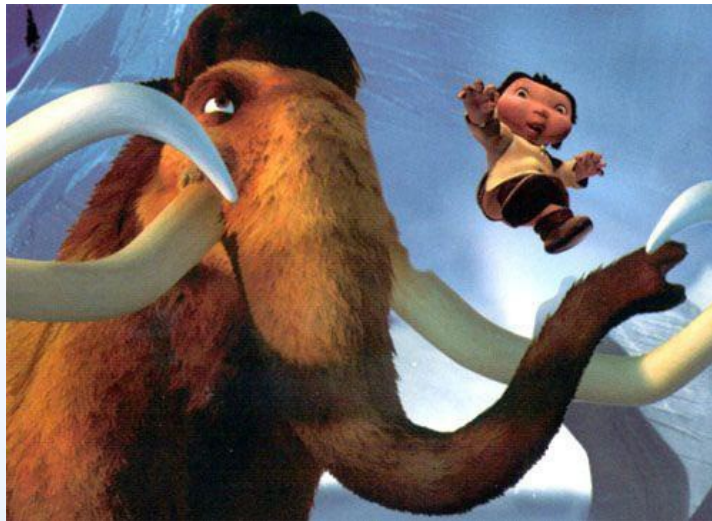


US EPA's Next Generation Compliance Program – Are We Ready for Big Data?

Adrienne Nemura
Geosyntec Consultants

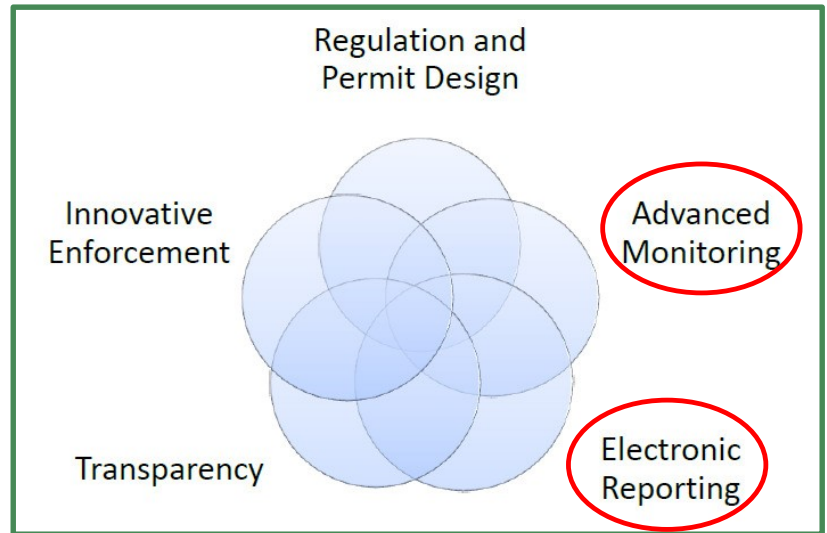
Ohio WEA Government Affairs Workshop
March 17, 2015

- ☐ What is NextGen?
- ☐ What is Big Data?
- ☐ What are Some Practical and Legal Considerations?
- ☐ Do We Still Have Issues with Small Data?



US EPA's Next Generation Compliance Program

- ❑ Led by US EPA/OECA building on US EPA's Clean Water Act Action Plan
- ❑ Allows US EPA to write regulations and permits with “built in” compliance (save on gov't resources!)
- ❑ Use existing and emerging monitoring and information technologies



OECA = Office of Enforcement and Compliance Assurance

<https://www.epa.gov/compliance/next-generation-compliance-strategic-plan-2014-2017> (Oct. 2014)

- Targeted enforcement plan to address “challenges” from
 - Largest dischargers
 - Confined animal feeding operations (CAFOs)
 - Sewer overflows
 - Contaminated water that flows from industrial facilities
 - Construction sites
 - Runoff from urban streets

