



建設ICTマスター養成講座

基礎養成編 選択分野別ソフトウェア実習

ビッグデータ

株式会社フォーラムエイト

 **FORUM 8®**

セミナースケジュール

13:30-14:10 ビッグデータ解析の概要

- BIGデータの概要
- BIGデータの事例

14:20-15:40 UC-win/Roadを用いた交通流データの分析体験

- Rコマンド、R図形、R統計
- 運転LOGデータ、交通流データ 等の分析

15:50-16:20 ビッグデータ解析事例

- 簡単な可視化事例
- UC-win/Road とビッグデータ
- 道路整備効果の算出事例

16:20-16:30 質疑応答

ビッグデータの特徴

ビッグデータとは、「様々な形をした、様々な性格を持った、様々な種類のデータのこと」データの量(Volume)、データの種類(Variety)、データ発生頻度・更新頻度(Velocity)の3要素からなります。

大きさ

- ・TBテラバイト
- ・レコード
- ・トランザクション
- ・表形式、ファイル

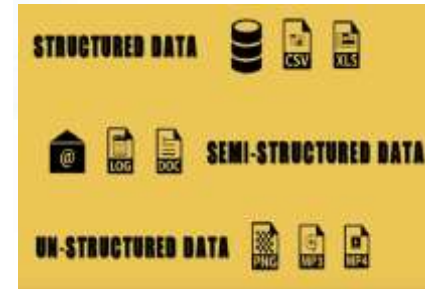
データ自体に意味はない。
大切なのはデータから
傾向・予兆といったストーリー
を見い出すこと。

3要素 の関連

- ・バッチ処理
- ・逐次処理
- ・リアルタイム処理
- ・ストリーム処理
- ・構造化データ
- ・非構造化データ
- ・半構造化データ
- ・上記すべて

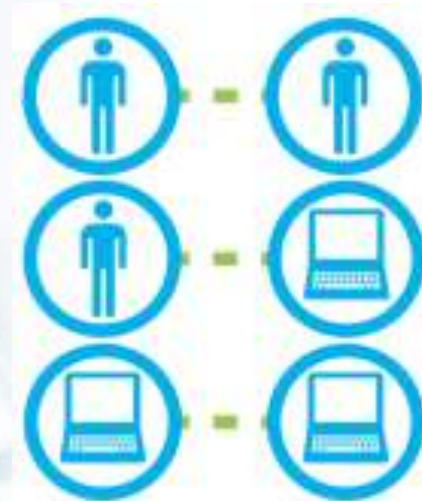
速さ

多様性



ビッグデータはどこから生まれる？

- 人 <---> 人
 - Facebook、e-mail、blog、twitter、コミュニティ、携帯電話、...
- 人 <---> 機器
 - Google検索、Apple Siri、eコマース、パソコン、携帯ゲーム、電子楽器、津波、地震、台風警報システム...
- 機器 <---> 機器
 - CNCコンピュータ数値制御、精密機器生産デバイス、GPS、各種センサー、モニターリングシステム...

