Standards, Security and Sharing

Precision Agri-Food Scoping Study and Assessment RFP 4

KEYSTEP *Growth & Finance*

Presentation Outline

- Study Objectives
- Big Data Reference Architecture
- Health Sector - Regulatory Framework, Standards, Data Sharing…
- Banking Sector - Regulatory Framework, Standards, Data Sharing…
- Agriculture - Conceptual Frameworks
- Agriculture - Regulatory Framework, Standards, Data Sharing...
- Data Sharing Agreements
- Conclusions
- Recommendations

2

Authors

Dr. Kevin Lim – Lead Investigator
- Geospatial Data and Systems Design & Implementation Consultant
- Major geospatial projects in defence, forestry and agriculture in the past

Dennis Nazarenko – Project Manager and Technical Research
- Senior consultant focused on emerging technologies and their commercial adoption

Peter Kallai – Project Oversight and Results Analysis
- Expertise in strategic program development for emerging technologies

About KEYSTEP
- 20 years in emerging tech with over 200 studies in various Canadian sectors for leading R&D organizations, universities, industry groups and companies.
- Dozens of projects relevant here in agri-food, spatial data, biotech, big data and ICT infrastructure
- Major public R&D program financing success included $2-$20 million/year

3

Problem Statement

- Data collected, validated, documented, stored and shared in different ways, at different locations across organizations
- Limits our ability to leverage agriculture data for research and other purposes which:
  - Is collected by diverse scientific interests
  - With differing considerations on standards, sharing and privacy
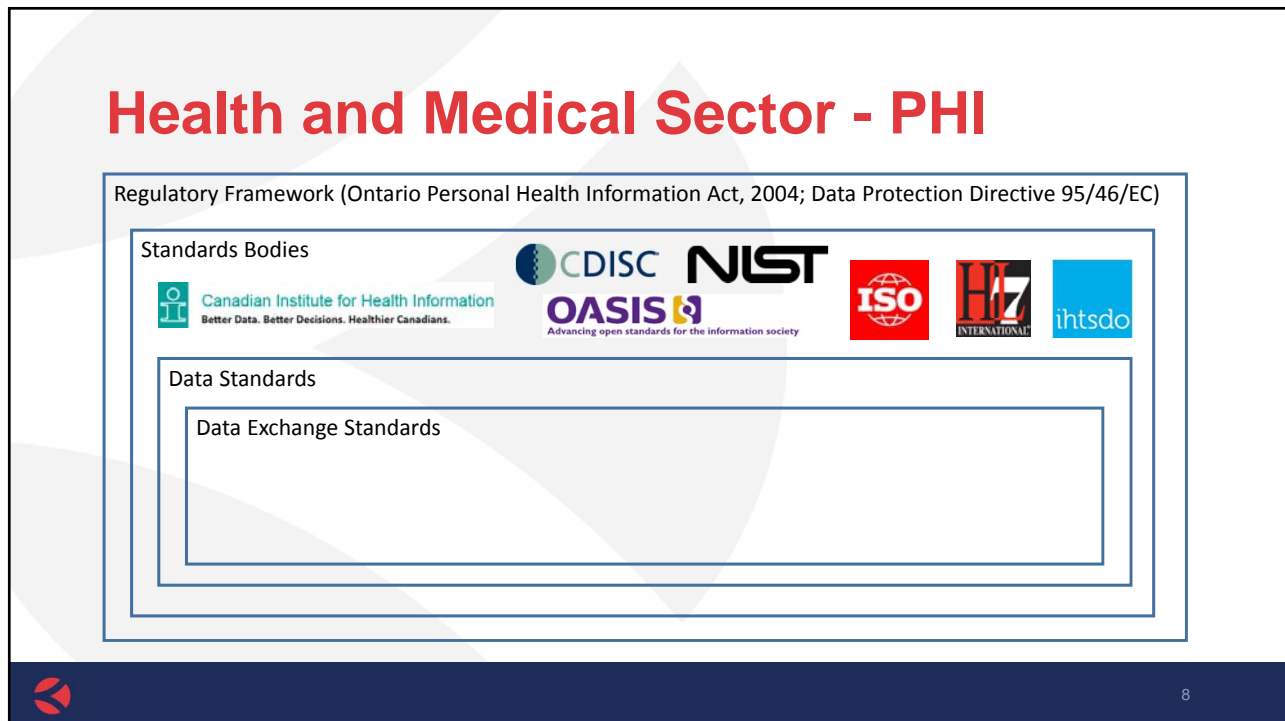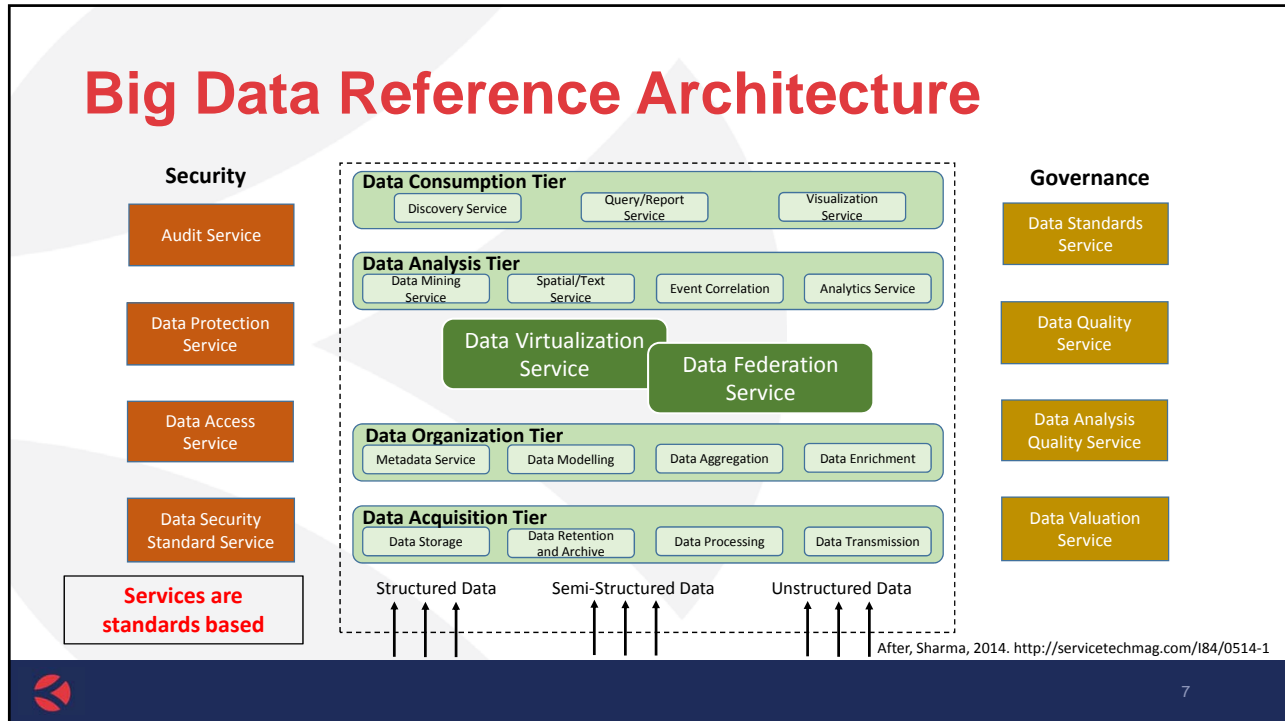 - Is increasing in volume, frequency and locations

4

Study Objectives

1 Understand how Big Data management needs (policies, standards, privacy, secure data exchange, data ownership and other issues) have been addressed by other sectors.

2 Understand how other players in precision agri-food technologies, both commercial and public, are addressing the issue of protocols and standards that protect the security and integrity of data.

3 Recommend a course for the Ontario agri-food research community to adopt as the standard going forward with reference to specific technical protocols.