Plan Goals

- Target enforcement to most significant pollution problems
- Improve transparency and accountability by providing better public access to data
- Strengthen enforcement performance

<https://www.epa.gov/compliance/about-clean-water-act-cwa-action-plan> (Oct. 2009)

Implementation of the CWA Action Plan

- Switch to e-reporting with automated compliance evaluations and improved transparency
- New paradigm to compel compliance via public accountability, self-monitoring, e-reporting...
- Re-tool key NPDES permitting and enforcement practices
- Coordinate permitting, compliance, and enforcement to improve agency performance in protecting and improving water quality

<https://www.epa.gov/compliance/clean-water-act-cwa-action-plan-implementation-priorities-changes-improve-water-quality>
(May 2011)

NOUN

Computing

Extremely large data sets that may be analyzed **computationally** to reveal patterns, trends, and associations, especially relating to human behavior and **interactions**:

*‘much IT investment is going towards **managing** and maintaining big data’*

http://www.oxforddictionaries.com/us/definition/american_english/big-data


Definition of BIG DATA

Popularity: Bottom 30% of words

: an accumulation of data that is too large and complex for processing by traditional database management tools

<http://www.merriam-webster.com/dictionary/big%20data>

Rohan Deuskar

CEO and Co-Founder, [Stylitics](#) 

Twitter: [@RohanD](#) 

“ **Big data refers to the approach to data of “collect now, sort out later”...meaning you capture and store data on a very large volume of actions and transactions of different types, on a continuous basis, in order to make sense of it later.** The low cost of storage and better methods of analysis mean that you generally don’t need to have a specific purpose for the data in mind before you collect it. ”

Deirdre Mulligan

Associate Professor, [UC Berkeley School of Information](#) 

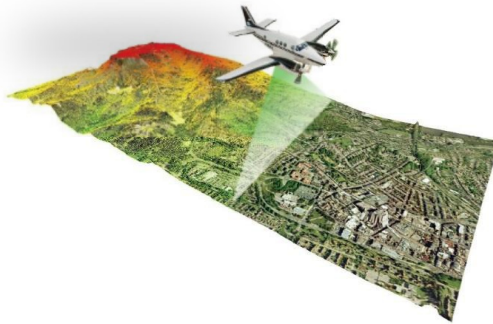
“ Big data: **Endless possibilities or cradle-to-grave shackles, depending upon the political, ethical, and legal choices we make.** ”



<https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know> Bernard Marr, Mar. 6, 2014

BIG DATA: VOLUME

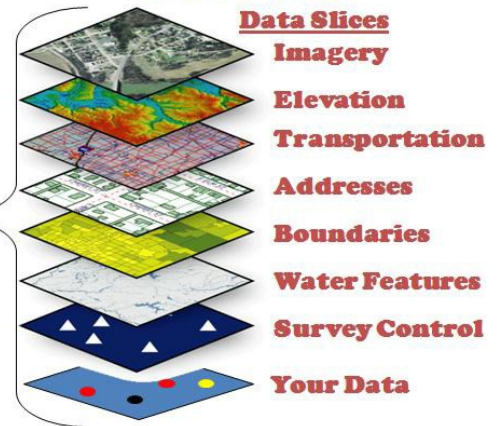
Vast amount of data generated every second. Use software to bring distributed data together.