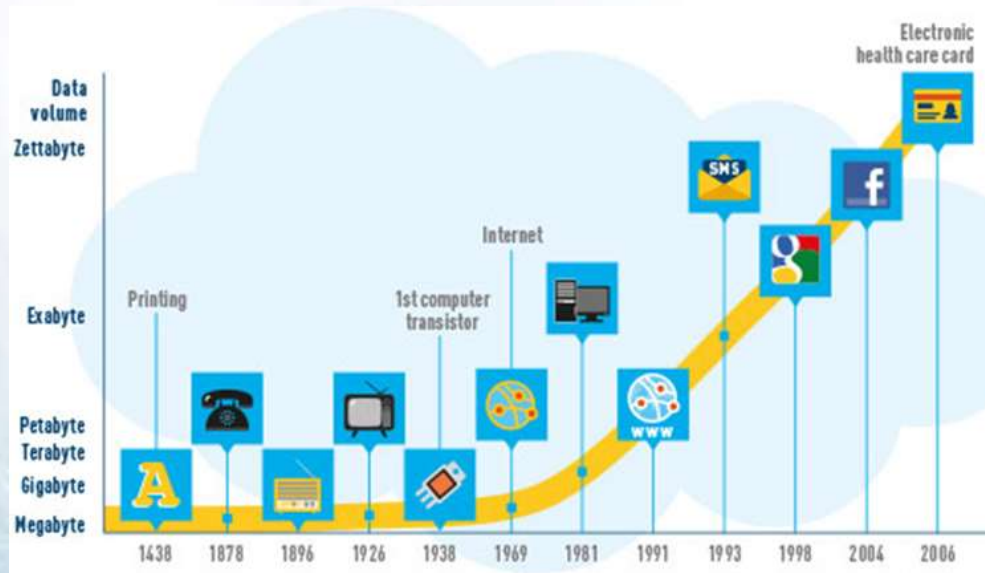


大量のデータ、ますます多くのデータ

ある調査によると、世の中の9割以上のデータは、ここ数年で飛躍的に急増

202x年全世界のデータ量は、44ZBに達する見込

Kilobyte (kB) ^[G 2]	10^3 Byte
Megabyte (MB)	10^6 Byte
Gigabyte (GB)	10^9 Byte
Terabyte (TB)	10^{12} Byte
Petabyte (PB)	10^{15} Byte
Exabyte (EB)	10^{18} Byte
Zettabyte (ZB)	10^{21} Byte
Yottabyte (YB)	10^{24} Byte



大量のデータ、ますます多くのデータ

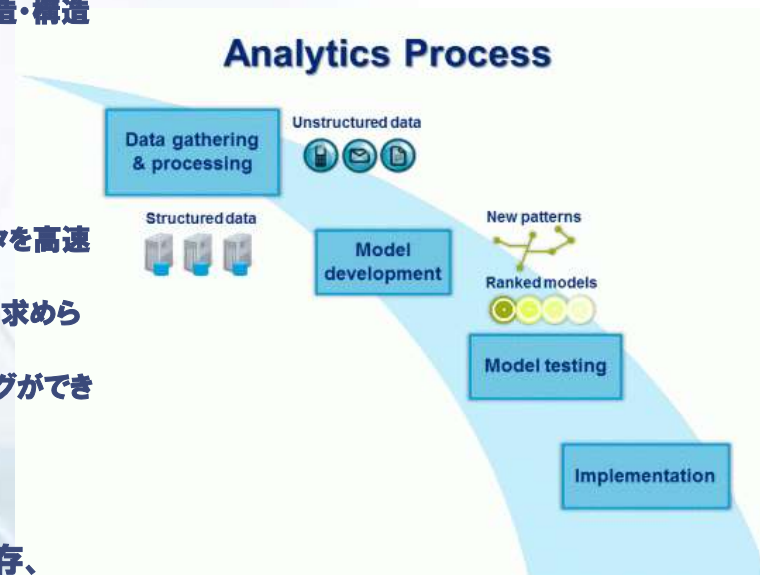
例えば、

- Googleは毎日24 PB超のデータ処理を行っている
(1 PB=1000 TB)
- Facebookは1時間毎に1千万枚以上の写真を受信し、
30億のメッセージ、毎日約100 TBのデータ処理
- YouTubeのクライアントは1分毎にアップロードの映像
は48時間も超えている
- Twitterの通信量も毎日4億件を突破



ビッグデータにある2つの課題の解決策

- データ量が毎日増え続け、保存バックアップが課題
 - 伝統的なライブラリ (RDBMS) は非構造・半構造・構造データの保存に問題がある
- ビッグデータの処理、解析、モデリングが難しい
 - ランダムな非構造・半構造・構造のビッグデータを高速で処理、計算ができる
 - 並列・分散型の処理がサポートでき、スピードも求められる
 - 十分なアルゴリズムでデータ解析及びモデリングができる
- これら課題には、例えば、Hadoopの分散型保存、および、管理システムで有効解決



ビッグデータの事例 Case 1

- これはFacebookが2012年初めてIPO公募の際、公開した画像(Facebookによる通信量を図化)
- 当時Facebookの分析スタッフPaul ButlerがR言語、およびHadoopを利用して作成。



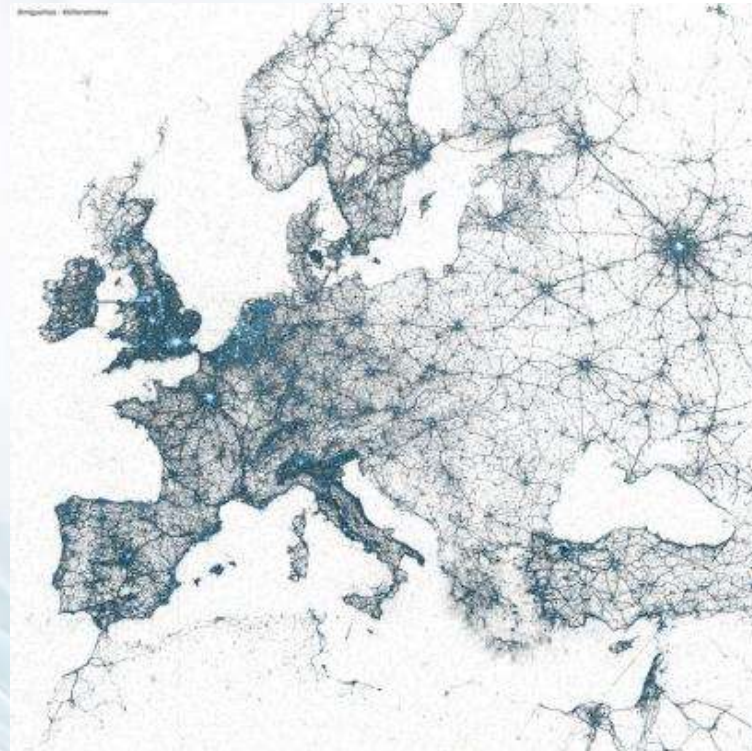
フェイスブック・チャートの説明

- 各自のFacebookは友達を持っている
- あなたの Facebookの友達が住んでいるシティを見つけ出す
- 世界地図上で、あなたの都市から友達の都市の間に白線を引く
- 世界地図上で、一番ハイライト表示されている都市はFacebookの利用者が最も多いことを示している
- この図から、ロシア、中央アジア、中国、アフリカでは Facebookの利用者が少ないことがわかる



ツイッター・チャートの説明

- これはTwitterのエンジニアがRを利用して作った画像
- 各ポイントは2009年からtwitter利用者の地理座標を図示
- 色の濃さはそのエリアでのtweetsメッセージの数を示している。密度は伝言の頻度を表す。
- 海上ルートにもtweetsを利用していることに注目。



ビッグデータの事例 Case 2

どちらの広告が効果的でしょうか？
(Eye Tracking Research)

モデルの目線に注目

ビッグデータの事例 Case 3

ウェブページ、タブレット、スマートフォン情報を見るとき、最初にどこを見ますか？
(eye tracking research)

クリック、ブラウズ、登録、購入、ダウンロードなどの目的…



WEBデザイナー、企業広告は、最も重要な情報を 第一注視点 に配置

ビッグデータの解析ツール R言語の概要 Spotfireの概要



R言語の使用アンケート

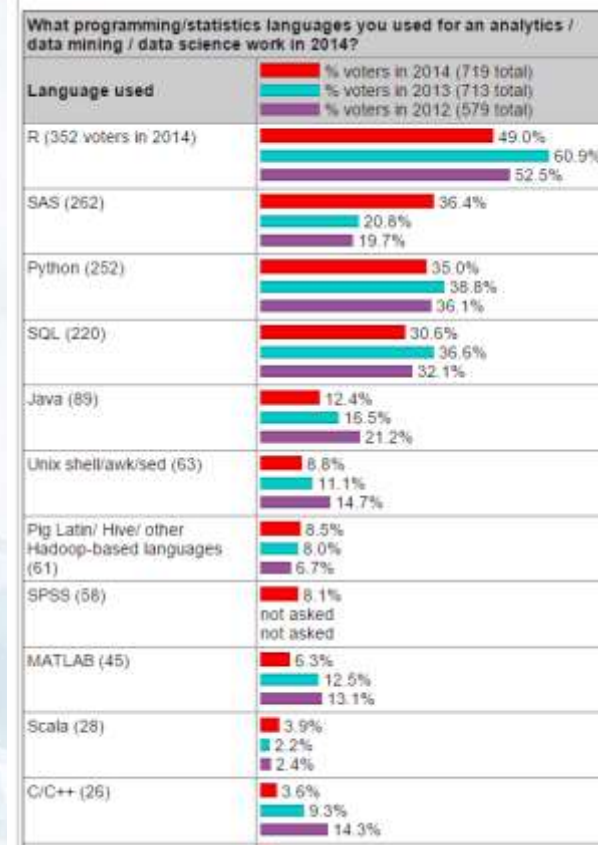
2012, 2013, 2014のアンケート調査で、
Rは3年連続第1位



“What programming/statistics languages you used for an analytics / data mining / data science work in 2014?”

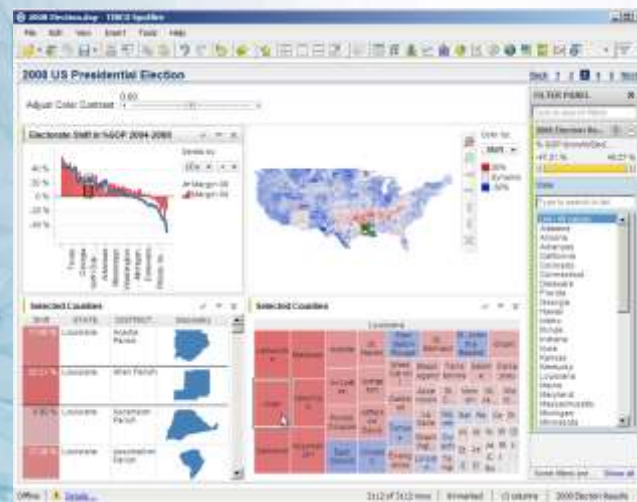
Kdnuggets は、Data Miningのコミュニティサイト。

Data Analyst, Data Miner, Data Scientistに関するソフト、ニュース、活用事例を提供。



Spotfire とは

- データ解析のためのプラットフォーム
- インタラクティブビジュアライゼーション + 統計解析
- BI + BA ソフトウェア (Business Intelligence and Business Analytics)
- TIBCO 社よりリリース



第2単元 操作体験実習

UC-win/Roadを用いた交通流データ の分析体験

- Rコマンド、R図形、R統計
- 運転LOGデータ、交通流データ 等の分析

操作実習資料をご覧ください

第3単元 **ビッグデータ解析事例**

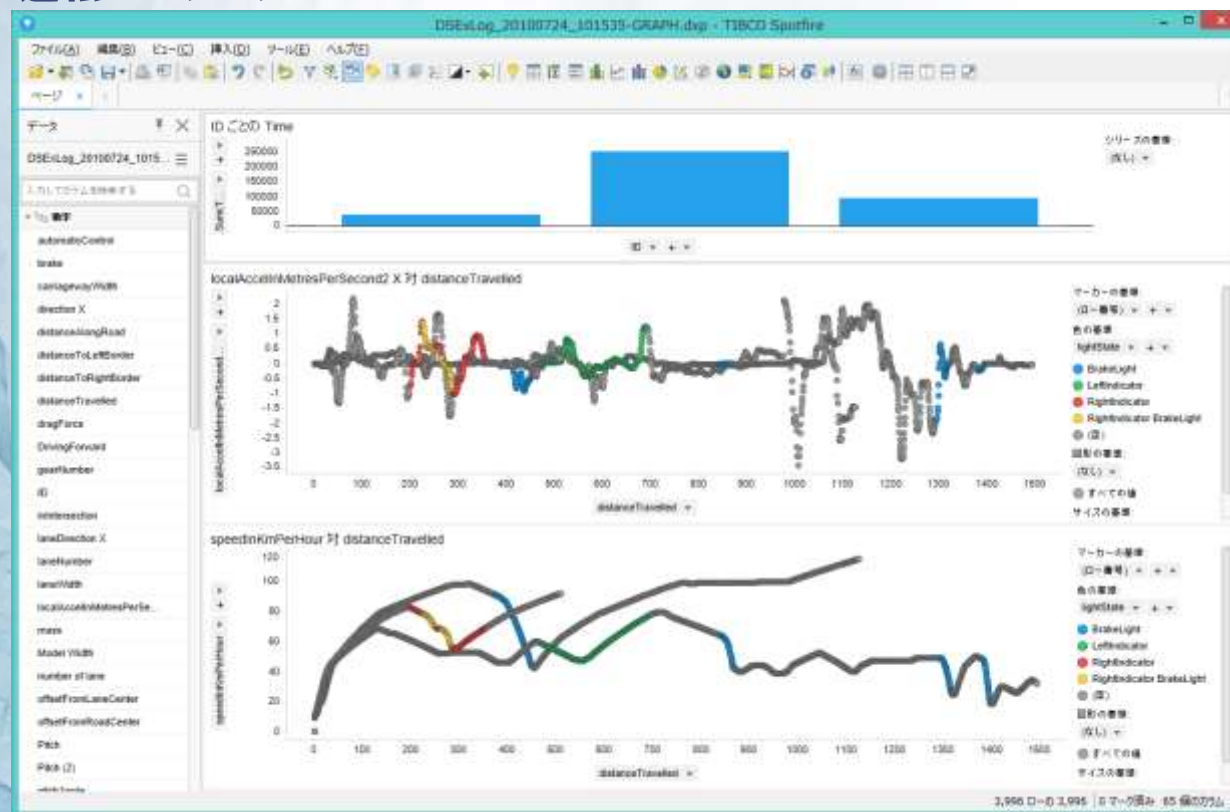
簡単な可視化事例

UC-win/Road と ビッグデータ

道路整備効果の算出事例

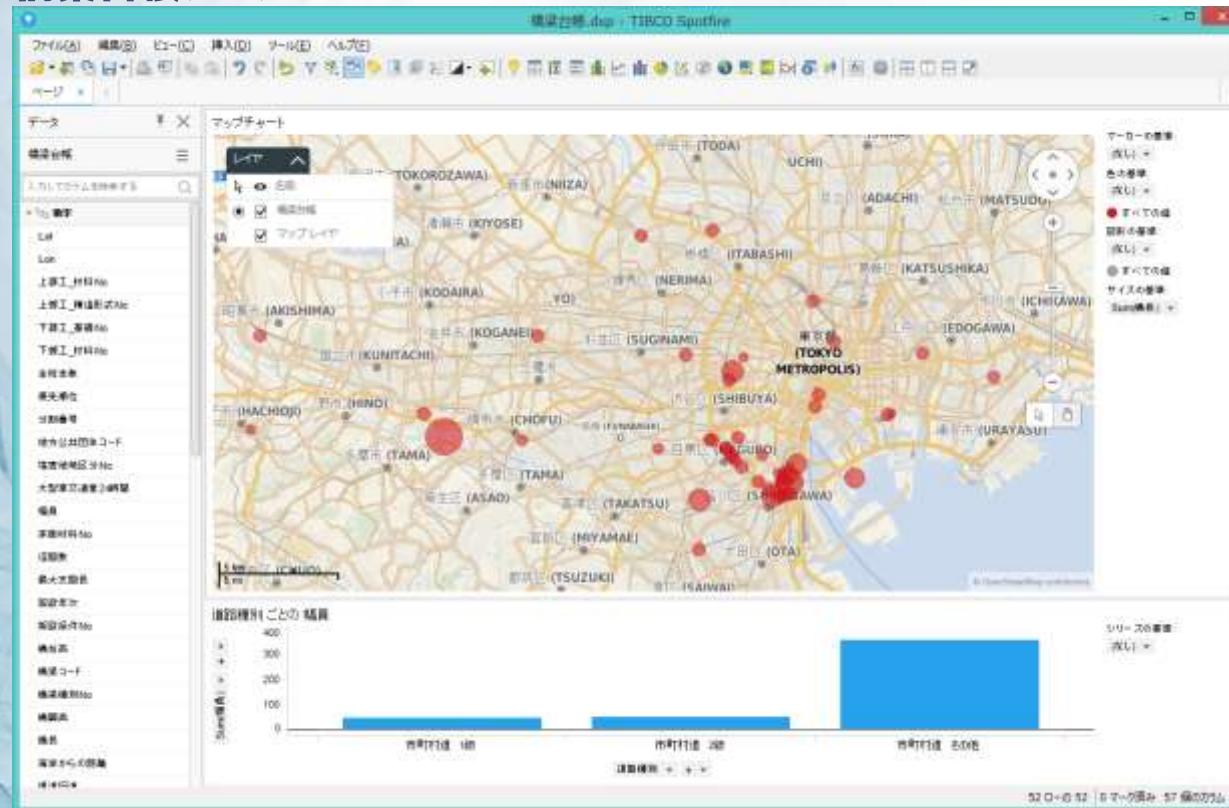
可視化事例（その1）

運転LOGデータ



可視化事例（その2）

橋梁台帳データ



UC-win/Road とビッグデータ

ドライブシミュレータで運転シミュレーションを行った際の運転ログは、膨大なビッグデータとなる。…分析に活用

- 被験者の運転特性分析
- 道路交通の分析



 **FORUM 8®**

道路整備効果の算出事例

UC-win/Roadの交通シミュレーション解析結果を用いて、
道路整備効果の算出を行います。

2車線道路(片側1車線)の交差点に右折レーンを設置した場合の
道路整備効果(便益)を分析



改良前と改良後(右折レーン設置)の解析モデルをUC-win/Roadで作成



ピーク時の交通状況(1時間分)を再現するシミュレーションを実行



LOGファイルに記録

改良前



改良後



道路整備効果の算出事例

＜走行時間短縮便益＞

走行台キロに費用便益分析マニュアル(国土交通省)に示されている車種別の時間価値原単位を掛け、総走行時間費用を改良前、改良後で、それぞれ算出し、その差分(改良前-改良後)が走行時間短縮便益になる。

走行時間短縮便益

	走行時間費用（1日あたり）		総走行時間費用（年間）	
	改良前	改良後	BTo(改良前)	BTw(改良後)
乗用車	13,826	12,260	5,046,478	4,475,061
バス	7,355	6,006	2,684,672	2,192,313
小型貨物	2,186	2,359	797,938	860,962
普通貨物	2,852	2,611	1,040,917	953,086
合計	26,219	23,237	9,570,004	8,481,422

1日あたりの走行時間短縮便益(円/日)

2,982 円

総走行時間短縮便益BT = BTo - BTw(円/年)

1,088,582 円

道路整備効果の算出事例

＜走行経費減少便益＞

走行台キロに費用便益分析マニュアル(国土交通省)に示されている車種別の走行経費原単位を掛け、総経費を改良前、改良後で、それぞれ算出し、その差分(改良前-改良後)が走行経費減少便益になる。

走行経費減少便益

	走行費用(1日あたり)		総走行費用(年間)	
	改良前	改良後	BRo(改良前)	BRw(改良後)
乗用車	5,432	4,809	1,982,805	1,755,465
バス	1,282	1,053	467,797	384,465
小型貨物	763	830	278,595	302,893
普通貨物	1,214	1,132	443,179	413,106
合計	8,691	7,824	3,172,377	2,855,929

1日あたりの走行経費減少便益(円/日)

867 円

総走行経費減少便益BR = BRo - BRw(円/年)

316,448 円

—インフラ分野のDX(Digital Transformation)とは—

<定義>

社会経済状況の激しい変化に対応し、インフラ分野においてもデータとデジタル技術を活用して、国民のニーズを基に社会資本や公共サービスを変革すると共に、業務そのものや、組織、プロセス、建設業や国土交通省の文化・風土や働き方を変革し、インフラへの国民理解を促進すると共に、安全・安心で豊かな生活を実現。

インフラ分野のDXに向けた取組

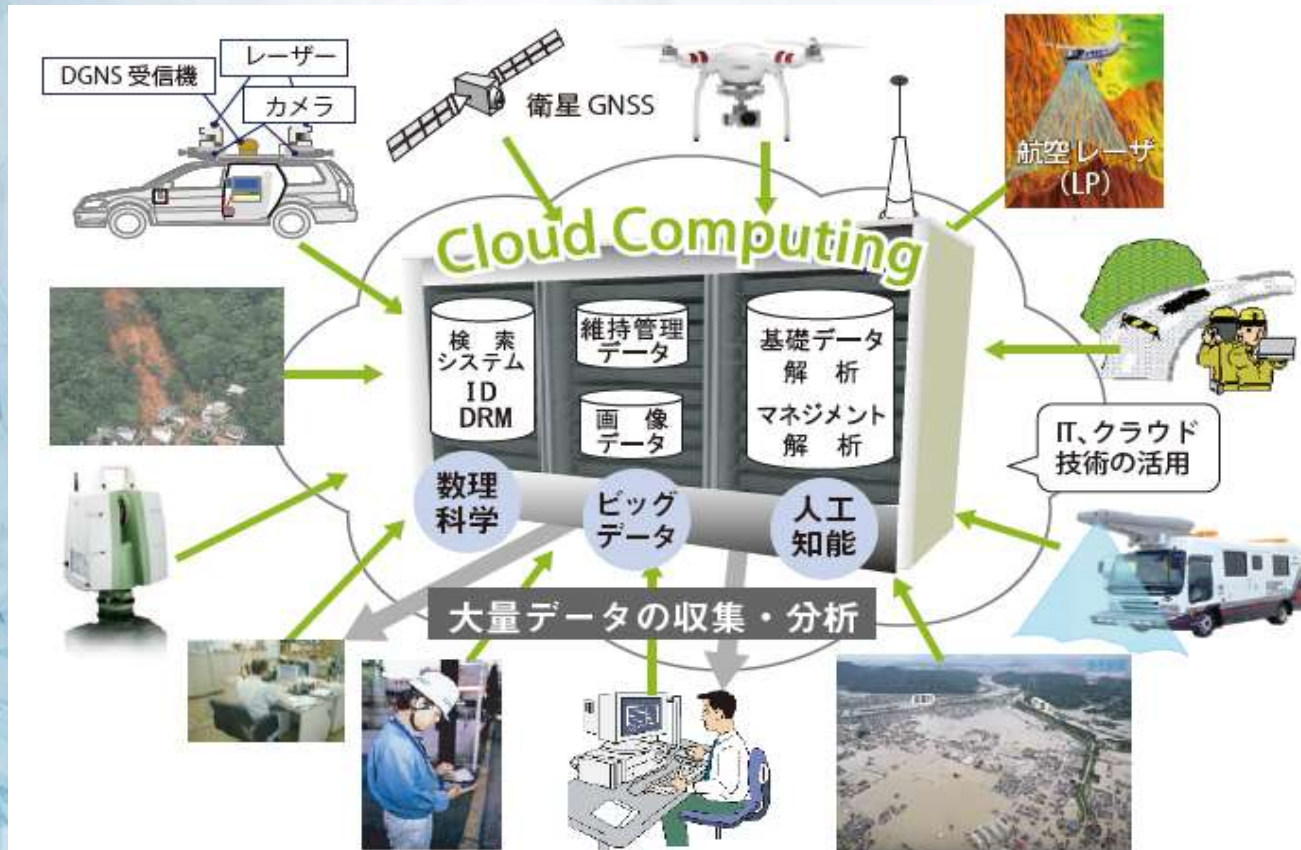


- 新型コロナウイルス感染症対策を契機とした非接触・リモート型の働き方への転換と抜本的な生産性や安全性向上を図るため、5G等基幹テクノロジーを活用したインフラ分野のDXを強力に推進。
- インフラのデジタル化を進め、2023年度までに小規模なものを除く全ての公共工事について、BIM/CIM※活用への転換を実現。
- 現場、研究所と連携した推進体制を構築し、DX推進のための環境整備や実験フィールド整備等を行い、3次元データ等を活用した新技術の開発や導入促進、これらを活用する人材育成を実施。

第1回 国土交通省インフラ分野のDX推進本部 配布資料2より
国土交通省 令和2年9月1日 第4回 BIM/CIM推進委員会 資料1より

—FORUM8 インフラ分野のDXへの取組み—

分析データのクラウド化



—FORUM8 インフラ分野のDXへの取組み—

現場にフィードバックされる



—FORUM8 インフラ分野のDXへの取り組み—

- 道路関連業務にて検索インターフェース上で個別システムまたは各種データに相互に関連付けて利用（登録・検索・表示・印刷出力）できるシステム
- 道路を中心に橋梁・トンネル・舗装・下水道・道路付属物のシステム群に注力
- 収集したデータを分析し、分析データもクラウド化して共有化を行い、最終的には現場にフィードバックされる
- 社会インフラの老朽化、熟練技術者の減少、経費削減や作業時間の短縮、判定結果のばらつきの抑制など、様々な課題解決すべく、DX(デジタル・トランスフォーメーション)のコンセプトに則り、開発・活動を行う



—FORUM8 インフラ分野のDXへの取り組み—

インフラデジタルデータベースを構築し、国のDXセンターとの連携も視野に活動。



インフラデジタルデータベース 維持管理システム(橋梁編)(仮称)

名誉教授ドットコム株式会社との共同開発



Up&Coming 第131号より

ご清聴ありがとうございました。