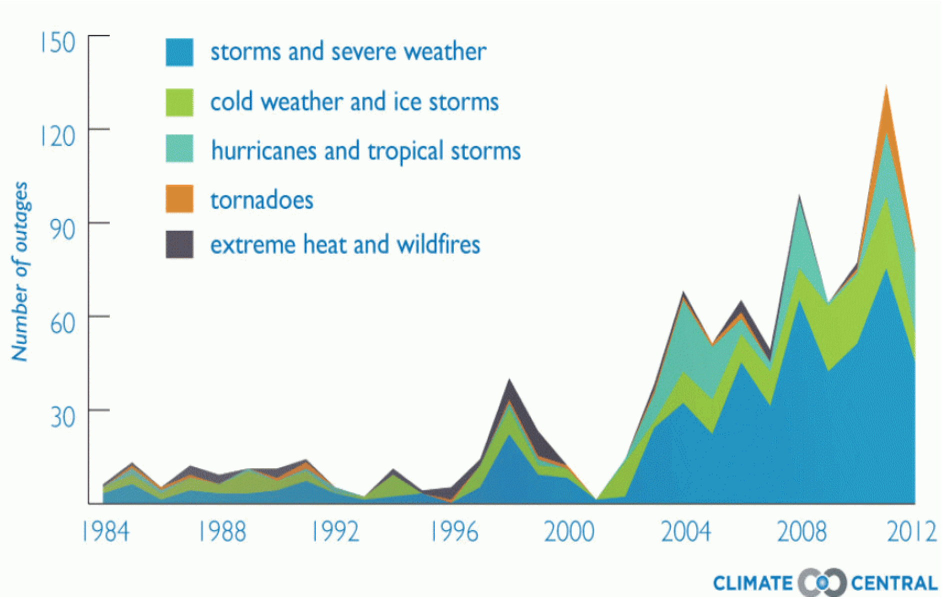
Example: Predicting outages/failures

- Transmission
- Distribution



Cause of outages



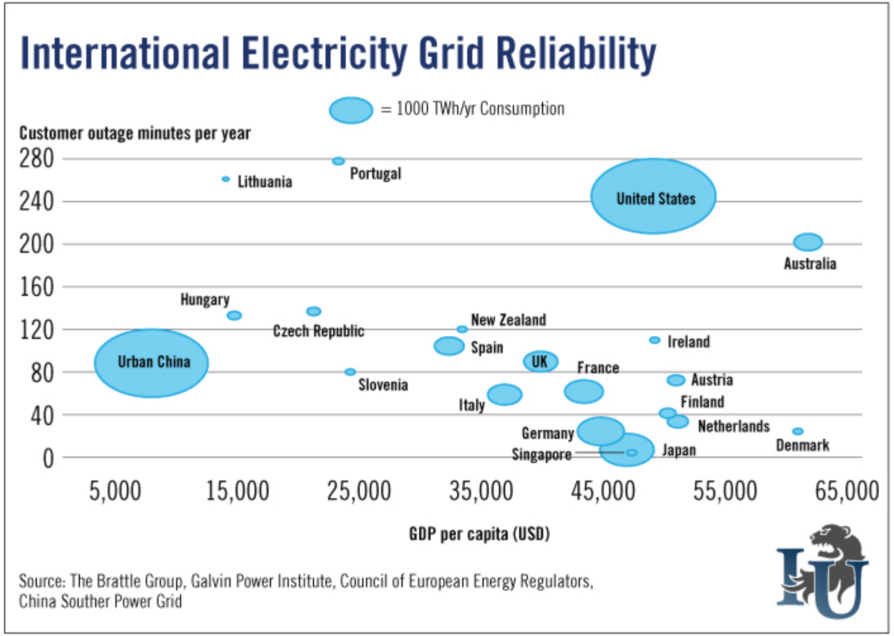


