

ERDC/ITL SR-20-6

Information Technology Laboratory



**US Army Corps
of Engineers®**
Engineer Research and
Development Center



USACE Data Strategy Initiative

Insights: An Update on the USACE Data Strategy Initiative

April 2019 Edition

Cary D. Butler, Martin C. Kittrell, and Robert G. Dixon

August 2020

The U.S. Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdclibrary.on.worldcat.org/discovery.

To search for other technical reports published by ERDC, visit the ERDC online library at <http://www.erdclibrary.on.worldcat.org/discovery>.

Insights: An Update on the USACE Data Strategy Initiative

April 2019 Edition

Cary D. Butler (Retired) and Martin C. Kittrell
Information Technology Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Robert G. Dixon
Commander and District Engineer (Retired)
Little Rock District (SWL), Southwestern Division (SWD)
P.O. Box 867
Little Rock, AR 72203

Final report

Approved for public release; distribution is unlimited.

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

Under USACE Data Strategy Initiative, Work Unit Number B72BKG

Abstract

The U.S. Army Corps of Engineering (USACE) Commanding General, LTG Todd T. Semonite, announced 10 initiatives on May 2018 to improve execution through informed decision-making, enabling lower costs, and world-class engineering results today and tomorrow. Two of the initiatives were geared to transition USACE to data-informed decision-making through the use of data analytics.

The first of these two initiatives sought to implement a data strategy that included a doctrine and governance, through creation of a data management plan. The plan implemented tools to aggregate data across the organization and improve reporting. Dr. Cary Butler, USACE Chief Data Scientist, led this effort.

The second initiative sought to establish a dedicated USACE Innovations Team to build and recruit a skilled team to act as finders and enablers of innovative, enterprise solutions that enabled USACE to become a digital business. The USACE Chief Information Officer oversaw both initiatives.

Because goals and objectives overlapped in many areas and gathered momentum for these initiatives, leaders from both teams came together and decided that showing the business value of data analytics to improve decision making across the Corps should be a top priority.

This document was used to report on the progress of the initiatives and give a better understanding of the need for data-informed decisions throughout USACE.

DISCLAIMER: The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products. All product names and trademarks cited are the property of their respective owners. The findings of this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

Contents

Abstract.....	ii
Preface	iv
1 Introduction	1
1.1 Background.....	1
1.2 Objective and approach	1
2 Establish Data as a Strategic Asset	2
2.1 Leadership buy-in	2
2.2 Workforce development	3
2.3 USACE data ecosystem	4
2.4 USACE data governance.....	5
2.5 Demonstrate to motivate	6
3 Conclusion	8
Appendix: Insights Magazine April 2019 Issue.....	9
Report Documentation Page	

Preface

This study was conducted for the USACE Office of the Chief Information Officer under Project USACE Data Analytics, Work Unit Number B72BKG. The technical monitor was Dr. Cary Butler.

The work was performed by the Office of Technical Directors, U.S. Army Engineer Research and Development Center, Information Technology Laboratory (ERDC-ITL). At the time of publication, Dr. Jerrell R. Ballard, Jr. was Chief, Computational Science and Engineering Division; Mr. Quincy G. Alexander was Acting Chief, Software Engineering and Informatics Division, and Dr. Robert M. Wallace was the Technical Director for Data Analytics. The Deputy Director of ERDC-ITL was Ms. Patti S. Duett and the Director was Dr. David A. Horner

COL Teresa A. Schlosser was the Commander of ERDC, and Dr. David W. Pittman was the Director.

1 Introduction

1.1 Background

With the data revolution under way, the opportunity arose to revolutionize the speed and accuracy of daily decisions within the U.S. Army Corps of Engineers (USACE). Accurate information will be needed that is available at the speed of relevance or opportunities could be missed for timely decisions or uninformed decisions could be made. This document discusses up-to-date enterprise efforts to turn USACE's historic and dynamic data into one of its biggest strategic assets, which is already being accomplished by the Army and other successful global businesses.

LTG Semonite stated in Semonote #18 (<https://team.usace.army.mil/sites/HQ-CL/PDT/USACEDA/default.aspx>) "Our data strategy enables informed decisions at all levels; absolutely essential for the Revolution!!" To assist in this data strategy, a Data Governance Board was created. This board consists of technical and senior leader experts where their goal is to create and support a data-driven environment that consistently and accurately delivers authoritative, timely, relevant, and secure data to decision makers.

1.2 Objective and approach

The data strategy for USACE to implement and understand the importance of data is:

- leaders, at all levels, should understand the importance and value of their data;
- the workforce should become data-savvy and able to draw insights on their own from data;
- centralized repositories should be created that would be easily accessible, securable, and easy to navigate;
- quality data that can be trusted; and
- demonstrations should be conducted to show the value of data.

2 Establish Data as a Strategic Asset

USACE leaders are seeing the value in data that was just considered a byproduct of missions. This data should now be used as a strategic asset; used to harvest insights to improve decision making. In turn, this will allow USACE to remain a world-class organization.

Since some business areas will handle the transition at a faster pace, this will take some time. To help in this effort, a Data Strategy Initiative has been put in place, which includes five lines of effort (LoE). These efforts are explained in detail in Sections 2.1 through 2.5. Army agencies already making similar transitions will be crucial in providing feedback across organizations in order to refine and update the approach.

2.1 Leadership buy-in

Leaders at all levels must learn to trust their data, instill data discipline, and educate the workforce to understand and value what data can show them and their missions. Senior leadership at USACE has created a Data Governance Board (DGB) to define clear goals.

The four objectives for this effort are:

1. **Support Executive Governance Meeting (EGM) activities:** Leaders regularly gather for professional development and to hold discussions about issues throughout the organization. Feedback on the Data Strategy Initiative will be also given during this activity.
2. **Support DGB activities:** The USACE DGB is the senior recommending directive body for development of coordinated USACE enterprise positions on data strategy, data standards, data quality, and data interoperability. This board serves as the senior adjudication body for the USACE Data Strategy Initiative; maintains the inventory of USACE authoritative data sources; serves as a certification/waiver approval authority for targeted data standards as delegated by the Deputy Commanding General (DCG); and collects and disseminates best practices and lessons learned for the data community. Within their charter, priorities and deliverables are listed that will be addressed by the LoE.
3. **Promote leadership development:** Executive courses concerning data analytics will be implemented through the USACE Data Strategy

- Initiative. These courses will help USACE leaders gain insight into how industry is using data and applying data analytics within their organizations. Real-world examples will be given to demonstrate that treating data as a strategic asset could improve decision making.
4. **Sync with leadership:** As new leadership comes onboard, opportunities will be available concerning the value of data and the current state of effort of the USACE Data Strategy initiative. Sub-tasks to inform leadership of the state of the USACE Data Strategy Initiative include:
 - a. The District Commander Course (DCC) and the General Officer-Senior Executive Service (GO-SES) data analytics modules: Modules describing the value of USACE data will be created for the DCC and the GO-SES courses that will be given regularly by leadership. These modules will give an overview of the importance of data as a strategic USACE asset and establish the doctrinal framework for District operations. It will also go over specific tactics, techniques, and procedures for success.
 - b. Regular USACE Data Strategy Initiative reports: A digital magazine, *Insights*, will be regularly published to inform USACE on the status of the Data Strategy Initiative.
 - c. USACE Data Analytics web presence: A SharePoint site has been created for all relevant documents about USACE Data Analytics. The site can be found at <https://team.usace.army.mil/sites/HQ-CI/PDT/USACEDA/default.aspx>.

2.2 Workforce development

The demand for employees trained in data analytics has exploded since the realization that this field could bring real value to USACE. These employees should be skilled in data science and be accessible across the organization, which will require a transition in thinking. Data teams (data engineers, data scientists, and data analysts) should be seen as an investment rather than an expense.

The four objectives for this effort are:

1. **Position description:** To help with the process of making decisions, District Data Officers (DDO) were incorporated. The DDO will also act as an enterprise-wide resource for sharing data and help answer questions using data across organizational boundaries. Over time, this

- DDO community will become a key player in the USACE decision process. A Position Description (PD) has been written to assist Districts in hiring.
2. **PROSPECT training/certification:** Most DDOs will come onboard without data analytics training, so a PROSPECT curriculum is being developed to provide the foundational knowledge required to serve as a DDO. This will include a certification process that will recognize the individuals that complete all 3 courses as a Subject Matter Expert (SME) for USACE Data Analytics.
 3. **USACE Data Analytics Center of Expertise (DACX):** Primarily built to create a knowledge base of expertise in data and data analytics, the DACX would accelerate the cultural transformation to manage and use data in informed decision making. The DACX would also develop the expertise and methodologies required to assist Districts and Divisions as USACE data is transformed into a strategic asset for the organization, focus on making data analytics a repeatable process, and support the creation of data analytics capabilities down to the District level. It would also lead a community of practice across the enterprise focused on maturing DA capabilities and promote the sharing of methods and data essential for decision making. It would also serve as a clearinghouse for information capture and flow between the field, CIO, and ACE-IT. The objectives of the CX are: 1) support USACE leadership's goal to improve decision making; 2) provide the field with technical support, best practices, analyses, and the overall introduction to DA; 3) help reduce redundancy and promote the sharing of DA advancements; 4) ensure Divisions and Districts progress at a similar pace in becoming efficient with using data to make decisions; and 5) work with data owners to ensure consistency, easy access, and quality of data.
 4. **Data workforce of the future planning:** This plan will help guide USACE in hiring and training as data becomes an integral part of the organization.