The Real World



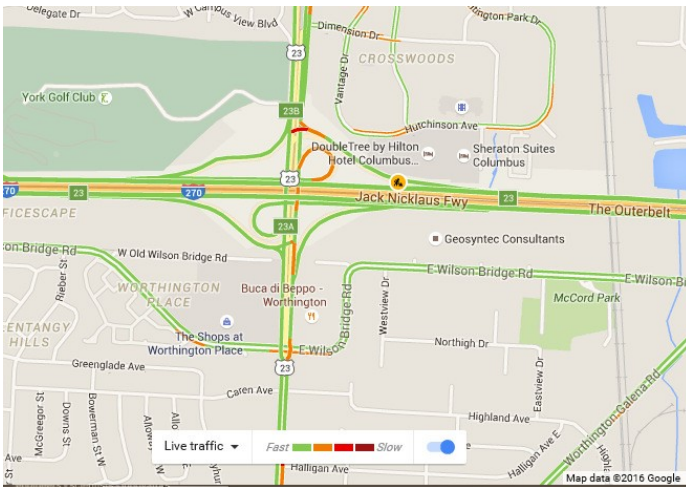
GIS World Model



BIG DATA: VELOCITY

Speed at which new data is generated and is “pushed out” to interested parties.

PROCESS pH MEASUREMENT PRODUCT GUIDE



BIG DATA: VARIETY

Harnessing different types of data
(structured and unstructured)
including photos, sensor data,
videos, etc. and bring them together
with more traditional, structured data.

VIDEO: Daredevil Swimming in River Des Peres Seen By EPA as Justification for Cleanup

Posted By [Chad Garrison](#) on Wed, Aug 31, 2011 at 11:41 am



YouTube

Chris Kline picks his bum after a swim in the
River Des Peres in 2008.



BIG DATA: VERACITY

Messiness or trustworthiness of the data. Volume is often used to make up for the lack of quality or accuracy.

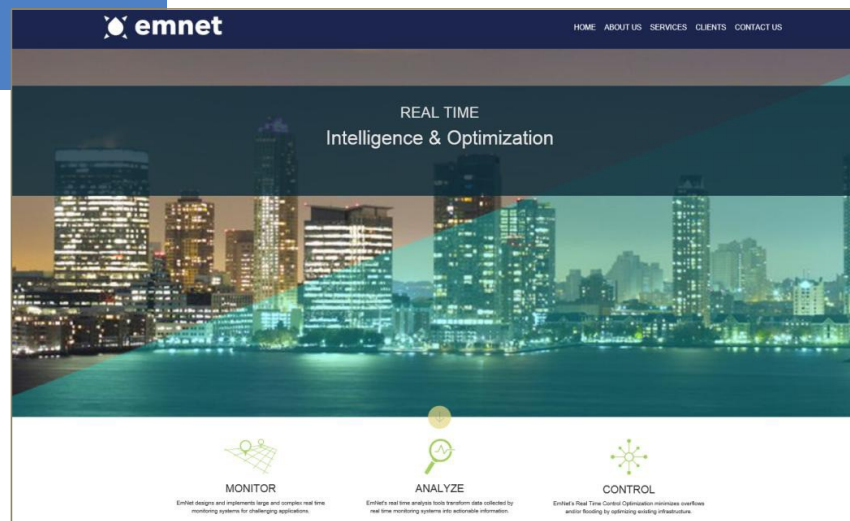
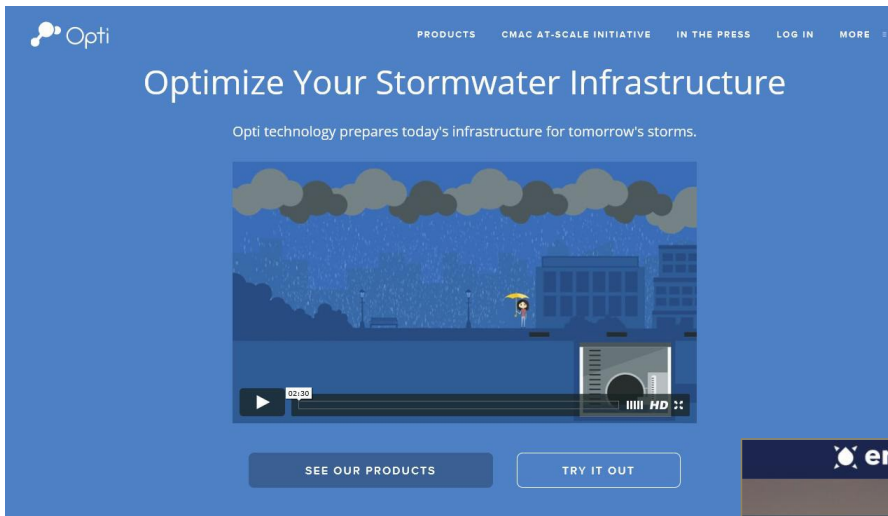


1,127,857
Comments Received*



BIG DATA: VALUE

Can a business case be made for collecting and leveraging big data?