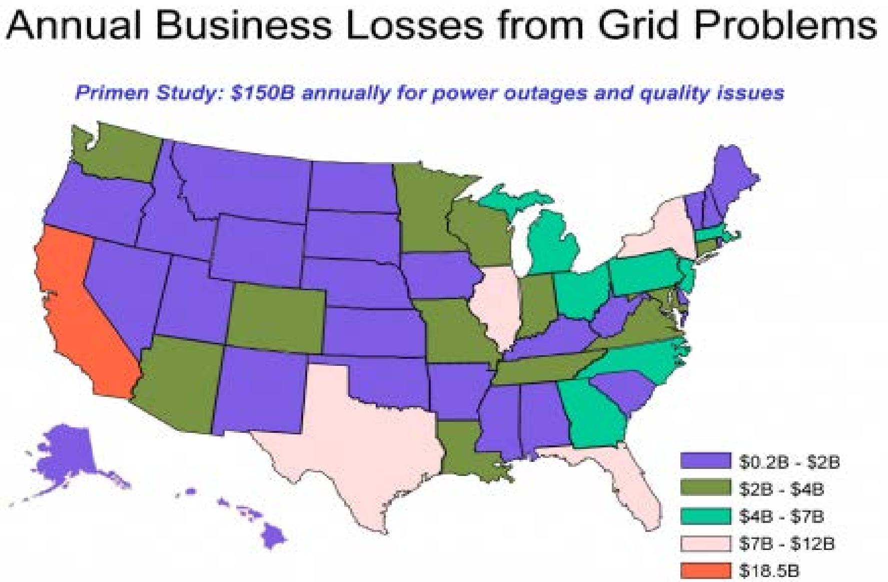
TEXAS A&M UNIVERSITY®

TEES

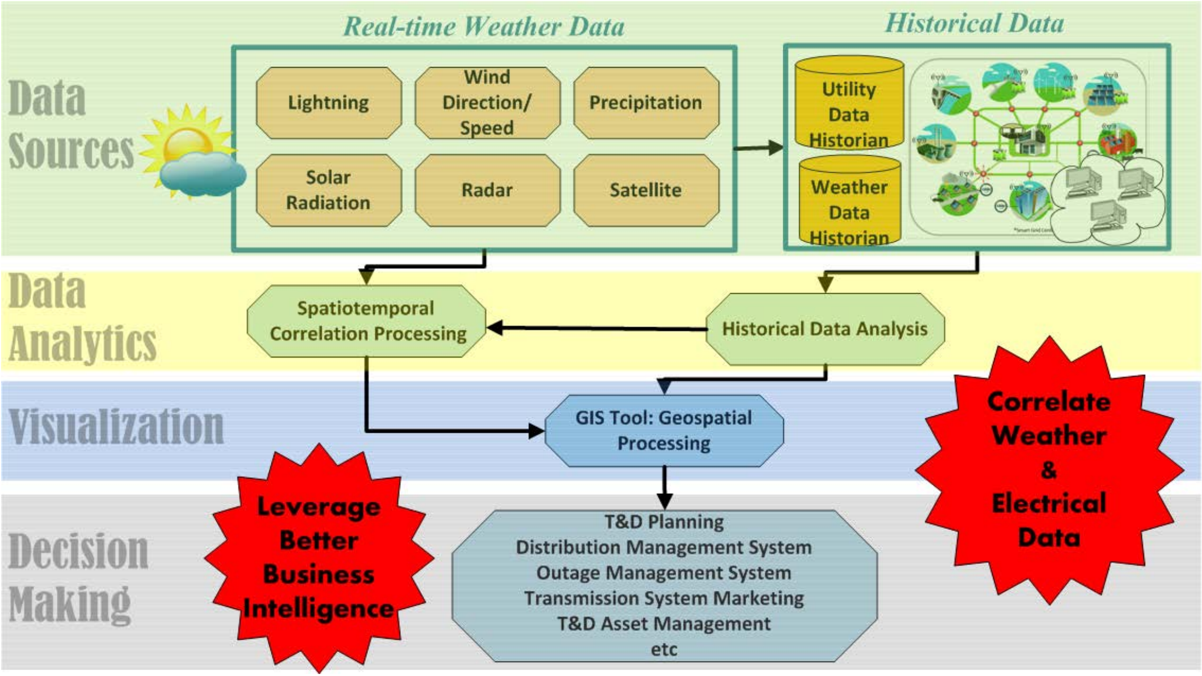
SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

