

# The Opportunity of Extreme Data



**John Weigelt**  
National  
Technology Officer  
Microsoft Canada

@Thumbtackhead



# ENABLING CANADA'S DIGITAL ECONOMY



The screenshot shows the 'Digital Economy Consultation' website. At the top, there is a header with the Canadian flag, 'Government of Canada / Gouvernement du Canada', and the 'Canada' wordmark. Below this is a banner for the 'Digital Economy Consultation' with the URL 'digitaleconomy.gc.ca'. A navigation bar includes links for 'Français', 'Home', 'Contact Us', 'Help', 'Search', and 'canada.gc.ca'. The main content area features a video player with a welcome message from Minister Clement, titled 'Welcome - Consultation on Canada's Digital ...'. Below the video is a transcript section titled 'Read Welcome Message Transcript' with text about the importance of digital technologies and a call to action 'How to Participate'. A sidebar on the left contains links for 'Home', 'Participation', 'Information', 'Rules', and 'Connect'. A right sidebar contains 'Ground Rules', 'Terms of Use', 'Consultation Paper', 'What About Copyright?', and 'News'. At the bottom, a status bar indicates 'This consultation is open for 60 more days.'

Government of Canada / Gouvernement du Canada

Canada

Digital Economy Consultation  
digitaleconomy.gc.ca

Français Home Contact Us Help Search canada.gc.ca

Home » Digital Economy » Homepage

**Home**  
Home  
Log in

**Participation**  
How to Participate  
Idea Forum  
Submissions

**Information**  
What About Copyright?  
Consultation Paper  
News  
Resource Area  
About  
FAQ

**Rules**  
Ground Rules  
Terms of Use

**Connect**  
Stay Connected  
Contact

Welcome - Consultation on Canada's Digital ...

The Honourable Tony Clement  
Minister of Industry

Read Welcome Message Transcript

Digital technologies are critical to every aspect of our economy and society. That is why a strategy for the digital economy is needed to ensure that Canada is positioned to benefit from the opportunities that it presents.

All Canadians have a role to play in helping shape Canada's digital future. Your perspectives, suggestions, ideas and submissions will be important inputs in the creation of our digital strategy. We appreciate your interest and participation.

**How to Participate**

This consultation is open for **60** more days.

**Ground Rules**

**Terms of Use**

**Consultation Paper**

**What About Copyright?**

**News**  
Minister Clement's Welcome Message  
Speaking Notes - Canada 3.0  
Media Advisory  
News Release - May 10th, 2010  
Archives »

Our goal for Canada is to have a world leading digital economy; to be a nation that creates, uses and supplies advanced digital technologies and content to improve productivity across all sectors.

# CANADA IS SPIKEY



Black Tusk 2



SanGatiche 2000



Payton Chung 2008



Dylan Kereluk 2005



The background of the image is a dark, stylized globe. Overlaid on the globe is a complex network of glowing blue lines and nodes. The nodes are represented by small, bright blue circles with concentric rings around them, suggesting signal transmission or data points. The lines are thin and curved, connecting the nodes across the globe. The overall aesthetic is high-tech and futuristic, with a focus on global connectivity and data flow.

PROJECTING EXCELLENCE

A large satellite dish antenna is silhouetted against a vibrant sunset sky. The dish is mounted on a tripod-like stand. The sun is low on the horizon, casting a warm orange glow over a range of mountains. The sky is filled with soft, wispy clouds. The overall scene conveys a sense of global communication and technology.

# GLOBAL APPROACHES



# Transformational trends

Millennials will make up **75%** of the American workforce by **2025**



Over **80%** of new apps will be distributed or deployed on clouds in **2012**.



Digital content will grow to **2.7ZB in 2012**, up 48% from 2011, rocketing toward **8ZB by 2015**.



By **2016**, smartphones and tablets will put power in the pockets of **a billion** global consumers.



The world's **mobile worker** population will reach

**1.3 billion**

—over **37%** of the total workforce—by **2015**.