- Identify ways to help communities save money (real-time controls, energy savings, right-sizing maintenance programs)
- Reduce conservatism in calculations of NPDES limits and total maximum daily loads (TMDLs)
- Allow states to have more data to make actual assessments about attainment of uses
- Avoid regulation of “imaginary pollutant events”
- Integrate treatment plant operations with actual river conditions
- Focus management efforts on restoring designated uses



Bruce Roll, Clean Water Service, NACWA Water Quality Committee, Feb. 2, 2015

- Who will pay to collect big data?
- More data could mean more non-compliance or could show high frequency of compliance
 - More data could show that non-compliant points are really outliers
 - Are our water quality standards and permitting regulations calibrated to big data?
- More data could result in more or less stringent permit limits

- Real-time data and analytical tools rapidly evolving
- EPA's compliance and enforcement programs are trying to evolve
 - What if in compliance 99.9% of time but still have "1,000s of violations"?
 - Will permitting keep up?
 - Will water quality standards keep up?
- Can legal programs keep up with data management and reporting requirements for big data?
 - Chain-of-custody
 - Quality assurance / quality control
 - Access protocols


Do We Still Have Issues with Small Data?

Study the past, if you would
divine the future.

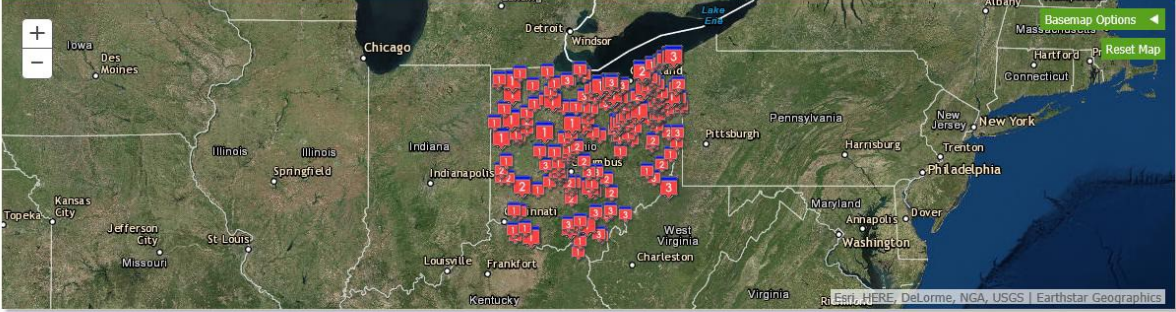
Confucius

You are here: [Home](#) » [Facilities](#) » [Facility Search](#) » Facility Search Results

Facility Search Results

 **New Jersey Clean Water Act data and some Clean Air Act data are frozen.** [Read More...](#)

[Hide Map](#) [Hide Table](#) [Hide Summary](#) [Modify Search](#)



[Basemap Options](#) [Reset Map](#)

[Customize Columns](#) [Download CSV File](#) [Download Excel File](#)

[Results Guide](#) [Reports Legend](#)

Facility Name	NPDES ID	Mapped	Reports	Street Address	City	State	FRS ID	Current Significant Noncompliance (SNC) Status	Quarters in Noncompliance (3 yrs)	Effluent Violations (3 yrs)	Inspections (5 yrs)	Formal Enf. Actions (5 yrs)
---------------	----------	--------	---------	----------------	------	-------	--------	--	-----------------------------------	-----------------------------	---------------------	-----------------------------

Search Statistics

193 Search Results
193 Facilities with a Current Violation/s
193 Facilities with Violations in the Last Three Years
15 Facilities with Formal Enforcement Actions in the Last Five Years
193 Facilities with Inspections in the Last Five Years

Facility Summary

Search Criteria
Search Type: Water
State(s): OH - Ohio
Permit Status: Effective; Expired; Administratively Continued; Pending; Retired
Permit Type: NPD - NPDES Individual Permit; NGP - NPDES Master General Permit; GPC - General Permit Covered Facility; UFT - Unpermitted Facility
Current Compliance Status: Effluent Violations - Significant Violations
Last Inspection: Within 3 year(s)
[Modify Search](#)

179 CWA Facilities

84 w/ 12 quarters of non-compliance

EPA DMR Pollutant Loading Tool Using the “EZ Search”

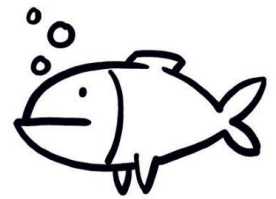
NPDES ID	Facility Name	Top Pollutant	Top Pollutant Pounds (lbs/yr)	Average Daily Flow (MGD)
		Solids, total dissolved	28,288,992	645
		Solids, total dissolved	836,953	295
		Residue, tot fltrble (dried at 105 C)	7,360,971	184
		Solids, total dissolved	196,262,877	142
		Solids, total dissolved	290,016,124	132
		Residue, tot fltrble (dried at 105 C)	217,216,793	103
		Residue, tot fltrble (dried at 105 C)	134,595,524	92
		Solids, total dissolved	174,143,226	78
		Solids, total dissolved	60,494,524	38
		Solids, total dissolved	17,097,414	31
		Solids, total dissolved	30,114,320	30

https://cfpub.epa.gov/dmr/ez_search_results.cfm

EPA DMR Pollutant Loading Tool Using the “EZ Search”

NPDES ID	Facility Name	Top Pollutant	Top Pollutant Pounds (lbs/yr)	Average Daily Flow (MGD)
OH0027481	SPRINGFIELD WWTP	Solids, total dissolved	28,288,992	645
OH0025852	IRONTON WWTP & SEWER SYSTEM	Solids, total dissolved	836,953	295
OH0024899	DEFIANCE WATER POLLUTION CONTROL PLANT	Residue, tot fltrble (dried at 105 C)	7,360,971	184
OH0024643	NE OHIO REGIONAL S D EASTERLY STP	Solids, total dissolved	196,262,877	142
OH0024651	SOUTHERLY WASTEWATER TREATMENT CENTER	Solids, total dissolved	290,016,124	132
OH0024741	SOUTHERLY WASTEWATER TREATMENT PLANT	Residue, tot fltrble (dried at 105 C)	217,216,793	103
OH0024732	JACKSON PIKE WASTEWATER TREATMENT PLANT	Residue, tot fltrble (dried at 105 C)	134,595,524	92
OH0027740	CITY OF TOLEDO DIVISION OF WATER RECLAMATION	Solids, total dissolved	174,143,226	78
OH0024660	WESTERLY ADV WWTP	Solids, total dissolved	60,494,524	38
OH0026328	MANSFIELD WWTP	Solids, total dissolved	17,097,414	31
OH0031062	EUCLID WWTP	Solids, total dissolved	30,114,320	30

- Springfield is 4.5X larger than Southerly! (but Springfield's design flow is 25 MGD)
- Ironton's flow is 10X Euclid, but Euclid discharges 35X more TDS than Ironton!



https://cfpub.epa.gov/dmr/ez_search_results.cfm

Uh oh



□ Adrienne Nemura, P.E.
Geosyntec Consultants
Ann Arbor, MI
anemura@Geosyntec.com
(734) 476-0357