2.3 USACE data ecosystem

This LoE supports the technology requirements of the Data Strategy Initiative. There is not a tool that exists right now that converts data to insights, so the immediate focus is being placed on investing in technology to capture, collect, organize, share, and analyze USACE data. Microsoft's Azure is providing the information technology (IT) platform for this data strategy.

The three objectives for this effort are:

1. **Data migration:** The process of selecting, preparing, extracting, and transforming data and having it permanently transferred from USACE's legacy system to a cloud environment.
2. **Data architecture:** A plan to establish models, policies, rules, and standards that will govern which data is collected, how it is stored, arranged, and used.
 - a. Enterprise data configuration management plan: Integrate and maintain Information Exchange Specifications and associated Controlled Vocabulary (Terms of Reference) and database requirements and fund them to comply. (Data Quality).
3. **Data Analytics Workbench (DAW):** Piloting a data science workbench in a cloud environment where users learn to extract data, train algorithmic models, and explore business intelligence and data analytics methodologies. Mobile District's existing Azure cloud and data repository for the Dredging Quality Management (DQM) program is currently being used. Workflows transforming data into insights are also being developed. This pilot includes publishing guidelines for final capabilities and deliverables.

2.4 USACE data governance

This LoE involves caring for the data. If the data is not accurate, confidence will be lost within the organization. When data gaps occur in the data, decision-makers tend to augment the data with their intuition and experiences.

The five objectives for this effort are:

1. **USACE Data Catalog (UDC):** A UDC will be constructed in a cloud environment. It will be a fully managed cloud service where users can discover USACE data sources needed. With the UDC, any user (analyst, data scientist, or developer) can discover, understand, and consume data sources. The catalog includes a crowdsourcing model of metadata and annotations to provide a single, central place for USACE data users to contribute their knowledge and build a community and culture of data.
2. **Authoritative Data Sources (ADS):** Define authoritative data sources and associated processes, including 1) inventory of authoritative data sources; 2) develop an internal USACE register; 3)

- eliminate redundant data sources; 4) identify conflict areas where multiple Data Stewards stake claims to the same data source; 5) provide guiding principles for workforce adaptations; and 6) create an established taxonomy to be used across the organization.
3. **Quality control measures:** Ensure USACE incorporates a comprehensive data quality management process as part of the data production and maintenance activities. Identify shared data elements across Automated Information Systems (AIS) and apply community-driven consensus on data quality and data standards guidance.
 4. **Update USACE Information Management Enterprise Data Management Policy (ER-25-1-110):** This document was last updated in July 2013 and contains obsolete and outdated information.
 5. **Sync with Investment Review Board (IRB) requirements:** Identify Data as an Asset (DAAS) metrics for use in IRB proceedings.

2.5 Demonstrate to motivate

This LoE encompasses the cultural changes required for a transition into a data-driven organization. Decision makers across all levels will be required to rely on what data is telling them. Making such a transition requires a great deal of trust to be placed on the accuracy and validity of the data. Such a transition requires users to know the answers given are correct and possess the confidence to allow the data to persuade their decisions.

The six pilots for this effort are:

1. **Navigation Business Line Pilot:** This pilot will link several dredging program disparate data silos and apply data analytics to attain a complete picture of program value and inform dredging decision makers on dredging methods and activities.
2. **Hydropower Business Line Pilot:** Conduct a pilot study to use data and machine-learning algorithms that allow decision makers to understand how the environment, maintenance, operations, and expenditures interact with the Hydropower Business Line.
3. **Military Construction Pilot:** Conduct a pilot study in military construction to demonstrate how data can drive the decision process.
4. **Emergency Operations Pilot:** Conduct a pilot study in emergency operations to demonstrate how data can drive the decision process.

5. **Flood Risk Management Pilot:** Conduct a pilot study in flood risk management to demonstrate how data can drive the decision process.
6. **District-25 Pilots:** Identify and support Districts using advanced data analytical processes to derive insights from data.

3 Conclusion

As USACE transitions to the Data Revolution, education of its leaders is paramount. Training for the leaders will be conducted in the value and techniques of modern data analytics. Legacy data silo practices should be abandoned and the use of individual Excel spreadsheets should be phased out. Instead, data should be available for every employee to see, understand, and exploit the richness that of the data. Innovation will come with the understanding and access of the data, which will allow USACE to remain a world-class organization.

This strategy identified five LoE that contain 16 objectives and six pilot demonstrations. Efforts have begun in all lines of effort and completion of all objectives and demonstrations are slated to be completed by 2025. Once all are completed, a sustainable data strategy will be embedded within the workflow.

This strategy will be considered a success when the goals below are met.

- A continuous leadership buy-in on the value of USACE data as a strategic asset.
- An adequate workforce within USACE that can analyze data and provide insight for decision makers.
- Technology in place to adequately capture, collect, organize, share, and analyze USACE data.
- USACE data considered authoritative, accurate, and accessible.
- The value of USACE data to aid in decision making has been markedly proven through pilot demonstrations.

Appendix: Insights Magazine April 2019 Issue

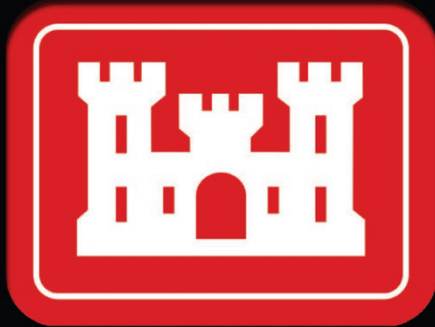
The complete issue can be seen below or located at

<https://team.usace.army.mil/sites/HQ-CI/PDT/USACEDA/default.aspx>.

Insights

AN UPDATE OF THE
USACE DATA STRATEGY
INITIATIVE
Volume 2 - APRIL 2019

A Revolutionary USACE
Informing
Decisions
Through Data



INTRODUCTION

Information at the speed of relevance

The Data Revolution is under way, presenting an opportunity to **REVOLUTIONIZE** the speed and accuracy of our daily decisions within the Corps of Engineers. We'll need the most accurate information that is available at the speed of relevance. Without it, we could miss opportunities for timely decisions, or worse yet, possibly make decisions that is not informed. This second update brings you up-to-date on enterprise efforts to turn USACE's immense volume of historic and dynamic data into one of its biggest strategic assets. The Army and other successful global businesses are already leveraging the opportunity.

"Our data strategy enables informed decisions at all levels; absolutely essential for the Revolution!!" said Commanding General LTG Todd Semonite in his personal note on Semonote #18 (see page 3). It is that note where he gave us our marching orders to change the culture of the Corps and become a data-informed decision-making organization.

The newly created Data Governance Board is already bringing together our best technical and senior leader experts! Their goal is to create and support a data-driven environment that consistently and accurately delivers authoritative, timely, relevant, and secure data to decision makers across the enterprise.

There are a lot of moving pieces to USACE's data strategy:

- We want leaders at all levels to understand the importance and value of their data;
- We need our workforce to become data-savvy, able to draw insights from data on their own;
- We are working on centralized repositories that need to be easily accessible?, securable, and easy to navigate;
- We need quality data that we can trust; and
- Several demonstrations are to be conducted to show the value of data to our organization.

The Data opportunity we have goes well beyond a **FOUNDATIONAL TASK** of doing routine tasks to a high standard. **DELIVERING OUR INCREASED PROGRAM** on time and within budget requires **ENGINEERING SOLUTIONS FOR THE NATIONS TOUGHEST CHALLENGES** using 'Big Data'. I am proud of our Districts, Divisions, and Business Lines that are already on board with turning their data into a strategic asset.

This Data Revolution is already occurring all around us. Thank you for leading and leveraging this opportunity.

ARMY STRONG.....BUILDING STRONG!




**Major General
Mike Wehr**

Major General Mike Wehr

Deputy Chief of Engineers/Deputy Commanding General

Call To Action


SemoNOTE



US Army Corps of Engineers

**A MESSAGE FROM
LTG SEMONITE**

08 January 2019



Call To Action

SemoNOTE #18: Data Strategy

SemoNOTE #18: Data Strategy

To the USACE TEAM:

At the 2nd Quarter FY18 Executive Governance Meeting I issued a call to arms, for all leaders at every level, to REVOLUTIONIZE how the Corps does business. "We have to change the culture for the Corps of Engineers if we wish to remain a WORLD-CLASS organization." Cultural TRANSFORMATION begins with allowing informed decision-making at every level. A sound DATA STRATEGY encompassing strong data governance and quality analytics drive those informed decisions.

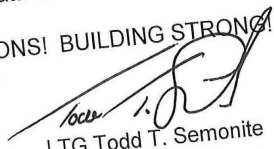
To facilitate a sound DATA STRATEGY, I have tasked the Deputy Commanding General to establish a Data Governance Board. The Board will develop a data strategy that oversees the Army's Strategic Intent for USACE: capturing the full business value of our data investments; allowing data to drive continuous improvements and innovation; accepting information as a strategic asset; leveraging analytical insights to improve business performance; and incorporating information governance across established information sources.