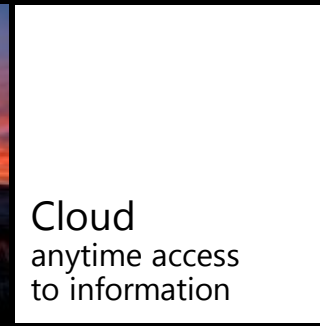
## IT megatrends



Mobility  
ubiquitous  
connectivity



Social  
transforming  
communication



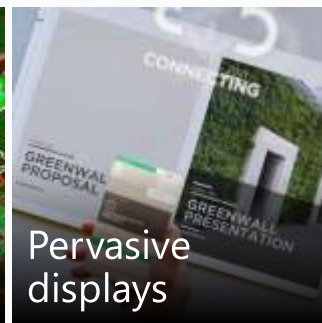
Cloud  
anytime access  
to information



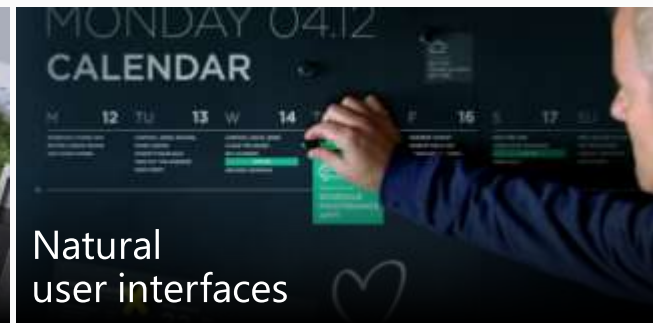
Big data  
enabled by  
machine learning



The computing  
ecosystem



Pervasive  
displays



Natural  
user interfaces

**65%** of companies are deploying at least one **social software tool**.



**70%** of businesses are either using or investigating **cloud computing solutions**



**80%** growth of unstructured data is predicted over the **next five years**.



# The Many Device World

System on a Chip Designs  
Powerful Mobile Devices

Graphics Processing Units  
High Quality Graphics

Explosive Data Growth  
Ubiquitous Sensors and Media

Inexpensive Embedded Computing  
Everyday Smart Objects, CIP

Mobile Device Growth  
Smart Phones and Feature Phones

New Software Models  
Social Networks, Clients + Clouds...



Solar power, white spaces bring 16Mbps  
broadband to towns without electricity



<http://ars.to/XDIL8h>



# THE CHALLENGE

"Cloud" is often misunderstood

The collage consists of three overlapping web page screenshots. The top-left screenshot is from 'techvibes', featuring a navigation bar with 'News', 'Events', 'Jobs', 'Directory', and 'Q&A'. Below the navigation bar, there's a section titled 'Cloud Computing Confusing to...' with a sub-headline 'Vancouver's Soligear Partners With...'. The top-right screenshot is from 'CDN' (www.computerdealernews.com) and 'itbusiness.ca', showing a 'Recent News' section with links to 'Managed Services' and 'Cloud Service'. The bottom-center screenshot is from 'BUSINESS INSIDER', specifically the 'SAI: Enterprise' section. It has a navigation bar with 'Home', 'TechMedia', 'Finance', 'Markets', 'Politics', 'Strategy', 'Entertainment', and 'Advertising'. Below the navigation bar, there's a headline '51% Of People Think Stormy Weather Affects 'Cloud Computing'' by Jay Yarow, dated Aug. 30, 2012, 12:14 PM, with 5,991 likes and 11 comments.

# Microsoft's Datacenter Evolution



[illegible]



# Common questions

- Uncertainty
- Human Resources
- Governance
- Security
- Privacy
- Interoperability
- SLA
- Operations Integration



# Common Myths

- It's all Internet
- All clouds are equal
- You'll lose ownership of your data
- You can't mix and match cloud services
- Providers arbitrarily pile data
- Data Tsunami response to warrants
- Provider staff Cloud Browse
- Law Enforcement browses the Enterprise Cloud
- The Cloud will displace all other technologies

# Getting Comfortable

- Select a service to be provided
- Assess the compliance environment
- Review the current “as-is” environment
- Conduct preliminary PIA & TRA
- Pilot the service
- Assess the risk delta
- Review the SLA
- Build out the business case
- Decide and manage the risk



# A Whole New Perspective on our World



**Imagine Medicine  
in the 1300's**



**Barbers had  
limited tools**



**In the early 1600s  
Galileo and van  
Leeuwenhoek  
began  
experimenting  
with the  
Microscope**



**Remember how  
you felt the first  
time you saw  
images from the  
Hubble  
Telescope?**

141,000 aircraft paths in 24Hrs



Photo credit: Aaron Koblin.