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
Denmark
Singapore
Japan
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


Annual Business Losses from Grid Problems
Primen Study: $150B annually for power outages and quality issues
$0.2B - $2B
$2B - $4B
$4B - $7B
$7B - $12B
$18.5B

BD Framework



BD Data Properties

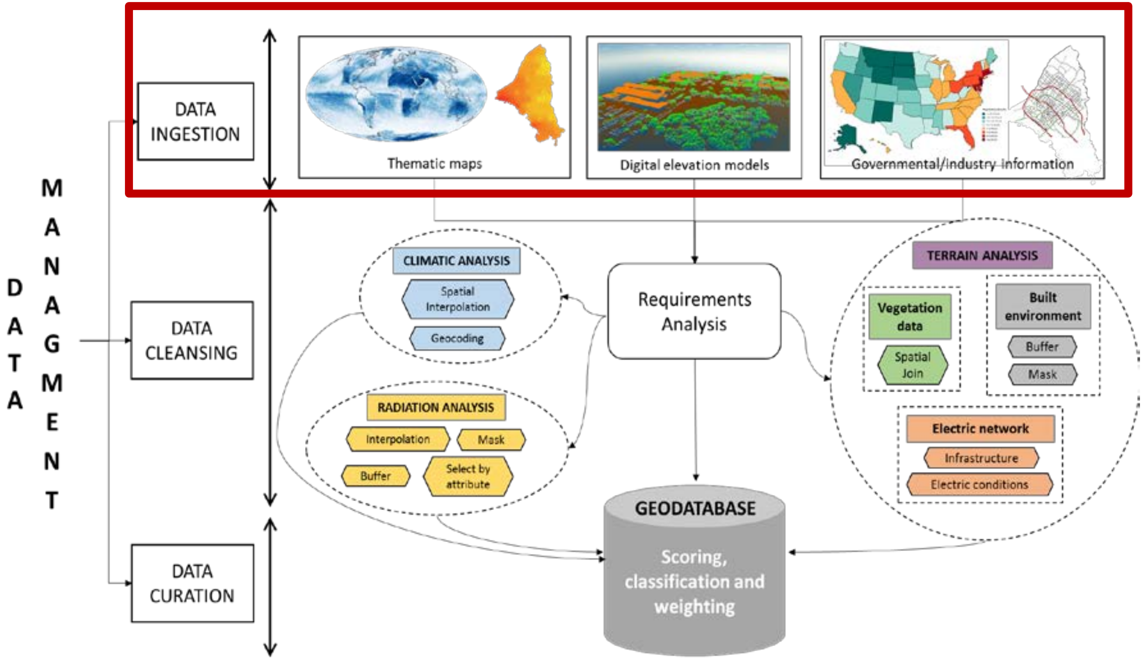
			VELOCITY	VOLUME	
	Source	Data Type	Temporal Resolution	Spatial Resolution	Measurements
VARIETY	Automated Surface Observing System	Land-Based	1 min	900 stations	Air Temperature, Dew Point, Relative Humidity, Wind Direction, Speed and Gust, Altimeter, Sea Level Pressure, Precipitation, Visibility…
	Level-2 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)
VARIETY	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

The Processing Steps

- Preprocessing: extract the data for the full graph of the network, and provide precise location of outages
- Spatiotemporal Correlation: correlate every network component with weather parameters
- Prediction Algorithm: graph based – combination of GCRF and logistic regression