4 Recommend how to achieve greater uniformity and consistency in data transmission, formatting and storage methods.

5 Recommend how to achieve greater uniformity and consistency in data analysis tools, algorithms and associated software products.

5

Approach

- Non-Agriculture Sectors
  - Survey cross sector initiatives
    - Banking and Finance
    - Health and Medical
  - Assess approach to issues of data standards and security as well as governance and processes that facilitate data sharing
- Evaluate Precision Agriculture Industry
  - Universities, institutions and international programs
  - Relevant reference points

6

Big Data Reference Architecture

Security

Audit Service
Data Protection Service
Data Access Service
Data Security Standard Service

Services are standards based

Data Consumption Tier
- Discovery Service
- Query/Report Service
- Visualization Service

Data Analysis Tier
- Data Mining Service
- Spatial/Text Service
- Event Correlation
- Analytics Service

Data Virtualization Service

Data Federation Service

Data Organization Tier
- Metadata Service
- Data Modelling
- Data Aggregation
- Data Enrichment

Data Acquisition Tier
- Data Storage
- Data Retention and Archive
- Data Processing
- Data Transmission

Structured Data   Semi-Structured Data   Unstructured Data

After, Sharma, 2014. http://servicetechmag.com/I84/0514-1

Governance

Data Standards Service
Data Quality Service
Data Analysis Quality Service
Data Valuation Service

7

Health and Medical Sector - PHI

Regulatory Framework (Ontario Personal Health Information Act, 2004; Data Protection Directive 95/46/EC)

Standards Bodies

CDISC   NIST
Canadian Institute for Health Information
Better Data. Better Decisions. Healthier Canadians.
OASIS
Advancing open standards for the information society
ISO   HL7 INTERNATIONAL   ihtsdo

Data Standards

Data Exchange Standards

8

Health and Medical Sector

Source: Paranjape (2015)

9

Health and Medical - Data Standards

Data Standard	Description
American Medical Association Current Procedural Terminology	Specify codes for the billing of medical procedures.
American Nursing Association	Data sets and terminologies that support nursing practice
CDISC - Controlled Terminology	Clinical Data Interchanges Standards Consortium (standard expressions (values) used with data items within CDISC-defined datasets.
Logical Observations Identifiers Names and Codes (LOINC)	LOINC codes are universal identifiers for laboratory results and clinical observations (e.g., vital signs, outcomes management, and research).
Unified Medical Language System® RxNorm	RxNorm provides normalized names for clinical drugs and links its names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software.
Unified Code for Units of Measure (UCUM)	Code system intended to include all units of measures being used in science, engineering, and business.

10

5

Health and Medical - Data Exchange Standards

Standard	Description
ASC X12 - EDI 837 Health Care Claim	The EDI 837 transaction set is the format established to meet HIPAA requirements for the electronic submission of healthcare claim information.
Digital Imaging and Communications in Medicine (DICOM)	DICOM is the standard to create, store and transmit medical images and associated information between systems.
HL7 Version 2.x Messaging Standard	Defines a series of electronic messages to support administrative, logistical, financial and clinical processes.
HL7 Version 3 Messaging Standard	In response to difficulties with lack of clarity in the HL7 v2 standard, HL7 created a new standard based on a strict modeling methodology and XML encoding, with robust built-in support for terminology.

- Various types of EDI systems are implemented depending on the organization's business needs

11

Precision Agri-Food

Regulatory Framework- no framework for data ownership, no framework for privacy, no framework for security

Standards Bodies

OGC® Making location count.   ISO   NIST

Data Standards

Data Exchange Standards

Emerging Shared Systems Addressing Data Sharing, Privacy, Standards

Future Internet Public-Private Partnership Programme (2010 -)