# Ontario Medical Association Turns to DMTI to Develop Province-Wide Routing Table for Physician Incentive Research

## BY THE NUMBERS

- ▶ **142 Billion:** Number of total records (entries in the database) delivered to the Ontario Medical Association; the largest delivery in DMTI's history
- ▶ **280 Years:** Amount of compute time it would have taken to build the routing table prior to DMTI's new routing algorithm being developed
- ▶ **2:** The total tables the Ontario Medical Association (OMA) can now rely on for various applications related to the analysis of physician incentive programs

The Ontario Medical Association (OMA) is a professional organization that represents Ontario's 34,000 physicians. The organization's mandate is to represent the political, clinical and economic interests of the medical profession.

## JUST THE BASICS

- ▶ **Goal:** Build a database for all of Ontario with the shortest drive times and distances between any two postal codes in the province
- ▶ **Challenge:** Previous mapping software and solutions not powerful enough to construct table
- ▶ **Outcome:** DMTI delivered two tables of 142 billion records, providing the functionality to show the shortest distance and time between any two postal codes in Ontario, allowing for complex statistical analysis of physicians and facilities across the province
- ▶ **What they Used:** DMTI Professional Services, Platinum Postal Code™ Suite - Local Delivery Unit Boundaries and CanMap® RouteLogistics

*"DMTI Spatial helped us solve our analysis requirements by producing the data we needed in a reasonable time and at a realistic cost. They turned a project that was previously thought to be unfeasible into a reality."*

Boris Krallj, Executive Director, Economics and Chief Economist, OMA

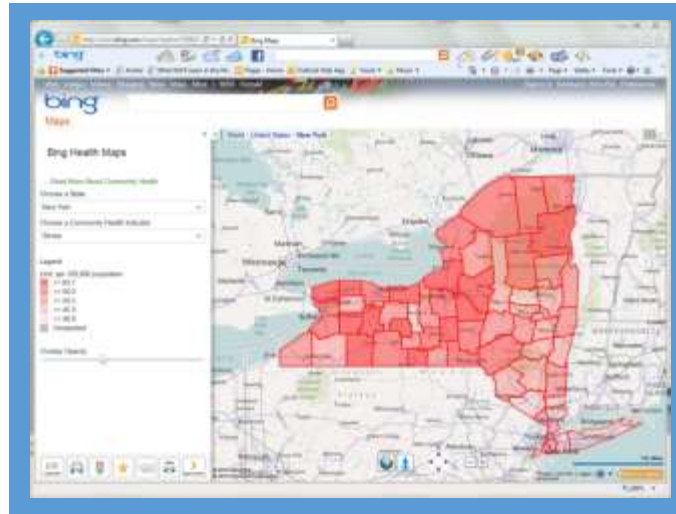




# Advanced Simulation

Predicting and planning for all scenarios

Through the use of large amounts of historical data and computing resources, organizations of all sizes can explore “what if” scenarios