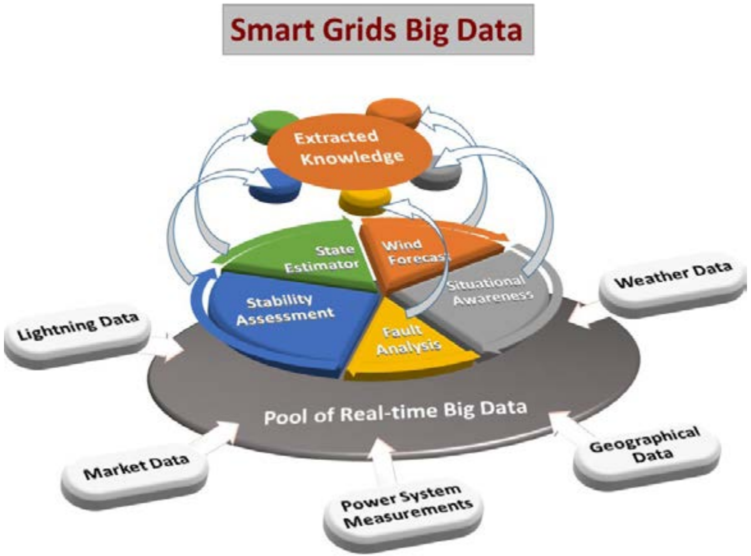
Data Integration

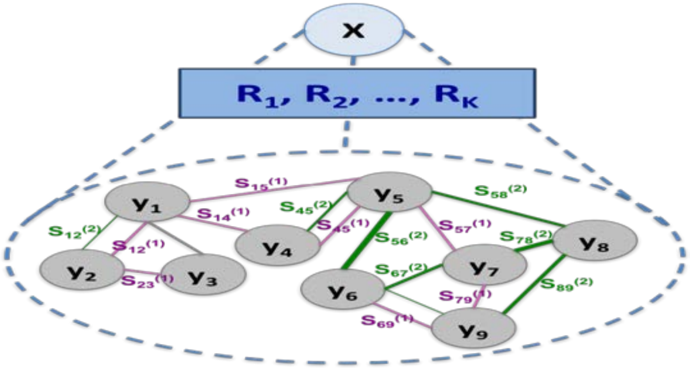


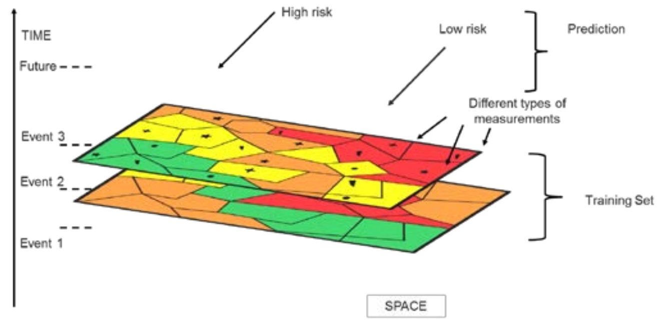
TEXAS A&M UNIVERSITY®

TEES | SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

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.

Weather Driven Risk Analysis

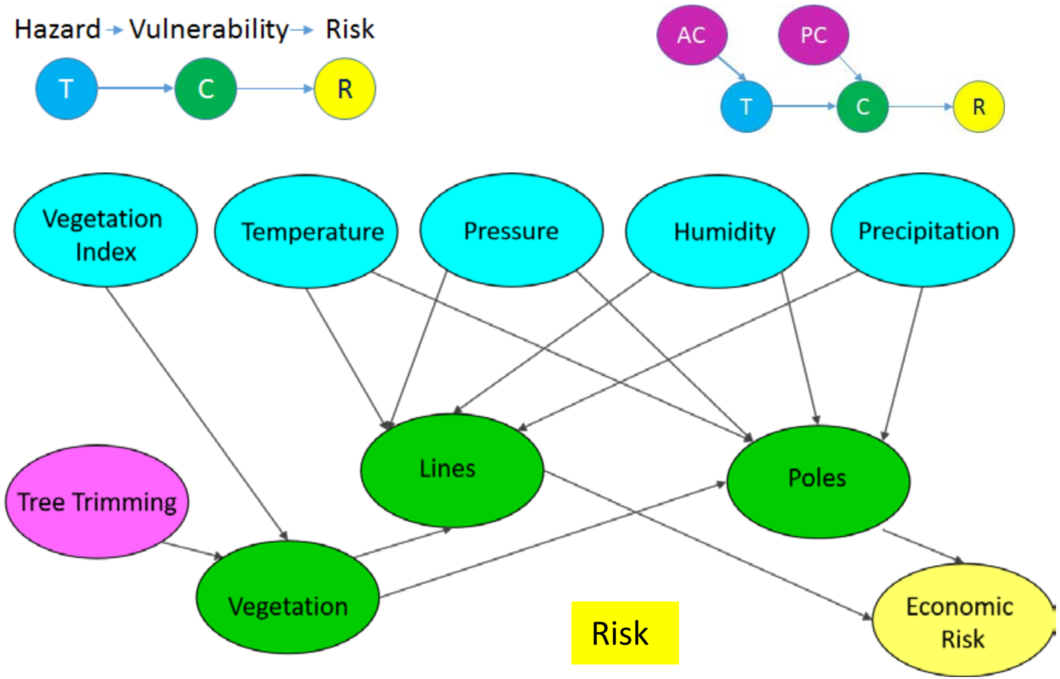
Risk = Hazard x Vulnerability x Economic Impact