Projects being leveraged for precision agriculture in Europe

SmartAgriFood (5 M€)
40 partners
http://smartagrifood.com

FIspace (13.5 M€)
Business Collaboration
28 partners
http://www.fispace.eu/

FIWARE accelerators (17 M€)
Service Delivery Framework
http://www.fiware.org/

Sjaak Wolfert, 2016

13

EU SmartAgriFood - Conceptual Cloud Architecture
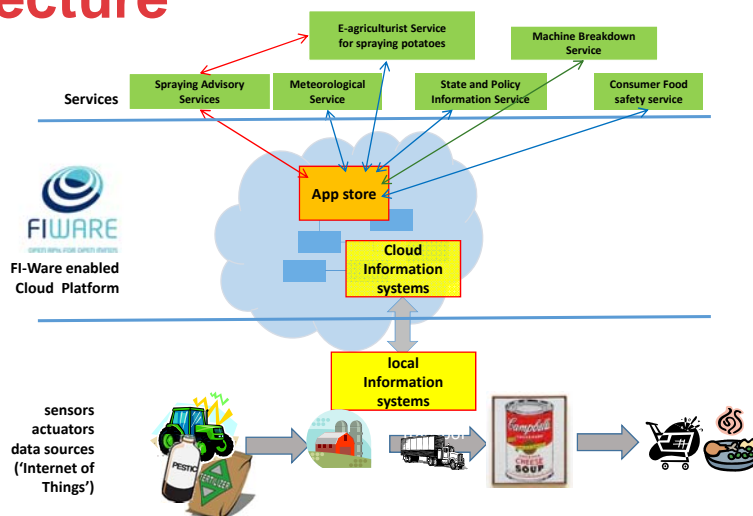
14

Infrastructure Components
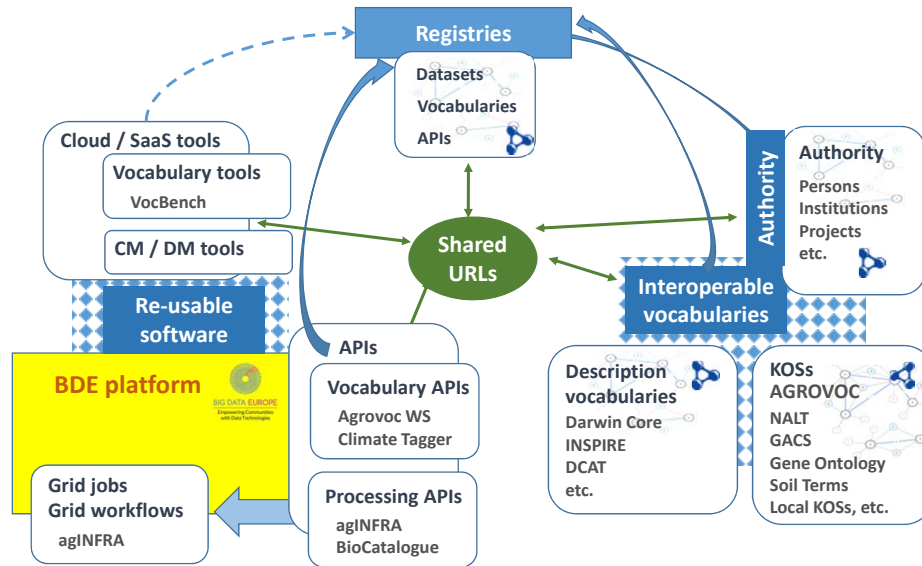
Pesce, 2015

15

Precision Agriculture – Data Standards

Data Standard	Description
W3C-Standards	Rule Interchange Format (RIF), Web Ontology Language (WOL), Simple Knowledge Organisation System (SKOS) and Resource Description Framework (RDF) are suitable for the interchange of rules and a unified semantic between systems.
ISO-Standards	ISO 19115 describes the composition of a meta data set; other ISO standards also have relevance to precision agriculture
OGC-Standards	GML (Geography markup language) describing features with geographic properties WMS (Web map service) provide presentations of geodata (maps as images) WFS(-T) (Web feature service [transactional]) offers access to vector data WCS (Web coverage service) defines interfaces and operations to request data excerpts from a specific area. WPS (Web processing service) executes GIS functionalities or algorithms to geodata SWE (Sensor Web Enablement) defines a suite of web service interfaces and protocols for sensor communication

16

Data Sharing Agreements – Framework

Data Infrastructure

Data Life Cycle

Data Management Plan

Data Sharing Agreement

Source: Backlund Jarquín, 2012
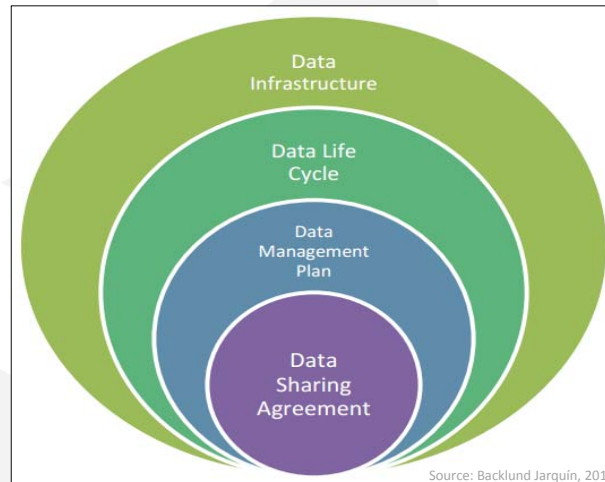
Data Sharing Agreement within the Life Cycle of Research Data Management

17

Data Sharing Agreements - Key Elements

Content Area	What to Consider
Title	• Examples: Memorandum of Understanding, Business Agreement, Data Sharing Agreement
Data Overview	• A description of the information to be gathered, used or made available • The nature and scale of the data that will be generated or collected
$$/Resources	• Funding/compensation requirements • Other resource sharing and allocation
IP/Legal	• Copyright, Intellectual Property rights. • Ethics, Privacy, Security, Levels of confidentiality • Other legal requirements
Ownership	• Ownership of the data
Storage	• Data archiving, security, long-term preservation
Access	• How the data will be made available to various groups (ie research community, wider community, etc)
Approval	• Approval process for access to data (including levels of access)

18

Data Sharing Agreements – Key Elements

Content Area	What to Consider

Who/Roles	• Who should participate in agreement, stakeholders, data steward?
	• Who has responsibility for:
	   • Data collection, storage, management and analysis
	   • Access decisions, ownership

Training	• What training will be available to all partners? • Identify training needs and requirements for all partners
Analysis	• Who will analyze the data?
Dissemination	• Determine dissemination procedures • Means of release of the data • Mechanism for interim findings release prior to final data set completion
Timeline	• Data archive maintenance, data collection, when will data be disseminated • Agreement length

19

Conclusions

- Health care, banking digitization and sharing have occurred. Agri-food digitization and sharing are behind, but …. coming
- Health care and banking sector data frameworks with policy, governance and standards can inform Agriculture
- Strategic precision agri-food programs and projects are being developed in Europe with multi-country projects
- Examples and protocols exist that can be applied to an Ontario initiative
- Need an IT systems approach, dedicated resources that can develop ICT, policies, standards, besides the application program expertise
- Need to define an entry point, application: (1) Share current research data? (2) Develop a new shared system?

20

Recommendations

- Develop a program and resources with an overall systems approach to include standards, security and data sharing frameworks. Develop shared systems for Ontario.
- All standards, data access, privacy, exchange mechanisms, and governance concerns should be defined as part of the program & infrastructure design process:
    - Define metadata standards based on implementation focus
    - Adopt a Service Oriented Architecture strategy
    - Adopt recognized Web Interface standards for web services
    - Define an Identity and Trust Model
    - Encrypt Sensitive Data
    - Implement Sensor Web Enablement standards
    - Define Data Exchange Standards
    - Leverage best practices in Data Privacy and Security
    - Implement Third Party Audits of Cloud Services
- Consider the feasibility of collaboration with European initiatives to leverage their approaches to Big Data governance and standards.

21

Acknowledgements

Thank you to these organizations and their researchers and staff for input and advice.

Acknowledgements