Most of Industry, DoD, and the Army share the vision of data as an ENTERPRISE ASSET, driving informed decision making and innovations. USACE currently uses standard, ad hoc reporting to answer "what happened" questions, as well as dashboards to diagnose why it happened. These tools have limited effectiveness if we don't efficiently GOVERN the underlying data supporting these tools. This includes AUTHORITATIVE data determinations; exploitation of resident data in all USACE relevant systems and investment in reporting and dashboard capability. USACE will employ analytics both Predictive (How will it happen – forecasting/data mining) and Prescriptive (How can we make it happen – data analytics/predictive) to optimize data to make better enterprise informed decisions.

I challenge our leadership to accept data as an enterprise asset and to invest in tools and governance processes that meet the Army's intent. The DCG will contact specific senior leaders to discuss the Data Governance Board's way ahead.

ESSAYONS! BUILDING STRONG! ARMY STRONG!

Our Data Strategy Enables Informed Decisions at All Levels; Necessary Essential for the Revolution!



LTG Todd T. Semonite
54th Chief of Engineers
Commanding General, USACE

SemoNOTE

Call To Action



USACE DATA ANALYTICS STRATEGY



Five efforts under way to establish data as a USACE strategic asset

USACE has been in the data business for a long time. But until recently, data was considered just a byproduct of our missions and used primarily for reporting.

Now, USACE leaders are seeing value in that data – so much value that our commanding general wants us to use this vast amount of data to revolutionize how USACE does business so that we remain a world-class organization. He wants to begin treating our data as a strategic asset and use it to begin harvesting insights to improve decision making across the organization.

The transition to a data-driven organization will take time, however. It starts with the organization declaring data as a major resource. Meeting this objective lays the groundwork for USACE to start the transition. It will not happen overnight because advances in some business areas will outpace others. The transformation will be incremental with adjustments made along the way. In the end, data will be viewed favorably and used to

drive the business.

Five lines of effort (LoE) are being used in the Data Strategy Initiative. These efforts are not being executed serially. Action and refinements in each are ongoing and maturing throughout the entire transition period. Interactions with other Army agencies making a similar transition will occur. Feedback across the organization will be used to refine and update the approach.

LEADERSHIP BUY-IN

Leaders and decision makers at all levels of the Corps must learn to trust their data, instill data discipline, and educate their workforce to understand and value what data can show them and their missions. USACE senior leadership has defined clear goals through the creation of a Data Governance Board (DGB).

Four objectives are in this effort:



INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE



Support Executive Governance Meeting (EGM)

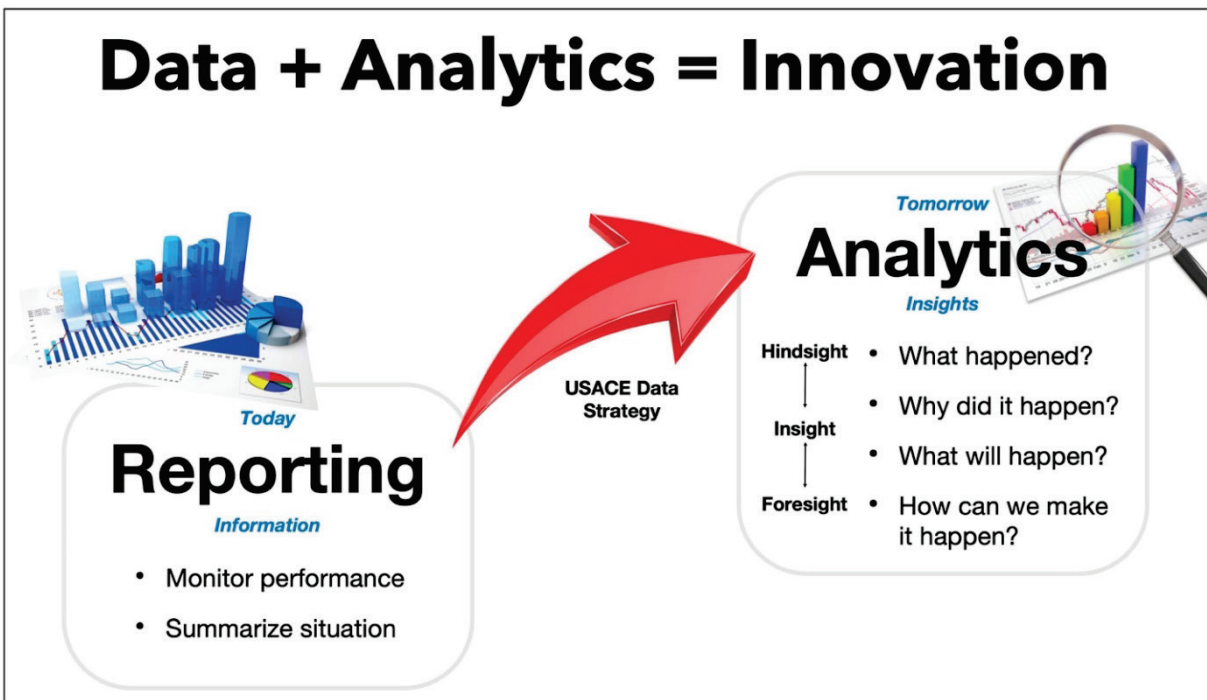
activities: Leaders from across the Corps gather regularly for professional development and to discuss issues throughout the organization. Regular feedback on the Data Strategy Initiative will be given.

Support Data Governance Board activities: The USACE Data Governance Board (DGB) is the senior recommending directive body for development of coordinated USACE enterprise positions on data strategy, data standards, data quality, and data interoperability. This board serves as the senior adjudication body for the USACE Data Strategy Initiative; maintains the

inventory of USACE authoritative data sources; serves as a certification/waiver approval authority for targeted data standards as delegated by the Deputy Commanding General (DCG); and collects and disseminates best practices and lessons learned for the data community. Within their charter, they have listed several priorities and deliverables for the board which are being addressed by these lines of effort.

Promote leadership development (the value of data):

A set of recommended executive courses about data analytics will be supported through the USACE Data Strategy Initiative. These courses will help USACE



leaders gain insight into how industry is using data and applying data analytics within their own organizations. With real-world examples, it will be demonstrated how treating data as a strategic asset can improve decision making across the Corps.

Sync with leadership: New generations of USACE leaders will have to be sold on the value of data as they come into the organization. Opportunities will be leveraged to update leadership concerning the value of data and the state of current effort of the USACE Data Strategy initiative. Sub-tasks to inform leadership of the state of the USACE Data Strategy Initiative include:

- **DCC and GO-SES data analytics modules:** Modules about the value of USACE data will be constructed for use in the District Commander Course and the General Officer-Senior Executive Service (GO-SES) courses given regularly by USACE lead-

ership. These modules will give an overview of the importance of data as a strategic USACE asset. The DCC and GO-SES courses establish the doctrinal framework for District operations, as well as specific tactics, techniques and procedures for success are held for new District Commanders and Deputy District Commanders on a regular basis.

- **Regular USACE Data Strategy Initiative reports:** A digital magazine, called Insights, will be published regularly to inform USACE about the status of the USACE Data Strategy Initiative.
- **USACE Data Analytics web presence:** A USACE SharePoint team site has been established for open access to all relevant documents about USACE Data Analytics. The site can be found at <https://team.usace.army.mil/sites/HQ-CI/PDT/USACE-DA/default.aspx>.



WORKFORCE DEVELOPMENT

The growing consensus that data analytics can bring real value to USACE has led to an explosion in demand for employees who are trained in data analytics — individuals who have an understanding of analytics techniques and can apply them to real USACE problems.

These individuals skilled in data science, however, must be accessible across the organization. This will require changing mindsets to see data teams (data engineers, data scientists, and data analysts) not as an expense but a true investment, and understanding data analytics is more about the skilled people in the data teams and less about technology and tools.

Four objectives are in this effort:

Position Description: The incorporation of District Data Officers (DDO) will bring the proper skills needed to gather data and use it to assist leaders in answering questions and making decisions. The DDO will also act as an enterprise-wide resource for sharing data and help answer questions using data across organizational boundaries. Over time, this DDO community will flourish in capabilities and become a key player in the USACE decision process. A Position Description (PD) has been written to assist Districts in hiring.

PROSPECT Training/Certification: A large majority of the DDOs will come into the job not being formally trained in data analytics. A PROSPECT curriculum is being developed that will provide the foundational knowledge required by each person selected to serve as a DDO. It will include a certification process so that each person who completes all three courses will be recognized as being a Subject Matter Expert (SME) for USACE Data Analytics. The organization of the certification program is based on three tiers of training. All three tiers of certification will be required for DDOs.