- Probability of hazardous weather conditions
- Depends on *Weather Forecast*
- Pick a moment in time (or a period of time) and estimate probability of hazardous conditions

- Probability that hazardous conditions will cause an event in the network
- Depends on *Historical Weather and Outage Data*
- Learn from the historical data what may happen if hazardous conditions occur

- Expected economic impact in case of an event
- Depends on the type of economic loss that the user wants to consider
- Identify type of economic loss that is of interest for the study and calculate it

Example : Vegetation Risk Model

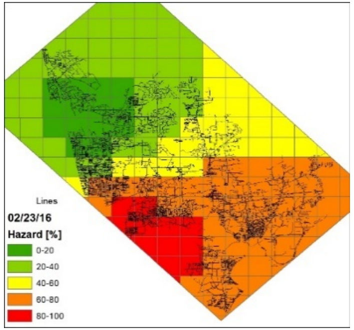


P. C. Chen and M. Kezunovic, "Fuzzy Logic Approach to Predictive Risk Analysis in Distribution Outage Management", *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2827-2836, November 2016.

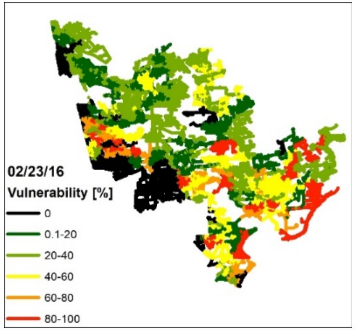
T. Dokic, P.-C. Chen, M. Kezunovic, "Risk Analysis for Assessment of Vegetation Impact on Outages in Electric Power Systems", CIGRE US National Committee 2016 Grid of the Future Symposium, Philadelphia, PA, October-November 2016.