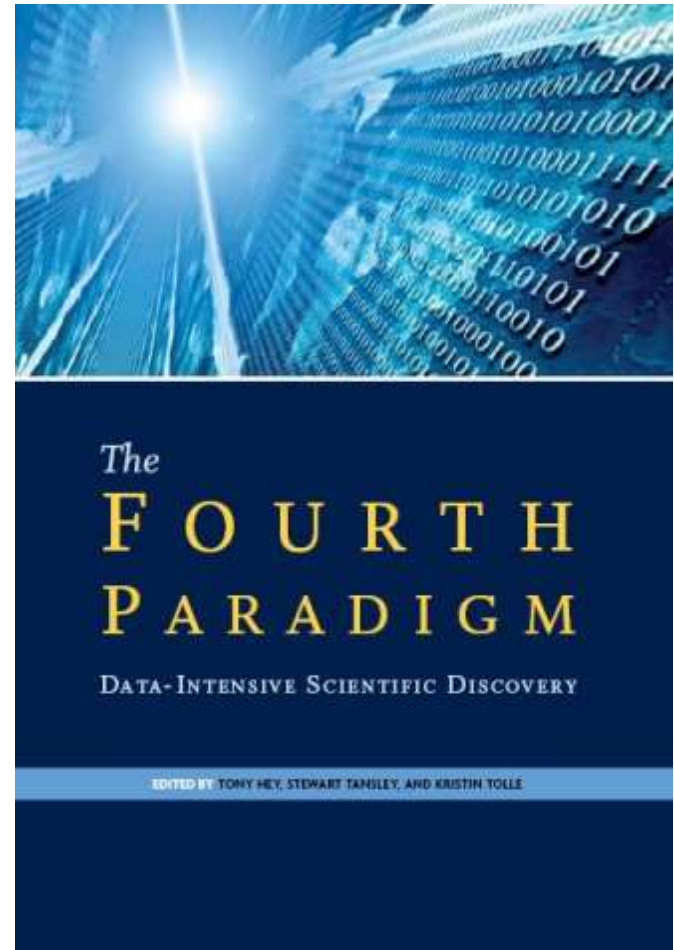


Advancing skills development



# Data Intensive Scientific Discovery

- Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets



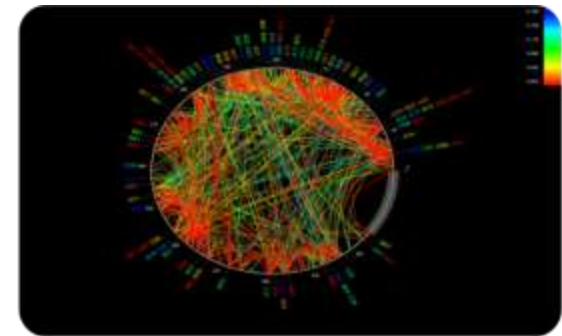
<http://research.microsoft.com/en-us/collaboration/fourthparadigm/default.aspx>

# Harnessing the Cloud to Fight Disease

Data-Intensive Scientific Discovery

PhyloD is Statistical tool developed by Microsoft Research used to analyze DNA of HIV from large studies of infected patients

- 100's of HIV and HepC researchers actively use it
- Typical job: 10 – 20 CPU hours;
- Extreme jobs: 1K – 2K CPU hours
- Large number of test runs for a given job (1 – 10M tests)
- Highly compressed data per job ( ~100 KB per job)



Cover of PLoS Biology  
November 2008



# Open Data

Not only should we share our data, but we should demand that governments and businesses share the data they prepare as well. Accessible raw data is his new objective for the world wide web. Data drives a huge amount of what happens in our lives... because somebody takes the data and does something with it.



Sir Tim Berners-Lee











# EYE ON EARTH

European Environment Agency 

Search  

STATIONS  
water stations



WaterWatch      







## LANDMARKS

- ☐ Community Centres
- ☐ Drinking Fountains
- ☐ Fire Halls
- ☐ Libraries
- ☐ Schools
- ☐ Translink Stops
- ☐ All Impark Parking Lots
- ☐ Metro Impark Parking Lots

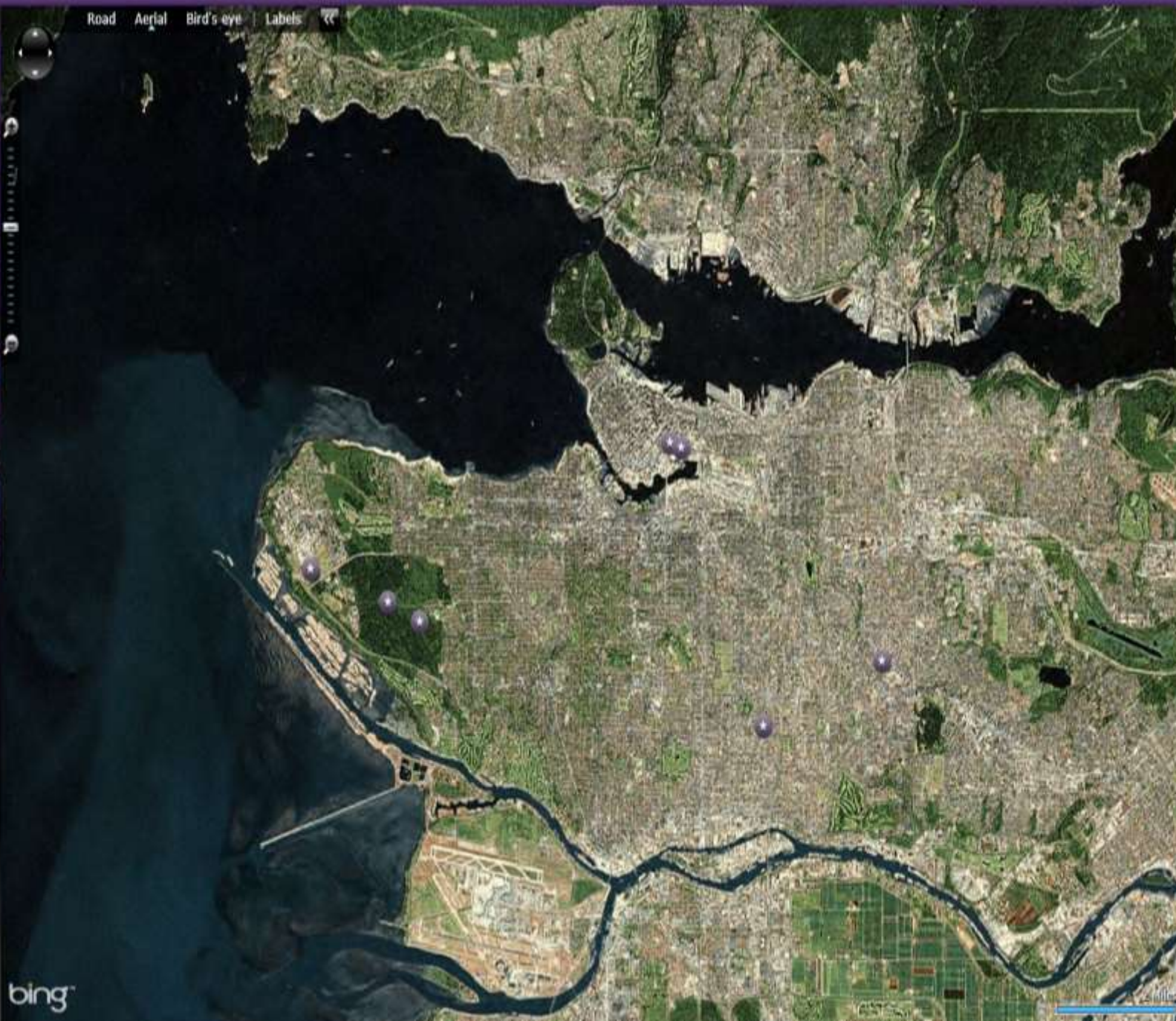
## SOCIAL LANDMARKS

- ☐ Shared Landmarks
- ☐ Nearby Tweets

## MAP ATTRIBUTES

- ☐ Hide all Map Attributes
- ☐ Facet Grid Boundaries
- ☐ Local Area Boundary Data

Road Aerial Bird's eye Labels





# REDUCING POLICY FRICTION



The Microsoft logo is displayed in white, italicized, sans-serif font on a solid blue rectangular background. The logo consists of the word "Microsoft" followed by a registered trademark symbol (®).

**Microsoft®**



John Weigelt  
johnwei@microsoft.com  
thumbtackhead.com  
@thumbtackhead