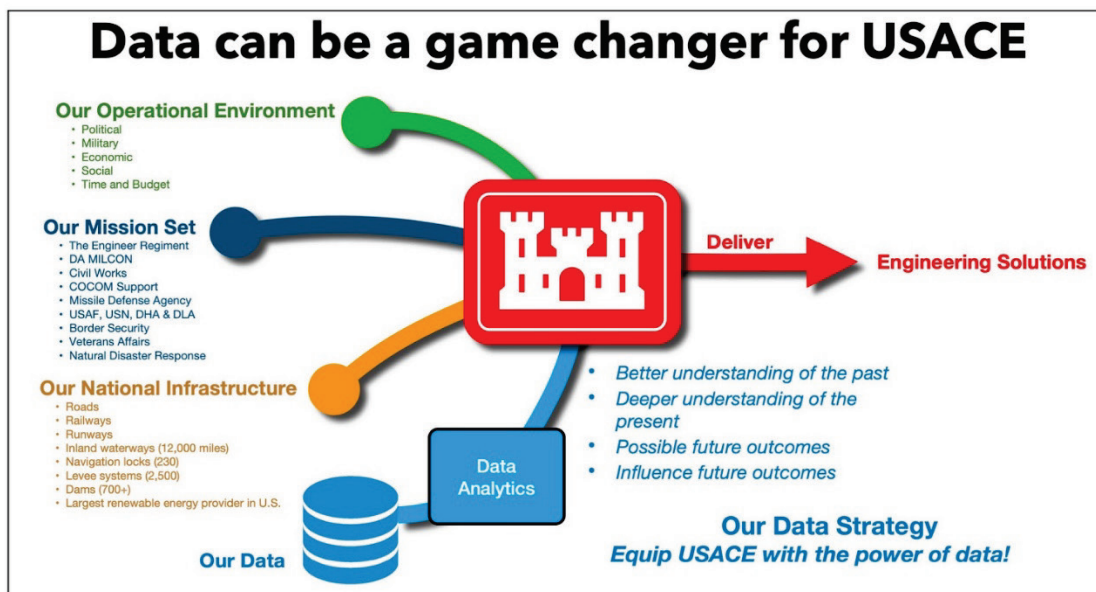
USACE Data Analytics Center of Expertise (DACX): Primarily built to create a knowledge base of expertise in data and data analytics, the DACX would accelerate

the cultural transformation of the Army and USACE – Districts, Divisions, and Business Lines – to manage and use data in informed decision making. The DACX would develop the expertise and methodologies required of a CX to assist Districts and Divisions as USACE data is transformed into a strategic asset for the organization. The DACX would focus on making data analytics a repeatable process for all Divisions and Districts and support the creation of data analytics capabilities down to the District level. It would also lead a community of practice across the enterprise focused on maturing DA capabilities and promote the sharing of methods and data essential for decision making. It would also serve as a clearinghouse for information capture and flow between the field, CIO, and ACE-IT. The objectives of the CX are: 1) support USACE leadership's goal to improve decision making through the use of data; 2) provide the field with technical support, best practices, analyses, and the overall introduction to DA; 3) help reduce redundancy and promotes the sharing of DA advancements as they occur in the field; 4) ensures Divisions and Districts progress at a similar pace in becoming efficient with using data to make decisions; and 5) work with data owners to ensure consistency, easy access, and quality of data.

Data Workforce of the Future Planning: What will our workforce require in order to work with USACE data so that it produces insights? This plan will help guide USACE in its hiring and training practices as data becomes an integral part of the organization.

USACE DATA ECOSYSTEM

This line of effort supports the technology requirements of the Data Strategy Initiative. Many think the purchase of a tool in this line of effort is all that is needed to magically convert data to insights, but no such tool exists. The immediate focus is on investing in the technology used to capture, collect, organize, share, and analyze USACE data. Microsoft's Azure will provide



the IT platform for this data strategy.

Three objectives are in this effort:

Data Migration: The process of selecting, preparing, extracting, and transforming data and permanently transferring it from USACE's legacy system to a cloud environment is being formulated.

Data Architecture: A plan to establish models, policies, rules, and standards that govern which data is collected, how it is stored, arranged, and put to use is being formulated for USACE data.

- **Enterprise data configuration management plan:** Integrate and maintain Information Exchange Specifications and associated Controlled Vocabulary (Terms of Reference) and database requirements and fund them to comply. (Data Quality).

Data Analytics Workbench (DAW): Piloting of a data science workbench in a cloud environment where users can learn to extract data, train algorithmic models, and explore business intelligence and data analytics methodologies. Mobile District's existing Azure cloud and data repository for the Dredging Quality Management (DQM) program is being used. Workflows for how data can be transformed into insights are also being developed. The pilot includes how final capabilities and deliverables can be published and used by decision makers across the organization.

USACE DATA GOVERNANCE

This line of effort involves caring for the data. The old principle of garbage-in, garbage-out still holds true. If the data is not viewed as accurate, then the organization



will lose confidence and will return to its past ways of making decisions. When data gaps occur in the data, decision-makers tend to augment the data with their intuition and experiences.

Five objectives in this line of effort include:

USACE Data Catalog: A USACE Data Catalog (UDC) will be constructed in a cloud environment. The UDC will be a fully managed cloud service where users can discover USACE data sources they need and understand the data sources they find. With the UDC, any user (analyst, data scientist, or developer) can discover, understand, and consume data sources. The catalog includes a crowdsourcing model of metadata and annotations to provide a single, central place for USACE data users to contribute their knowledge and build a community and culture of data.

Authoritative Data Sources (ADS): Define authoritative data sources and associated processes, including 1) inventory of authoritative data sources; 2) develop an internal USACE register; 3) eliminate redundant data sources; 4) identify conflict areas where multiple Data Stewards stake claims to the same data source as their own; 5) provide guiding principles for workforce adaptations; and 6) create an established taxonomy to be used across the organization.

Quality control measures: Establish a data quality control program to ensure USACE incorporates a comprehensive data quality management process as part of their data production and maintenance activities. Identify shared data elements across Automated Information Systems (AIS) and apply community-driven consensus on data quality and data standards guidance.

Update ER-25-1-110: This document outlines the Information Management Enterprise Data Management Policy. It was last updated in July 2013.

Sync with Investment Review Board (IRB) requirements: Identify Data as an Asset (DAAS) metrics for use in IRB proceedings.

DEMONSTRATE TO MOTIVATE

This line of effort encompasses the cultural changes required to transition into a data-driven organization. This will require decision makers across all levels to move away from just using hunches and instinct to relying more on what data is telling them. Making such a transition requires a great deal of trust being placed in the accuracy and validity of the data. Just delivering insights from data is not enough. Such a transition requires users to believe the answers are correct and possess the confidence to allow the data to persuade their decisions.

Six pilots are planned in this area:

Navigation Business Line Pilot: This pilot will link several dredging program disparate data siloes and apply data analytics to attain a complete picture of program value and inform dredging decision makers on dredging methods and activities.

Hydropower Business Line Pilot: Conduct a pilot study to use data and machine-learning algorithms that allow decision makers to understand how the environment, maintenance, operations, and expenditures interact with the Hydropower Business Line.

Military Construction Pilot: Conduct a pilot study in military construction to demonstrate how data can drive the decision process.

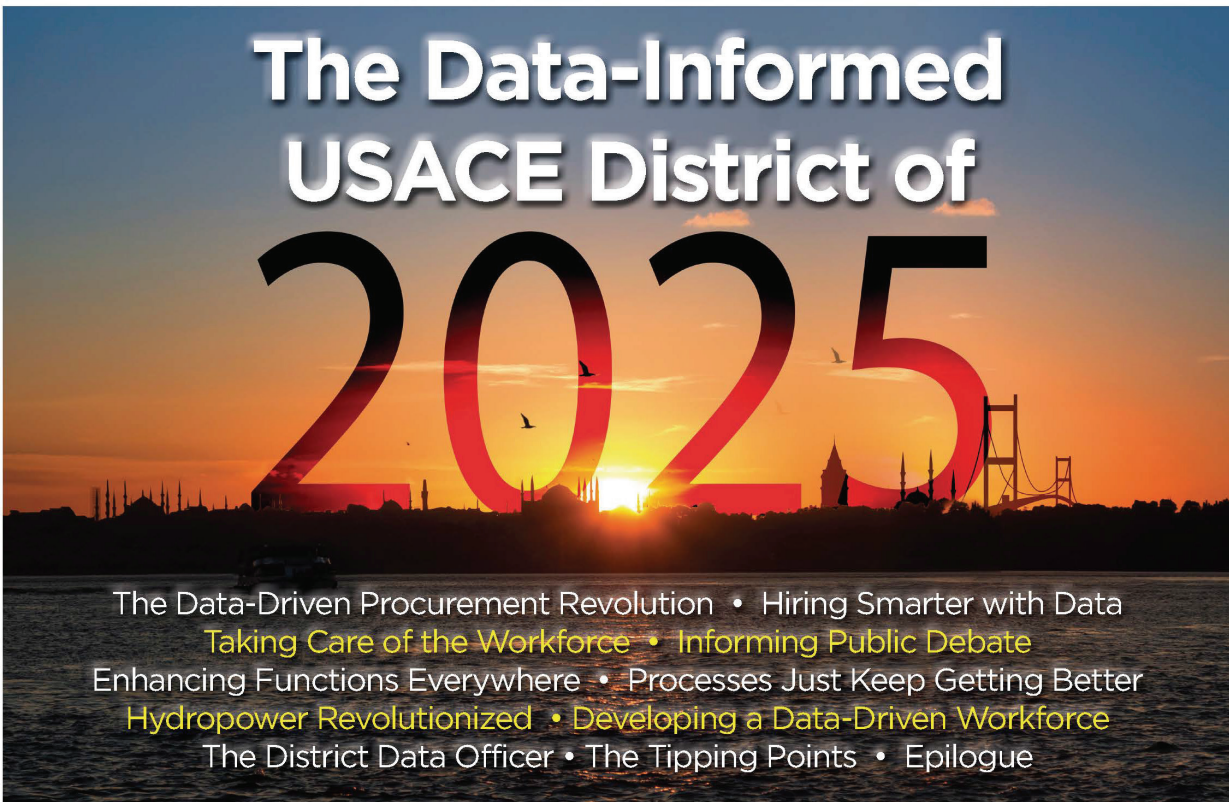
Emergency Operations Pilot: Conduct a pilot study in military construction to demonstrate how data can drive the decision process.

Flood Risk Management Pilot: Conduct a pilot study in military construction to demonstrate how data can drive the decision process.

District-25 Pilots: Identify and support Districts who are using advanced data analytical processes to derive insights from data.



A DISTRICT COMMANDER'S PERSPECTIVE



Every revolution begins with a vision of what might emerge as the struggle unfolds. With the start of the USACE data revolution in spring 2019, we set out to imagine what a District would look like once it learned to harness the power of data analytics. A work of fiction, this narrative is set in 2025 and looks back on the past six years in the District, seeing what has changed for the better. My hope is that this story helps inspire everyone who is vested in the success of USACE, encourages those data champions at every level, and that it unlocks the imagination of those who don't yet see the power of data.



By

COLONEL ROBERT DIXON
 SWL Commander & District Engineer

USACE DATA ANALYTICS INITIATIVE



should define clear goals and paint a picture of what success looks like. No longer should data just be collected and buried deep in the organization on someone's desktop computer or a siloed server. Data should be viewed as a critical resource for the organization much like electricity that runs USACE facilities or the Nation's waterways. The importance of data must be understood and appreciated.

Step 2 involves talent management. Individuals skilled in data science must be accessible across the organization. This will require changing mindsets to see data teams (data engineers, data scientists, and data analysts) not as an expense but a true investment, and understanding data analytics is about the skilled people in the data teams and less about technology and tools. The incorporation of District Data Officers (DDO) will bring the proper skills needed to gather data and use it to answer questions. The DDO will act as an enterprise-wide community for sharing and helping one another answer questions using data. Over time, this DDO community will flourish in capabilities and become a key player in the decision process.

As more DDOs are added to the organization, lessons learned from data discovered in one district can be shared and applied across district lines. At some point, Division Data Officers can be brought onboard to coordinate the sharing of data insights across the enterprise.

Step 3 encompasses technology. Some would argue tools exist that magically convert data to insights, but no such tools exist. The immediate focus should be on investing in technology used to collect, organize, and share USACE data. As mentioned earlier, data should be considered as a critical organizational asset and technology must be in place to properly care for this asset. The analysis of the data requires a combination of data scientists, decision-makers, and technology. All three are equally important.

Step 4 involves caring for the data. The old principle of garbage-in, garbage-out still holds true. If the data is not viewed as accurate, then the organization will lose confidence and will return to their past ways of making decisions. When data gaps occur in the data, decision-makers tend to augment the data with their intuition and experiences.

Step 5 encompasses the cultural changes required to transition into a data-driven organization. This will require decision makers across all levels to move away from just using hunches and instinct to relying more on what data is telling them. Making such a transition requires a great deal of trust being placed in the accuracy and validity of the data. Just delivering insights from data is not enough. Such a transition requires users to believe the answers are correct and possess the confidence to allow the data to persuade their decisions.

No longer should data just be collected and buried deep in the organization on someone's desktop computer or a siloed server.

Data should be viewed as a critical resource for the organization much like electricity that runs USACE facilities or the Nation's waterways.

The importance of data must be understood and appreciated.



NATIONAL SECURITY STRATEGY

"Data, like energy, will shape U.S. economic prosperity and our future strategic position in the world. The ability to harness the power of data is fundamental to the continuing growth of America's economy, prevailing against hostile ideologies, and building and deploying the most effective military in the world."



INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE

to meet scheduling needs. In our Medical Initial Outfitting line we were able to predict supplier challenges before they came to light. In some cases, we were able to adjust construction schedules to accommodate the lag, and in other cases we were able to intervene with the supplier to accelerate lagging equipment delivery. This wasn't possible before we started combining historical information in our databases with the market signals available through other government and open-source databases to allow our computers to conduct a predictive analysis.

One of the biggest game-changers for procurement was the automation of micro-purchases. No longer did we need a lengthy process for procuring routine items (especially when the cost of said items exceeded purchase card limits). The automated process taking advantage of the Internet of Things (the connection of machines and other objects to the internet) freed employees from wasting time ordering things manually. A simple example of this is how we now order printer cartridges. Instead of incurring the cost of monthly transactions for ordering ink, our printers detect their ink levels, analyze historical usage rates and shipping times, and contact the vendor directly. Before the user knows she is low on ink a new one arrives. No procurement specialist spends time finding three bids and printing out receipts. No supervisor has to review the purchase. Nobody wastes time on a routine task that a machine could do. The funny thing was that this capability was always there and was widely used in industry. We just had to get out

of our own way to turn this feature on.

We took that concept to nearly all of our supply actions, saving thousands of labor hours and procurement administrative costs. Yes, we had to overcome the resistance to the automated model, and it took Army and USACE leaders to advocate changes to the Federal Acquisition Regulation, but it was worth it. A few changes in the FAR and we saved hundreds of thousands of taxpayer dollars.

We were surprised at how much analytics reduced our procurement risk. Whether or not we articulate the risks we take in hiring contractors, we assume four major categories of risk every time we bring in a vendor to do work: execution risk, commercial risk, continuity risk, and compliance risk. Once we learned about how analytics helped tackle these risks, we found ways to make source selections with more confidence, and transferred those models to construction for better performance monitoring.

This transformation wasn't easy. We had to re-look the talent we were hiring for procurement. When we started there was a significant gap in data analytic skills needed to compete in the modern construction market. We mitigated some of that gap through the USACE PROSPECT courses in data analytics, and changed Position Descriptions and Specialized Experience to include experience in data. Of course, we never would have transformed our workforce the way we did without revolutionizing our hiring process.

ABOUT THE AUTHOR

Colonel Robert "Bob" Dixon is a USACE District Commander and self-described data geek. This narrative is written from the standpoint of where we might see ourselves in 2025 looking back. The description of where we are in 2019 is not meant to disparage, only to bring a critical view to the gap between what is and what might be. There are undoubtedly people already working towards these objectives, or may have already achieved them. If that is the case, the author thanks them for doing that and asks that these capabilities be pushed to the district level as soon as possible. The author also recognizes that achieving all of these things over the next half decade is highly unlikely, but we won't get very far if we don't start asking now. The opinions expressed in this narrative are the author's and don't necessarily reflect SWL, SWD, USACE or the US Army's official position.





Hiring Smarter with Data

In the past we used to rely on the “tried-and-true” “post-and-pray” method of finding talent. We did this through a laborious process designed to screen out people and limiting the number of people who would apply. Forced to use OPM’s USA Jobs and CHRA’s antiquated business rules, we mostly relied on the pure luck of timing, hoping our ideal candidate happened to be online and looking at our Job Opportunity Announcement (JOA) during the seven days it was available online and would be patient enough to navigate the process. That all changed when we adopted an Active Recruiting Posture.

Our Talent Acquisition Specialists (TAS) (we had to make up that position since USACE didn’t have recruiters for its 35,000-person workforce) came onboard in 2018 and began analyzing the market for talent. They quickly concluded that they didn’t have the right tools to find the right talent, let alone recruit them. That all changed in 2019 when USACE obtained the authority to use USA Jobs’ resume search tool and OPM’s online data analytics programs (note: that is an aspirational statement... the tool exists now in 2019 but USACE is not allowed to use it!). These tools, combined with Google Analytics on our public-facing websites gave our TAS insights into who was interested in careers with us. By analyzing hits on our recruiting page (which abandoned stale, outdated USACE standardized

templates) and our JOAs, we were able to determine where our most interested candidates were located. TAS and PAO began to focus messages in those areas.. We then began the dialogue with potential employees long before a JOA was online. When we released the Job Announcement, we already had several high-quality candidates ready to apply, with explicit instructions on how to navigate the CPAC processes.

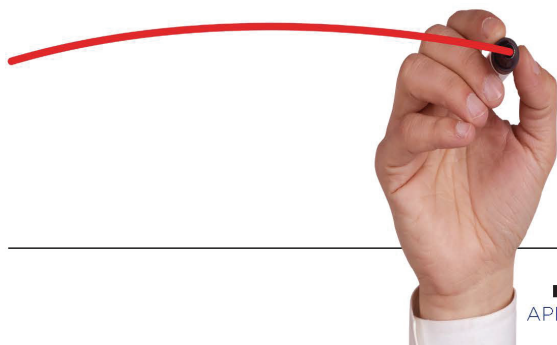
With enhanced hiring authorities brought about by the “Revolutionize the Corps – Updating Authorities and Policies” line of effort, we are now able to locate, recruit, hire and onboard the right talent for tomorrow in weeks instead of months.

Analytics also helped our leaders determine our workforce needs. Predictive analytics helped us forecast the workforce-workload analysis with greater fidelity and confidence. These models continue to evolve, but leaders at the District and Division levels no longer spend days or weeks on this analysis: it is done continuously and instantaneously by algorithms running in the background of our automation systems.

As we saw the increased need for data analytics skills, we used our new market research skills to identify those markets where the skills were prevalent – and were able to grow the workforce in both quantity and quality.

By using data-informed active recruiting we dramatically increased our quality of hire. Our Quality of Hire analytics program helped us understand our best talent streams and what recruiting techniques were most successful. It was with some surprise that the data revealed some counterintuitive insights about where our best employees came from and how we could best reach them. By augmenting our hiring managers’ judgment with machine learning, we are now capable of recruiting and assessing the workforce we need in terms of increased quality, not just fill rates.

Join our team





Taking Care of our Workforce

Taking care of our workforce has become easier with data analytics as well. Our models now reliably predict attrition rates by month, season, or year based on market conditions, employee engagement, time and attendance, and local environmental factors. It's actually a little eerie how accurate these models have become.

Knowing our attrition rates helps us set our hiring goals, and also helps us understand some of the reasons people leave. Moreover, we got to know our workforce in new, more meaningful ways. In the past, we relied on exit interviews, which amounted to a series of anecdotes (i.e., not data), to reveal why an individual made a departure decision. We also relied on surveys (e.g. FEVS or Command Climate Surveys) to get feedback on how engaged our workforce is. But data analytics offered us some new insights.

Once we were able to correlate at-work engagement, measured by time and attendance, work output, login rates, how long people spent watching cat videos on YouTube, etc., we were able to identify when employee engagement levels changed. Since we used anonymized data for this analysis (for privacy reasons), we did not try to intervene at the individual employee level. However, we were able to see when drops in engagement were correlated with specific events or changes at the district, or whether they were correlated with some external factor (e.g., the weather, government shutdown, etc.). This gave leaders insights into how our workforce responded to certain conditions, and helped us determine the types and timing of morale-enhancing messaging or events.

Informing Public Debate

Multi-use lakes are among the most contentious environments when it comes to public planning. Water supply competes with hydropower, recreation, fish and wildlife, and Flood Risk Management for use of the water or storage capacity. Over the past five years we have been able to create models that reliably portray the risk factors associated with water allocation changes. Incorporating the impact of population growth rates, changes in economics, land use, and the cost of renewable energy, these models help planners optimize the use of multi-use lakes, and the data visualization tools help the public and political leaders better understand the tradeoffs between the myriad constituencies competing for these resources.



Our Public Affairs community plays a key role in this arena. Using analytics to understand who is contributing to public debate, how public sentiment is affected, and which medium is impacting our environment most has made our PAO exponentially more effective. No more hand-jamming PowerPoint Slides with yesterday's Facebook posts during disaster response, our PAO uses real-time, global stream monitoring for social and news media to form a 360-view of public sentiment and identify influencers during a crisis. Done by offsite servers, our tools analyze millions of posts across multiple media and uses natural language processing and machine learning to refine post categorization, way beyond basic positive, neutral or negative. This helps guide senior leader engagement in a very sophisticated and effective way and allows USACE to get their messages to the right audiences at the right times through the right avenues. There is no better way to protect our brand than to engage in the information sphere armed with in-depth, real-time analytics.



INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE



Enhancing Functions Everywhere

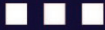
Nearly every business line and function benefited from the data-augmented workforce. Navigation maintenance planners predicted maritime traffic patterns and adjusted maintenance schedules based not on what industry says it does, but by what it actually does in the real world. The Navigation Data Center gets higher quality data directly from smart gates, Facilities and Equipment Maintenance (FEM) systems and river gauges and is able to give us insights we never had before. Better yet, they now push us information about our system where we had to pull it from them before.

Our Operations Project Managers have new insights about how our visiting public enjoys our lakes, parks, and campgrounds. By analyzing anonymized cell phone data, we are now able to tell how many people are on the lake at any given time, and determine the predominant activities. In several instances on a high-use recreation lake, our monitoring system alerted USACE Park Rangers to an increased level of water skiers in a historically overcrowded cove. Since the number of boats in the cove exceeded the maximum safe boats per acre, our rangers coordinated with the State Police Water Patrol to direct an officer to that cove. During the first summer of implementation, there was a 25% reduction in water skiing accidents on that lake.

Data Analytics has enhanced Environmental Stewardship (ES) by identifying high-risk areas for timber trespass and encroachment. Our ES analytics system

monitors publicly available open-source information and correlates new adjacent land acquisition, building permit issuance, and terrain analysis (from our own GIS system) to predict the most likely encroachment areas. Rangers are dispatched to at-risk areas and have been able to engage with offenders before the destruction of public property became irreparable on four occasions in the past two years.

Flood Risk Management is one of the business lines already fairly mature in its use of data when the Data Revolution began in 2019. The data collected from USGS stream gauges, the robust modeling, the Corps Water Management System (CWMS) and a host of data-driven hydrologists, made flood fight decision-making a more confident endeavor. But some significant risks remained in the system: computing power and bandwidth limitations that left our Flood Fight Teams without redundant offsite backup. Should a long-term power outage occur during a flood fight, we would be without access to the real-time data needed to make decisions. That changed once the USACE Data Revolution Group established governance guidelines that made all district data available in the CWMS Cloud architecture, making it possible for another District to access and analyze Little Rock data in a crisis. Thankfully, we only had to use that system once, but we were sure glad that Tulsa's Flood Fight Team had our back for the several hours we were down.



Processes Just Keep Getting Better

Project Delivery is the coin of the realm, and our Data Team has been essential in helping us deliver better projects. Over the past five years, we have used data analytics to understand where our processes are working and where they are lagging. In pre-award, we have improved our acquisition planning, our collaboration, and our scheduling based on the analysis of thousands of projects.

Identifying the factors that contributed to contract quality was another big gain. Our data team reviewed thousands of contracts from across USACE to see if there were common factors that contributed to cost or schedule growth. It turns out that there were! The elusive “leading indicators of failure” were hidden in un-analyzed data sitting in USACE databases. One great insight we gained was from analyzing the causes of contract modifications during execution. For SRM work, “unforeseen site conditions” was the leading cause

of contract modifications, accounting for 67% of all mods in the past 20 years across the enterprise. So we focused on reducing that number. A few select Early Contractor Involvement acquisitions along with use of Prequalified Contractor pools for site visits and on-site draft document reviews during the design phase helped the government prepare better technical requirements and helped contractors make better risk-informed decisions about their bids, reducing the cause for mods by 22% over a 2 year period.

Scheduling benefited from analytics as well. While standardized timelines for business processes are a nice goal, they are not very useful when they don't reflect reality. If a process takes, on average, 30 days to complete but our enterprise goal is 15 days, we are bound to under-deliver. Where did 15 days come from as a goal? After analyzing thousands of project timelines, our data team helps PDTs determine realistic timelines based on analysis, not hope. Under particular market conditions and for specific kinds of work, contracts can predictably

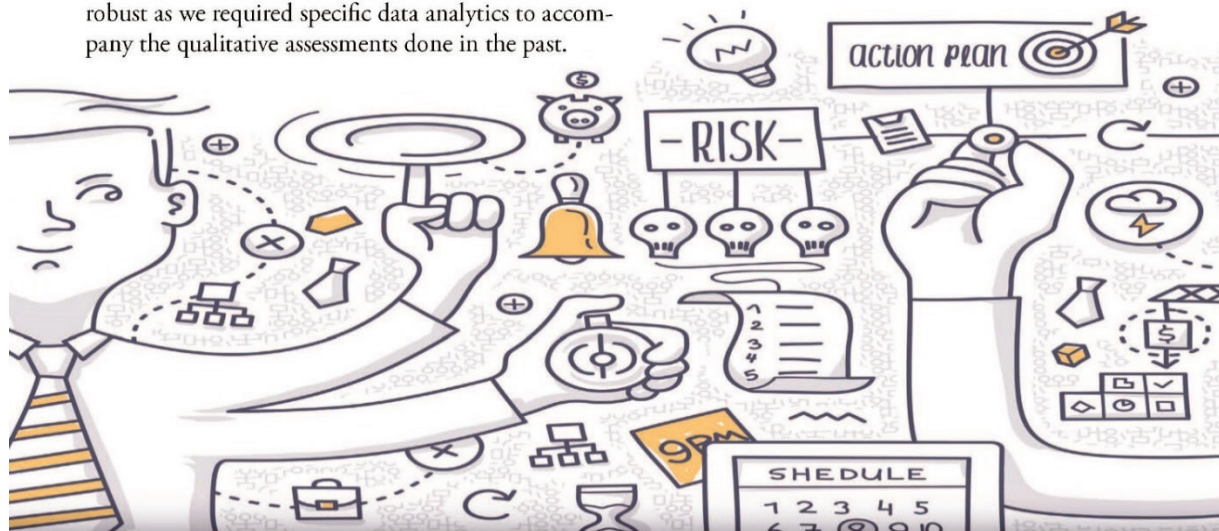


INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE

be awarded in 60 days, for example. Under different conditions (such as time of year, increases or decreases in non-governmental demand for the same products, likelihood of protest under these conditions, etc.), the same contract effort may take 90 days. By using our data team to forecast timelines based on actual conditions, we were able to more accurately communicate deliverable timelines to our stakeholders with a 92% reliability rate.

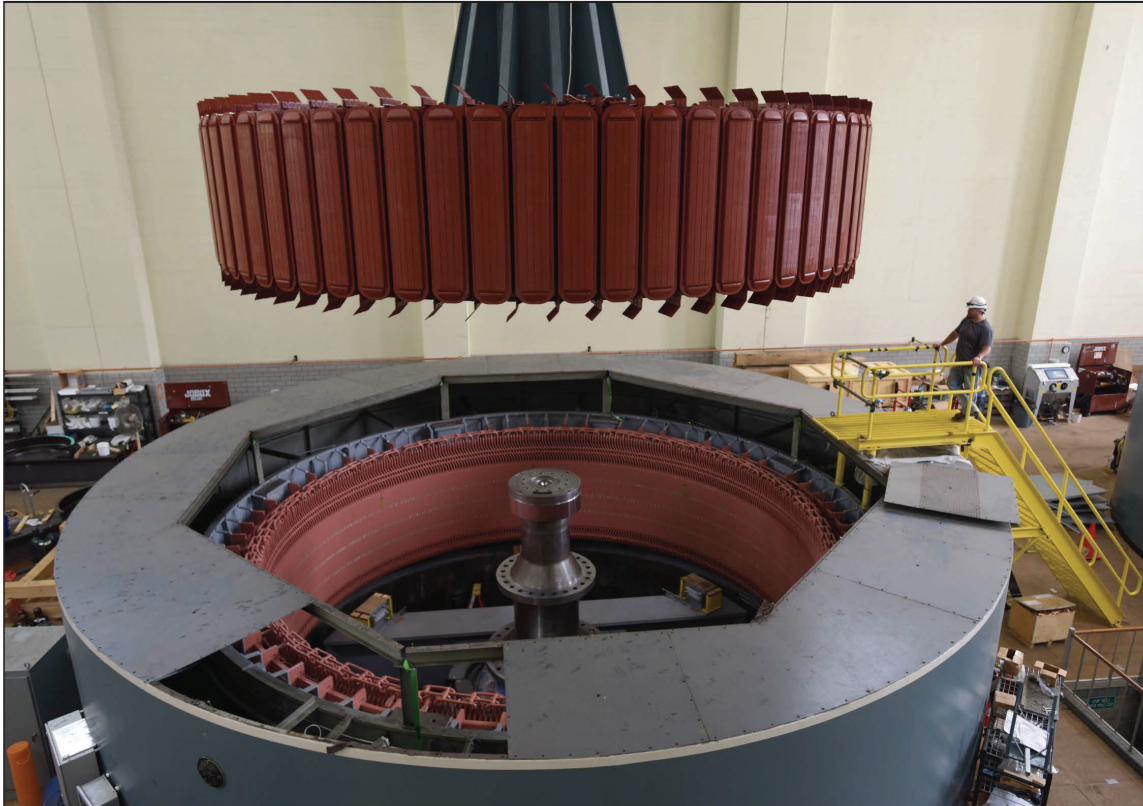
By setting our data team loose on our Project Delivery Business Process, we were able to gain some interesting insights that helped us deliver the program better. Besides data-informed scheduling and acquisition, we also found insights in resource estimating and change management processes. Analysis of Value Management yielded some insights, but likely the new protocols in data management were the most effective change that yielded long-term results for the District. No longer were “P-Drive” folders with cryptic filenames the standard. Every piece of information associated with the project is collected in the database of record – and is accessible by PDT members and leaders throughout the District. Additionally, the AAR process became more robust as we required specific data analytics to accompany the qualitative assessments done in the past.

Data Analytics has probably had the biggest impact on the budget process. In the past our budget process seemed to be largely risk-based. If a leader could articulate how badly the infrastructure may fail and the dire consequences associated with that failure, his or her budget had a chance of increasing. Failure to frighten budgeteers typically led to decreasing budgets. Now, our Civil Works budget is informed by FEM, datafied condition assessments, and the correlation between spending, labor use, environmental conditions, past infrastructure failure, consequence projections, usage rates, infrastructure age, and market conditions. This is based on the analysis of millions of maintenance and budget data points that combine to reveal not only our greatest risks but our also our highest payoffs. We never could have done this level of analysis by hand, and it was only after the Data Revolution that we were able to use machine learning to reveal these answers. This doesn't necessarily get us more funding every year, but it does help our civil works funding go further by helping us focus on the greatest needs. This leads to better resilience in our aging infrastructure.





Hydropower Revolutionized



Inspired by automated communications from his Ford Fiesta (his son's car texted him from the driveway that it would need new brake pads within the next 2500 miles, based not on miles or time but on the actual condition of the brakes) our Chief of Hydropower wondered out loud why his powerplant's critical systems weren't giving him the same courtesy. After all, if Ford can do it, surely the largest producer of Federal Hydropower ought to be able to set up diagnostic systems that inform deci-

sion makers of upcoming maintenance needs. This isn't the same as alerting him when something breaks. It's alerting him before it breaks. Now he can budget, order and receive parts, and fix the impending fault before we suffer an unscheduled outage. Better yet, the monitoring system in some cases can actually order the part for him!

Why is that so important? Because over the past four years we have transformed USACE hydropower from



INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE



an at-risk business line to a prosperous one, and secured federal hydropower's position in the renewable energy market. It was at risk in 2021 when preferred customers began shifting towards the highly subsidized solar and wind producers, and walking away from hydropower due to the increasing costs. Hydro couldn't compete with other renewables because our costs continued to rise at the same time our competitors' costs were dropping.

This is a national energy strategy issue: Hydro does more than just produce electricity. It stabilizes the grid. It produces when the wind stops and the sun goes down. It starts and stops quickly and reliably while other forms of power generation take hours to do the same thing, making hydro responsive to unpredicted peak demands. We couldn't allow this business to fail, but declining budgets and increasing O&M costs were a burden to the customers. They were ready to walk away, and take their customer funding for capital reinvestment with them. Thankfully, that all changed.

Using advanced telemetric sensors installed in 2019, plants were able to see stresses and strains on their infrastructure and units that they never saw before. Since every plant in USACE was outfitted similarly, COP leaders were able to draw conclusions about what conditions produced damaging stresses using a large data set from every plant in the fleet. Combining this new data with maintenance data collected in the Facilities and Equipment Maintenance (FEM) program, the whole business line was able to see itself in a whole new way. Maintenance began to shift from reactive to preventive. Machine learning algorithms evaluated acoustic signals and compared them to historic maintenance issues. The Hydro-monitoring system began finding correlations between changes in frequencies and the failure of specific hydropower subsystems. Noting these changes, the governing system would then signal a maintenance supervisor, recommend a fix, and then order the necessary parts. Eventually this system became so reliable that supervisors were only required to ap-

This is a national energy strategy issue: Hydro does more than just produce electricity. It stabilizes the grid. It produces when the wind stops and the sun goes down. It starts and stops quickly and reliably while other forms of power generation take hours to do the same thing, making hydro responsive to unpredicted peak demands.

prove about 20% of the orders: the rest of the time the machine would order it on its own. This "Internet of Things" approach freed up maintenance folks to focus on issues more pressing than ordering parts.

At a macro level, the Hydropower COP and the Hydrological Design Center now have the ability to offer insights to districts based on data gathered from across the fleet. This isn't the "lessons learned" or "best practices" PowerPoint level of sharing done in the past. These are real insights about the way Francis Turbines, for example, are behaving under certain environmental and load conditions based on thousands of sensors at dozens of plants. We are able to make minor adjustments to how we operate based on a whole fleet of evidence, not just what we know about our individual units. These "minor" adjustments contributed to a significant decrease in cavitation-induced repair requirements, saving roughly \$750,000 in the first year alone!

Data Analytics doesn't make Hydro less expensive than solar or wind, but it did reduce our overall costs by 25% (found in acquisition savings, reduced maintenance burdens as we operated the units more optimally, and by reducing unscheduled outages by 35% over 3 years).





Developing a Data-Driven Workforce

Not only did we help our folks see the value of data, we transformed our workforce to crave data. Trained to be capable of manipulating data, and informed about what analytics could do for them, our workforce began producing value at unforeseen levels. Moreover, they were armed with the right analytic tools that didn't require a degree in data science to operate. Using interactive, visual tools our workforce can create their own queries quickly, accelerating their ability to dive deeply into the data to gain insights they need to make decisions.

Productivity went up. Their decisions got better. They didn't need to form committees, conduct endless studies, and get permission from higher headquarters to make a decision; they knew what the right decision was because they had the evidence they needed when they needed it. Once they got a taste for this, they didn't want to go back to the information equivalent of the bread lines they were standing in before.

With a demonstrated increase in reliable decision-making, leaders were free to delegate more and more authority to our data-informed, committed workforce. Increases in decision-making quality built trust in the vertical team, and the risk of delegation dropped dramatically. Employee engagement increased, as did job satisfaction as evidenced by a 15% increase in positive returns on the FEVs. We are proud to be labeled a "World Class Place to Work" for the fourth year in a row!

But a data-driven workforce didn't materialize overnight, nor did the data skills emerge immediately. It took time and commitment from the whole Enterprise to make this happen. Without centralized funding for the District Data Officer and workforce education we would still be relying on hand-jammed analytics like we were in 2019.

The District Data Officer

Probably the biggest catalyst to the data revolution is the introduction of the District Data Officer (DDO). Piloted in Little Rock in 2019, the first DDO helped define how Districts can exploit the mountains of data it has access to. The DDO's original charter was simple: enable the use of data in new ways that create value for the district. Specifically, we wanted to find someone who could do the following:

1. Exploit available data to help leaders make better-informed decisions. More importantly, help others exploit their data so employees at every level can make better, data-informed decisions.
2. Assist leaders in enforcing data collection/entry standards (source of standardization rules should be Enterprise level). Assist District in developing/acquiring automated data collection capabilities. Improve data quality and how we measure data quality.
3. Ensure data is available, reliable, consistent, accessible, secure, and timely to support the mission and activities of the agency.
4. Help determine which decisions can be better data-informed. Help decision-makers ask the right questions of their data.
5. Govern FEM implementation from a data perspective. Help determine which data is the most relevant to district operations.
6. Help develop data visualization tools and/or dashboards that bring analyzed data to the users to produce insights.
7. Educate leaders on the use of data. Help develop a sense of data as a corporate asset. Help the district master business intelligence, data management, modeling, metadata management, etc.
8. Help determine how to manage the data we have, and how to digitize the analog data we have.
9. Help determine how mobile apps may assist in automating data collection, storage and retrieval.
10. Decrease the cost of managing data and increasing the value of the data we have
11. Identify innovative ways to acquire and use novel data sets
12. Determine if and how many additional data employees (data scientists, data managers, data engineers, data analysts, etc.) might be needed to maximize the district's use of available data. Coordinate with data partners at universities (nascent Data Science Intern Program).





The Tipping Points

There were several moments that served as tipping points for our Data Revolution: 1) When the Enterprise made the decision to invest in widespread data analytics (including funding for a District Data Officer and data education at the district level), 2) when our employees changed from using human prediction and started using highly reliable machine prediction to make reliable, informed decisions. It wasn't about the technology or the tools. It was about the cultural shift in our workforce that no longer accepted intuition alone, but began on insisting on reliable, analyzed data to make even every-day decisions. This only came about when 3) the right tools were in place to give us answers to questions in real time instead of requiring weeks or months to acquire, clean, analyze, and visualize data to inform decisions. Once data became democratized, there was no stopping the workforce.

The culture of trusting data didn't happen overnight. But once our workforce understood that the data didn't replace human experience and intuition, but actually augmented human judgement, they began to demand more access to data. That's when it really took off!

Our employees increasingly saw value in what they were getting from the new ways of exploring the data. They learned new ways of asking better and better questions, and our data teams responded by giving decision-mak-

ers at every level greater and greater access to the data they needed to be successful. Once leaders saw how well their data-informed employees performed, they "powered down" more and more decisions. No longer were teams of MSC and HQs people reviewing routine work or reports from districts by hand or backward-looking dashboards. Decision making returned to the lowest levels and the ability of higher headquarters to see anomalies that required intervention increased. In short, the Enterprise was able to transform into an organization that trusted its employees to make the best decisions, and had the ability to see the aggregation of those decisions in a new way. In fact, we are returning to a flatter organization in our district because we no longer need the additional levels of management in our empowered, data-driven organization.

Frankly, when I look back on the past

6 years I am amazed at how far we have come. Our district has always been a great place, and has done excellent work. But since the arrival of a Chief Data Scientist at USACE and the first District Data Officer in Little Rock, our productivity, effectiveness, reliability, and credibility has grown every year. I am proud of the workforce we have and the work we do, and am proud to have had a small part to play in the USACE Data Revolution.





▼▼▼ Epilogue

In 1872, John Henry, a “steel-driving man” raced a steam-powered rock drilling machine in constructing a railroad tunnel. John Henry famously won that contest, but never recovered from it and died shortly thereafter. This was a parable of the end of the age of manual labor, and ushering in of the age of the machine. The folk story was bemoaning the demise of the primacy of man and the loss of human dignity as a mere machine began to surpass human ability.

At some point we got past that and learned to harness the power of machines in ways that we didn't feel sacrificed human dignity.

In 2006 the world had another John Henry moment: Gary Kasparov, the best chess grandmaster of all time, faced off with IBM's Deep Blue, and AI-enhanced machine. The 6-game match came to a draw, the first time Kasparov had suffered such a “loss.” Kasparov soon after admitted that the human primacy in chess was over. “I give us only a few years. Then they'll win every match, and we may have to struggle to win even a single game,” said Kasparov who claims he can evaluate 1-3 moves per second. Deep Blue can examine 2-3 million possible moves in that time.

We have leaders and employees who are clinging to our human analytic primacy, and are rooting for the “John Henry with an Excel spreadsheet” analysts to keep us competitive. Some of us refuse

INSIGHTS: AN UPDATE OF THE USACE DATA STRATEGY INITIATIVE



to admit that the steam engine of today is a data lake and a workforce armed with Tableau, R, and Python. Meanwhile the world has gone beyond steam and is leaving us in our tunnels with our hammers wondering how they got so far ahead of us and why we are still in the dark.

This stuff isn't just a Little Rock District fantasy. This technology all exists right now. What we lack is the investment in and commitment of the workforce, the right data ecosystem, and the data governance protocols that will help us transform into a world-class, data-driven public engineering organization.

1. Leadership Support
2. Develop the Workforce
3. Develop Data Ecosystem
4. Data Governance
5. Demonstrate Value

The one big missing factor here is Investment. The new Army Data Strategy outlines an idea of where we need to go, but this is yet another unfunded requirement at the District level. The USACE business model which relies on overhead funds for district-level modernization and innovation is not compatible with a data revolution. Without centralized governance and funding, the 43 districts will limp forward on their own paths, and the Enterprise will continue to struggle with data coherence and suspect data quality.

If United Parcel Service had relied on its thousands of franchises and distribution centers to innovate in isolation, or to separately fund their own implementation of OASIS (UPS's delivery optimization system), they would be the post office instead of the billion dollar industry leader they are today. If USACE relies on project funding overhead to fuel the data revolution, we are unlikely to see great strides within the decade. If, however, we invest in the data revolution as an enterprise, we will see the return on investment quickly. Revolution startup costs may be high, but once established they will be self-sustaining, and the ROI will amaze even the most ardent skeptics.





CONCLUSION

USACE Data must be democratized

USACE's Data Revolution is right in front of the organization, begging to be exploited. USACE must first educate its leaders in the value and techniques of modern data analytics. Legacy data silo practices must be abandoned and the use of individual Excel spreadsheets that hoard data in dark caves must be phased out. Instead, we must democratize data so that every USACE employee can see, understand, and exploit the richness that is in the data. From this data come insights, and from insights come innovation. Only then will USACE's committed employees be empowered to achieve world-class status.

This strategy identifies five lines of effort that contain 16 objectives and six pilot demonstrations. Efforts have begun in all lines of effort and completion of all objectives and demonstrations are slated to be completed by 2025, at which time the tenets of a sustainable data

strategy will be embedded within the workflow of the entire organization.

This strategy will be considered a success when these goals are met:

- There is continuous leadership buy-in on the value of USACE data as a strategic asset.
- There is an adequate workforce within USACE that can analyze data and provide insight for decision makers at all levels;
- Technology is in place to adequately capture, collect, organize, share, and analyze USACE data;
- USACE data is considered authoritative, accurate, and accessible; and
- The value of USACE data to aid in decision making has been markedly proven through pilot demonstrations.

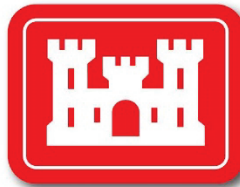
POINTS OF CONTACT

Cary Butler, PhD
USACE Chief Data Scientist
cary.d.butler@usace.army.mil

COL Robert "Bob" Dixon
SWL Commander and District Engineer
robert.g.dixon@usace.army.mil

LEARN MORE

<https://team.usace.army.mil/sites/HQ-CI/PDT/USACEDA/default.aspx>



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) August 2020		2. REPORT TYPE Final		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Insights: An Update on the USACE Data Strategy Initiative; April 2019 Edition				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Cary D. Butler, Martin C. Kittrell, and Robert G. Dixon				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER B72BKG	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Information Technology Laboratory U.S. Army Engineer Research and Development Center 3909 Halls Ferry Road, Vicksburg, MS 39180-6199; Commander and District Engineer (Retired) Little Rock District (SWL), Southwestern Division (SWD) P.O. Box 867, Little Rock, AR 72203				8. PERFORMING ORGANIZATION REPORT NUMBER ERDC/ITL SR-20-6	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Corps of Engineers Washington, DC 20314-1000				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES USACE Data Strategy Initiative					
14. ABSTRACT The U.S. Army Corps of Engineering (USACE) Commanding General, LTG Todd T. Semonite, announced 10 initiatives on May 2018 to improve execution through informed decision-making, enabling lower costs, and world-class engineering results today and tomorrow. Two of the initiatives were geared to transition USACE to data-informed decision-making through the use of data analytics. The first of these two initiatives sought to implement a data strategy that included a doctrine and governance, through creation of a data management plan. The plan implemented tools to aggregate data across the organization and improve reporting. Dr. Cary Butler, USACE Chief Data Scientist, led this effort. The second initiative sought to establish a dedicated USACE Innovations Team to build and recruit a skilled team to act as finders and enablers of innovative, enterprise solutions that enabled USACE to become a digital business. The USACE Chief Information Officer oversaw both initiatives. Because goals and objectives overlapped in many areas and gathered momentum for these initiatives, leaders from both teams came together and decided that showing the business value of data analytics to improve decision making across the Corps should be a top priority. This document was used to report on the progress of the initiatives and give a better understanding of the need for data-informed decisions throughout USACE.					
15. SUBJECT TERMS United States Army Corps of Engineers-- Operations research Data sets—Analysis Decision making Strategy – data processing					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			SAR