Results – Risk Maps

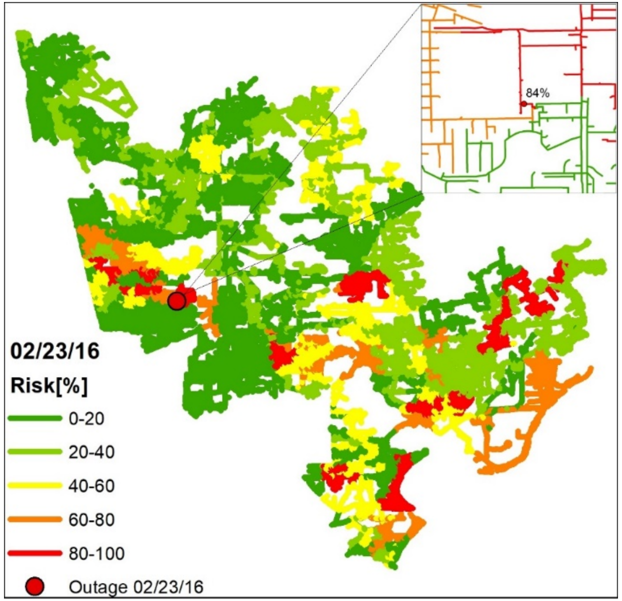
Weather
Hazard



Network
Vulnerability

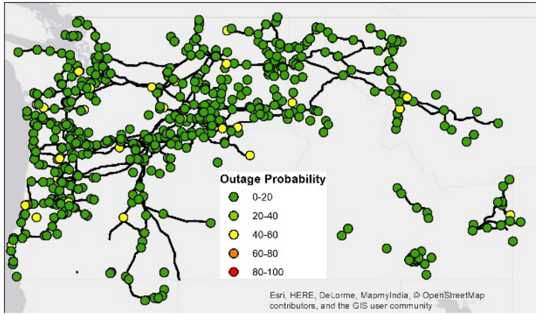


Risk Map

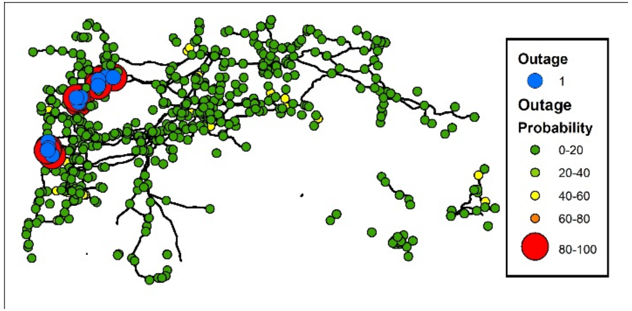


BD Analytics Outcomes

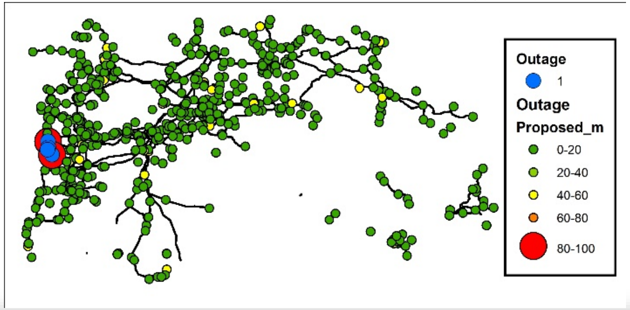
Probabilities of outages for no outage



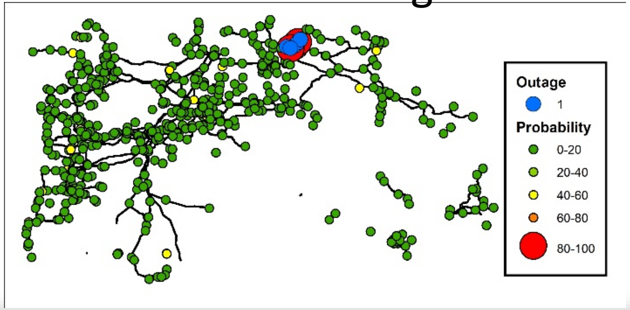
Probabilities of outages for lightning



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.)

- **Lessons learned:**

  - Assessment of risk is not meaningful without clear mitigation steps (design, component health, operating steps)

  - Big Data Predictive Analytics is cost effective and feasible if Big Data is readily available

  - Acceptance of Big Data Analytics depends on whether it is able to solve problems that otherwise are not solved

  - The target need to be great challenges with high returns if solved to justify the cost of implementation

  - The solutions are not necessarily intuitive, so extensive training and mind set change may be needed

Publications

T. Dokic, M. Kezunovic, “Predictive Risk Management for Dynamic Tree Trimming Scheduling for Distribution Networks,” IEEE Transactions on Smart Grid, (Accepted, In print).

T. Dokic, M. Pavlovski, Dj. Gligorijevic, M. Kezunovic, Z. Obradovic, "Spatially Aware Ensemble-Based Learning to Predict Weather-Related Outages in Transmission," The Hawaii International Conference on System Sciences - HICSS, Maui, Hawaii, January 2019.

M. Kezunovic, T. Dokic, R. Said, “Optimal Placement of Line Surge Arresters based on Predictive Risk Framework using Spatio-Temporally Correlated Big Data,” CIGRE Paris, August 2018.

M. Kezunovic, T. Dokic, "Predictive Asset Management Under Weather Impacts Using Big Data, Spatiotemporal Data Analytics and Risk Based Decision-Making," 10th Bulk Power Systems Dynamics and Control Symposium – IREP’2017, Espinho, Portugal, August 2017.

http://smartgridcenter.tamu.edu/resume/long_resume/Html/index.html#publ

Together -
building a prosperous future

where energy is

clean, abundant, reliable, safe, secure and affordable

TEXAS A&M UNIVERSITY®

TEES

SMART GRID CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION