Linked VISTA

A Metadata-Centric Approach to VA Master Data Management

February 24, 2015
Pacific Joint Information Technology Center
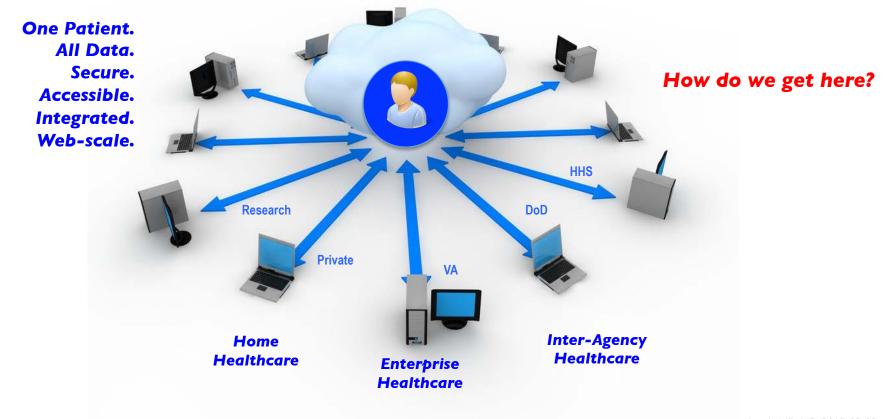*Maui, HI*

Rafael Richards MD MS

VHA Office of Informatics and Analytics
U.S. Department of Veterans Affairs

Vision: **Linked VISTA**

A Linked Data Approach to Patient-Centric Care

One Patient.
All Data.
Secure.
Accessible.
Integrated.
Web-scale.

How do we get here?

HHS

Research

DoD

Private

VA

Home Healthcare

Enterprise Healthcare

Inter-Agency Healthcare
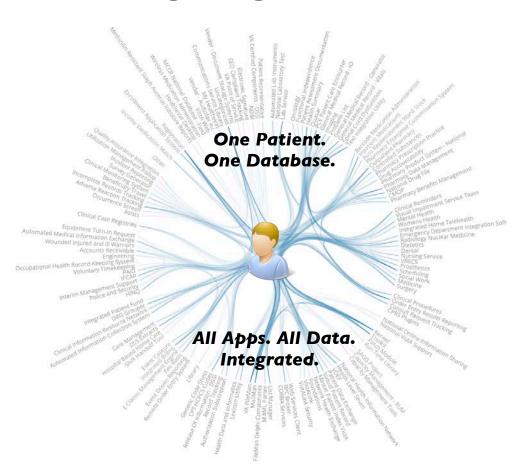
.ichards MD MS 2015-02-23

VISTA Overview

- *VISTA Architecture: Single-integrated EHR*
- *VA Enterprise: One Patient. Many VISTAs*
- *Problem: Isolated VISTAs. Fragmented data.*
- *Approach: Linked VISTAs.*
- *Vision: OneVA.*

A Patient-Centric
Single integrated EHR

**One Patient.
One Database.**

**All Apps. All Data.
Integrated.**

Patient-centric
Health Record

The data architecture of VISTA consists of over 160 modules for clinical care integrated within a single common multidimensional data engine (MDE).

In VISTA, both business logic (Applications) and data (Database) are managed within the multidimensional data engine. This provides the tight integration of applications to data, and to a single common integrated database.

The integration between VISTA applications (outer ring) and shared data (inner circle) is visualized, showing the shared data flow between applications.

VA Enterprise: *Many VISTAs*

1200 VA healthcare facilities throughout the US Supported by 130 VISTA systems

Care coordination for patients within VA requires integration of patient data across all 130 VISTA systems.

Problem Statement

VA is comprised of 130 unique VISTAs.
This fragments patient data and care.

RM Richards MD MS 2015-02-23

VISTA Silos

**One Patient.
Many Databases.**

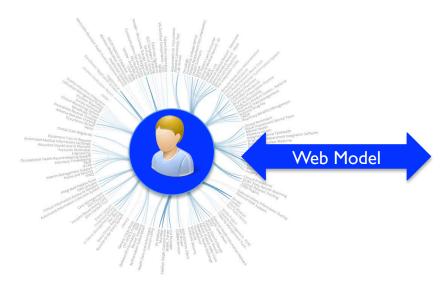
**Many VISTAs.
Many Models.
Fragmented Data.**

While each VA VISTA system is individually highly integrated, each system has its own distinct data model (shown as different color).

Different data models creates data silos — fragmenting patient data and fragmenting patient care.

Approach: Linked VISTA

Exposing VistA's Data Model securely as a Web-standard model enables web-scale integration.

Web-standard Integration requires:

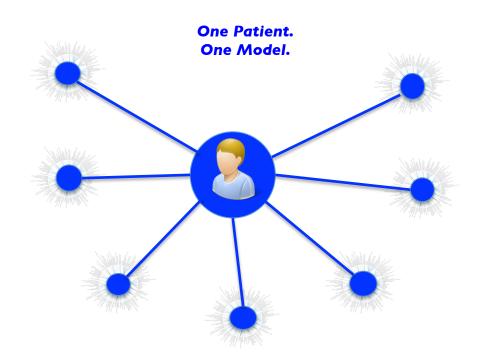**Common Data Model.**
Data Security.
Data Services.

A model-centric approach to VISTA data modernization would provide comprehensive granular, metadata driven access and security to all data.

This would allow publishing, linking, and exchanging, with patient-centric security, granular data with internal and external systems with international Web standards.

Linked VISTA

One Patient.
One Model.

Many VISTAs.
One Model.
Integrated Data.

Exposure and cross-linkage of VISTA's local models to a common VISTA model provides an integrated view of all patient data from all VISTA systems.

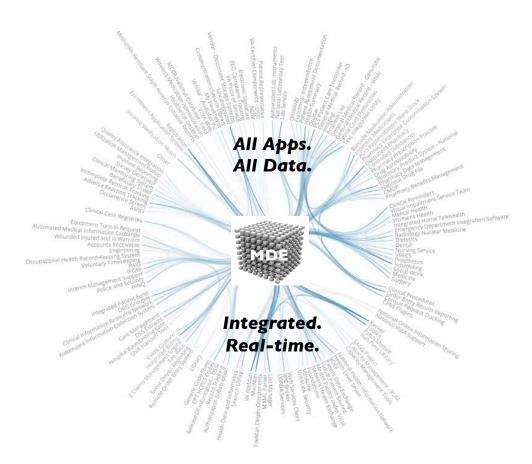# **OneVA**

VISTA Data Model (VDM)

A Path to VISTA Data Management:

- VISTA's Database
- VISTA's Data Model
- VISTA's Data Model Exposed
- Benefits of Leveraging VISTA's Data Model:
    - Master Metadata Management
    - Centralized Knowledge Management
    - Master Data Definition
    - Patient-Centric Security Model
    - Separating Business logic from Data
    - Query Access

Review: VISTA's Database

The foundation of VISTA is a high performance Multidimensional Data-integrated Application Engine in which all data and all applications are fully integrated in real-time with each other and to one single authoritative data source.

**All Apps.
All Data.**

MDE

**Integrated.
Real-time.**

VISTA's integrated application data engine. All 160+ applications are integrated with their data and logic inside the multidimensional data engine (MDE). This keeps transactional patient data and logic highly integrated for real-time use within one single data store.

VA uses the same healthcare industry-standard data engine that nearly 50% of hospitals in the US currently use as their core EHR database.
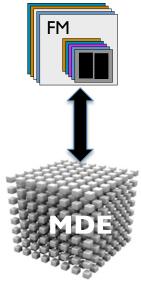
This same multidimensional data engine is also used as the real-time transaction engine of the world's five largest banks, the largest Wall Street trading systems, and over half a million ATMs in North America.

All real-time transactional operations in VISTA take place within the the multidimensional data engine (MDE). To provide consistent structure and model for the data, a data dictionary driven hierarchical data storage system is overlaid on top of the MDE. All VISTA applications read and write data to this hierarchical store using a file manager called Fileman.

File Manager

FM
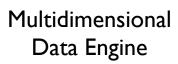
Hierarchical Graph Store

This layer provides the data structures, data models, and data management for VISTA. Fileman is the DBMS for this hierarchical store.

Apps Data

MDE

Multidimensional Data Engine

This is the data-structure flexible transactional core of VISTA. This core is a healthcare industry-standard for nearly 50% of the hospitals in the US. This is also used as the real-time transactions engine of the world's five largest banks, the largest Wall Street trading systems, and over half a million ATMs in North America.

Review: VISTA's Data Model (VDM)

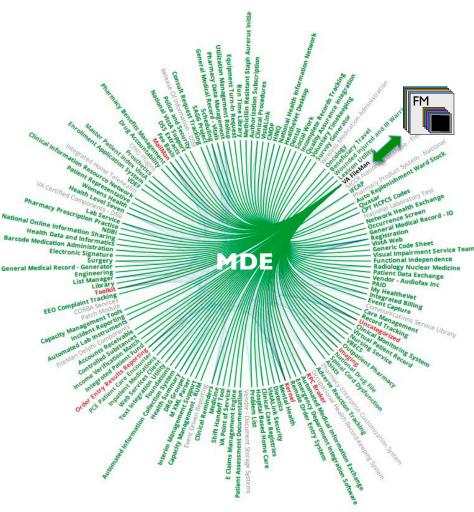All VISTA applications read and write data to VISTA's data engine through the NoSQL hierarchical data manager, Fileman (FM). Unlike many NoSQL databases, VISTA's database is not schema-less, but schema-driven. VISTA's schema is self-documenting through data dictionaries in Fileman. *Surprisingly, this fact has not yet been leveraged to expose and modernize the data model.*

All VISTA data is organized through its data model, managed by Fileman.
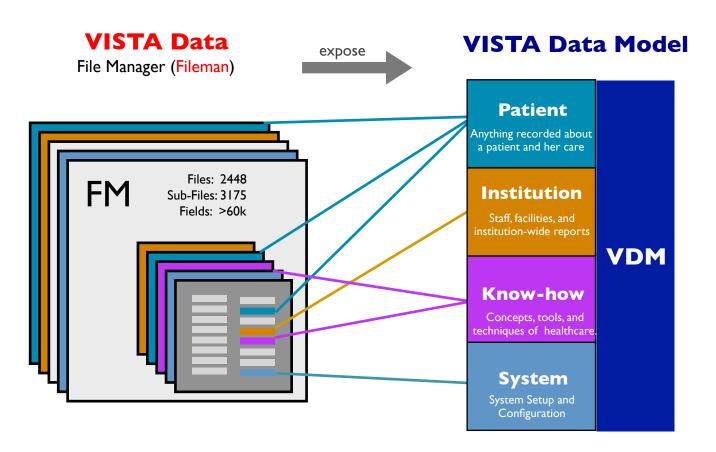
All VistA applications use Fileman to access, structure, and query all data in VistA.

The green lines represent the data flow of all VISTA application to Fileman for all read and write operations to the Multidimensional Data Engine (MDE).

RM Richards MD MS 2015-02-23

VDM: Expose and Leverage the Model

The first stage of data modernization is to expose and leverage VISTA's real, live operational data model. Since this is just metadata, there is no patient data involved. Unlike many NoSQL databases which are schema-less, VISTA's NoSQL database model is self-documenting through Fileman. This allows us to render this in a standard definition format.   In this new web-standard medium, data can be sorted, tagged, searched, and organized by data categories such as by patient, institution, know-how, or system information.

VISTA Data
File Manager (Fileman)

expose →

VISTA Data Model

FM

Files: 2448
Sub-Files: 3175
Fields: >60k

Patient
Anything recorded about a patient and her care

Institution
Staff, facilities, and institution-wide reports

Know-how
Concepts, tools, and techniques of healthcare.

System
System Setup and Configuration

VDM

Access: Thousands of RPCs, API's, HL7
Model: Unknown

Access: Single query access
Model: Consistent, Transparent

RM Richards MD MS 2015-02-23

VDM: Master Data Management

A benefit of a VISTA data model allows one to manage data logically across all VISTA application boundaries independent of the source of the data. This lets one manage data with much flexibility, including logically partitioning and managing the data using metadata tagged categories (such as Patient, Institution, Know-how, and Systems information).

VISTA Data Model

- *One may logically partition data by any class of data*, such as Patient data, Institution data, re-usable Know-how, and System configuration data.

- *One may extract and move all patient data from system to system with one operation*, making system configuration migration and patient record movement far more efficient.

- *One may apply security metadata or protocols to any of these logical classes* of data. For example a patient-centric security model for patient data, and thus enforce patient-centric controls on information exchanged.

Patient
Anything recorded about a patient and her care

Institution
Staff, facilities, and institution-wide reports

Know-how
Concepts, tools, and techniques of healthcare.

System
System Setup and Configuration

VDM

Benefits:

Patient Data Management
Extract and manage patient data with patient-specific security and metadata, allowing patient-centric controls on data access and exchange.

Institutional Data Management:
Institution specific data can be exchangeable and centrally manageable

Knowledge Management:
Common medical concepts, standards, and know-how may be identified and managed as a clearly defined class of VISTA data.

System Management:
The entire configuration of a system can be identified, extracted, and transported, and inserted from system to system

VDM: Patient-Centric Data Security

An exposed VISTA data model allows one to tag and partition certain classes of data separately from all other data in VistA. Specifically, this allows one to granularly partition any and all Patient data from all other kinds of data in VistA. This provides true, direct, "on the metal" security on patient data itself.

VISTA Data Model

- The most important class of data to apply security is the Patient data category. This will allow very granular patient-centric security on the data itself.

- Current VISTA security is based on actions one can take using a legacy menu system - which has no relationship to the data.

- With a data-centric security model, we can specify not just what type of data ("Mental Health Record"…) but whose data ("For patient X").

- This is much more specific and secure than the prior VistA security model.

Patient
Anything recorded about a patient and her care

Institution
Staff, facilities, and institution-wide reports

Know-how
Concepts, tools, and techniques of healthcare.

System
System Setup and Configuration

VDM

Patient-centric security model
Extract and manage patient data with patient-specific security and metadata, allowing patient-centric controls on data access and exchange.

This is compatible with security notions in Meaningful use Data Exchange: it can suppress even data that exists if there is no access permission.

It is essential to improve precision in data security to permit access to VISTA data securely. Otherwise one will have to reverse-engineer 3300 legacy RPCs and their one-off use of Kernel's menu options for each payload.

A patient-centric security model is much more appropriate, flexible, and secure as a foundation for patient data security than the current VistA security model. The current VistA security model provides security only _indirectly_, through legacy controls of a _menu system_ for a legacy roll and scroll terminal interface – which has nothing to do with the type of data at all (!).

VDM: *Analytics Driving Interoperability*

Comprehensive exposure and analytics of the VISTA Data Model will drive enhanced data use and interoperability as well as a major improvement in the structure of the database itself. To address these and other areas, focused reports could be generated from the model including:

| Report	Activity	
Inconsistencies between VistA data models	Drive dictionary and code fixes in various centers so that every center is running the same consistent model	→ **Enterprise Data Model**
Isolate centrally and locally managed "know-how"	Enables the next generation of enterprise knowledge services that seamlessly synchronize VistAs and other applications	→ Centralized Knowledge Management
Under-definition in the model	Too many ill-defined string values and not enough nuance ("zip code", "telephone number") can be defined, and provided additional metadata ("home", "work", "mobile") leading to a plan for incremental dictionary improvement	→ **Enterprise Data Definition**
Key logic performed within FileMan	The barrier between the data store and business logic will be laid bare. This report will encourage the movement of certain types of logic into FileMan and out of less maintainable procedural code.	→ Clean separation of business logic from data
Overlooked but highly valuable types of patient data (Some of this may have been hard-set by the application logic, and overlooked by Fileman or DD)	Improve VistA Data mining (for CDW etc.) and interoperability (more comprehensive electronic patient records). Without a complete model, how do you know what you're leaving behind?	→ Clinical Research → **Interoperability**
VistA model/ FHIR comparison (key types)	Show how a direct from VistA transformation can remove the need for redundant intermediate, hard to maintain procedural code	→ Accelerate Data Exchange
Isolate patient from other types of data	Enable patient-data access control rather than the crude option/API security now in VistA	→ **Patient-centric Security Model**

Linked VISTA
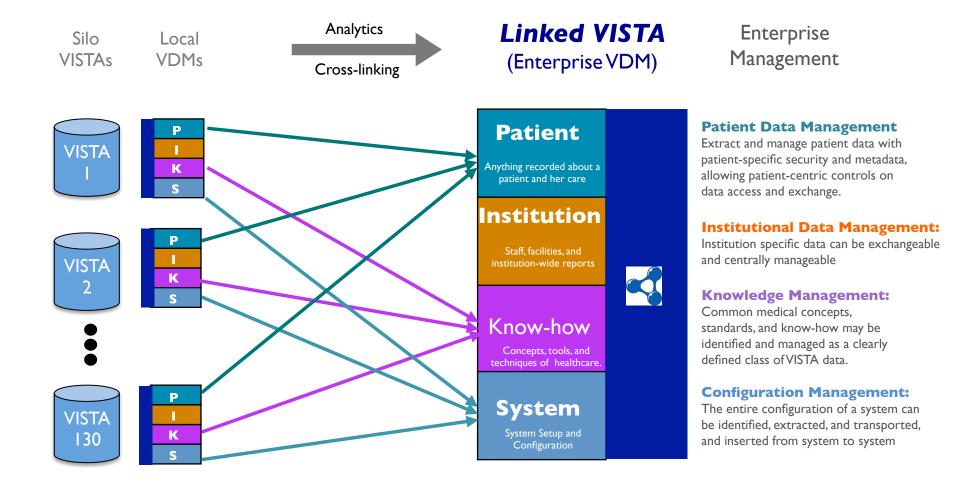
Features

- *Enterprise Cross-linked VISTA Model*
- *Comprehensive Data Access*
- *Web-standard representation language*
- *Web-scale semantic integration*
- *Knowledge Discovery*
- *Native terminology integration*

Linked VISTA: The Enterprise Cross-Linked VISTA Data Model

Exposing VistA's Data Model (VDM) and applying analytics allows one to cross-link all local VDMs to an enterprise VISTA data model (Linked VDM), providing the capability for Enterprise query and Enterprise data management. This leverages the capability of Linked Data to create a cross-VISTA (enterprise) data model.
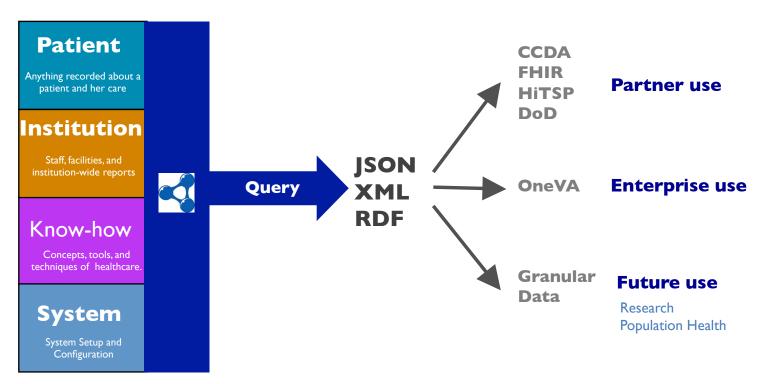
Silo VISTAs   Local VDMs   Analytics → Cross-linking   **Linked VISTA** (Enterprise VDM)   Enterprise Management

VISTA 1
VISTA 2
⋮
VISTA 130

P I K S

Patient
Anything recorded about a patient and her care

Institution
Staff, facilities, and institution-wide reports

Know-how
Concepts, tools, and techniques of healthcare.

System
System Setup and Configuration

Patient Data Management
Extract and manage patient data with patient-specific security and metadata, allowing patient-centric controls on data access and exchange.

Institutional Data Management:
Institution specific data can be exchangeable and centrally manageable

Knowledge Management:
Common medical concepts, standards, and know-how may be identified and managed as a clearly defined class of VISTA data.

Configuration Management:
The entire configuration of a system can be identified, extracted, and transported, and inserted from system to system

RM Richards MD MS 2015-02-23

Linked VISTA: Comprehensive Access

*One benefit of creating a cross-VISTA enterprise data model (Linked VISTA) is that it can be queried against **any** VISTA for **any** data with with **one** web-standard query.*
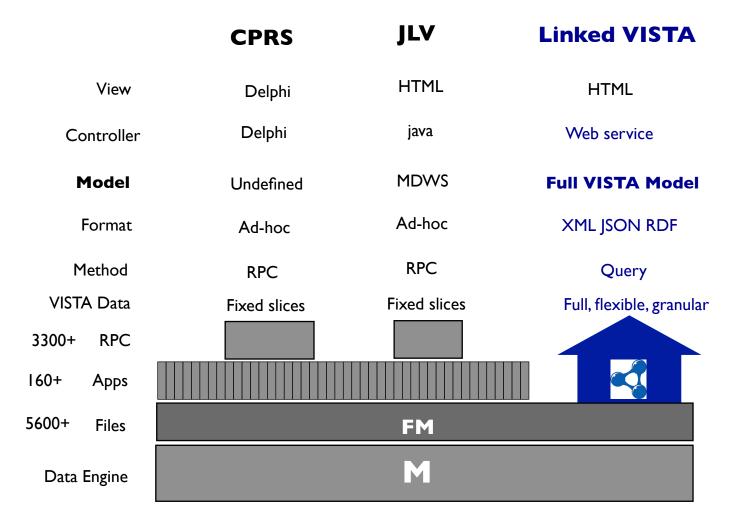
This would allow any authorized system to securely and directly query authoritative VISTA data in real-time with one standard query language. The output of these queries will be in all modern web-standard forms to maximize secondary use.

Linked VISTA

Patient

Anything recorded about a patient and her care

Institution

Staff, facilities, and institution-wide reports

Know-how

Concepts, tools, and techniques of healthcare.

System

System Setup and Configuration

Query → **JSON XML RDF**

CCDA
FHIR
HiTSP
DoD

Partner use

OneVA → **Enterprise use**

Granular Data

Future use

Research
Population Health

RM Richards MD MS 2015-02-23

Linked VISTA: Comprehensive Access

	CPRS	JLV	Linked VISTA
View	Delphi	HTML	HTML
Controller	Delphi	java	Web service
Model	Undefined	MDWS	**Full VISTA Model**
Format	Ad-hoc	Ad-hoc	XML JSON RDF
Method	RPC	RPC	Query
VISTA Data	Fixed slices	Fixed slices	Full, flexible, granular

3300+ RPC

160+ Apps

5600+ Files — **FM**

Data Engine — **M**

VISTA Data Model (VDM) can access all data spanning 160+ applications with full granularity because the fully exposed VISTA Data model bridges all applications. No legacy APIs, HL7, RPCs, or MUMPS. Just VISTA data. All of it. Real-time.

RM Richards MD MS 2015-02-23
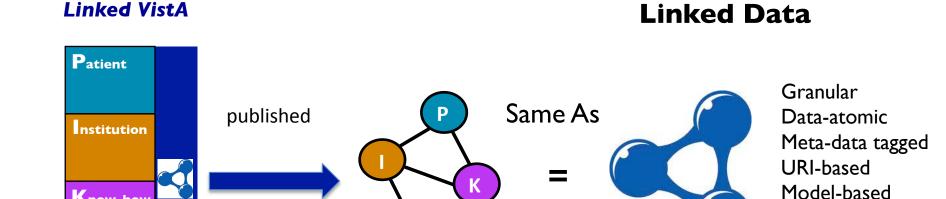
Linked VISTA: Web-standard Representation

Use of web-standard representation maximizes data re-use, meshing, and re-mixing with the maximum number of other sources of data for research and patient care.
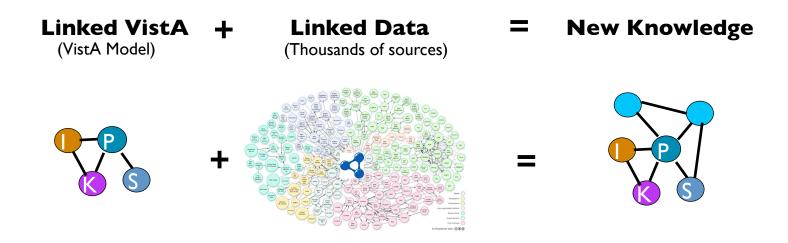
Linked VistA

Patient
Institution
Know-how
System

published →

P
I
K
S

Same As

=

Linked Data

Granular
Data-atomic
Meta-data tagged
URI-based
Model-based
Linkable
Web Standard

Linked VISTA: **New Knowledge Discovery**

Linked VISTA can participate in federated queries over unlimited number of other Linked Data sources, enabling meshing, enrichment, and ultimately, new knowledge discovery.
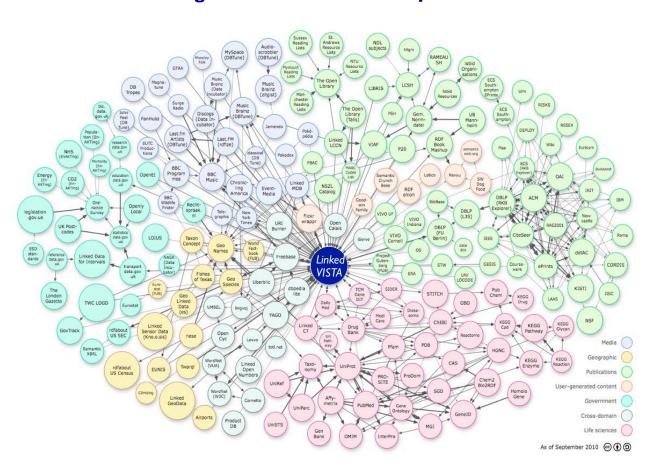
Linked VistA **+** **Linked Data** **=** **New Knowledge**
(VistA Model)   (Thousands of sources)

Because data in the VistA Data Model can be represented just like any other Linked Data resources, *one can mesh VistA data directly with unlimited sources of internal or external, public or private life sciences, and other scientific or healthcare related data sources.* This leads to discovery of new relationships between different sources of data - and new knowledge.

Representing VistA data in a Linked Data form supports real-time semantic integration with thousands of other other linked sources.

Linked VistA

VistA Data - represented in standard Linked Data form - can be interlinked with any and all other Linked Data sources.
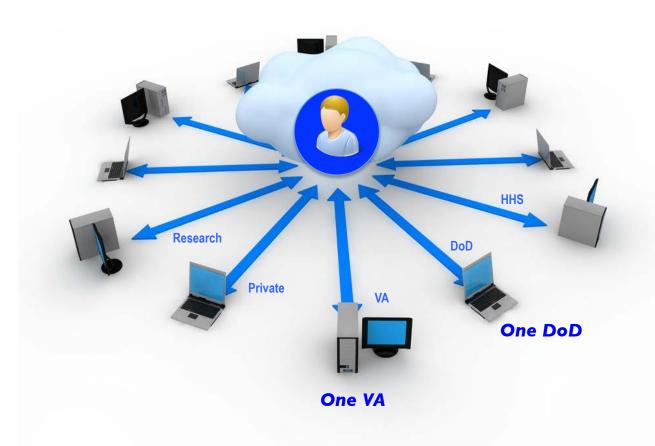
This will enable *meshing, enrichment, and augmentation of patient data with all other sources, providing a comprehensive view of all patient data from all locations, clinics, hospitals, or the home.*

Data sources also include *patient-generated, mobile device, TeleHealth, and any other Internet-enabled device data (Internet of Things).*

Vision: Linked VISTA

**Web-standards approach to provide
Patient-Centric Collaborative Care**

*One Patient.
One Model.
Integrated.
Web-scale.*

HHS

Research

DoD

Private

VA

One DoD

One VA

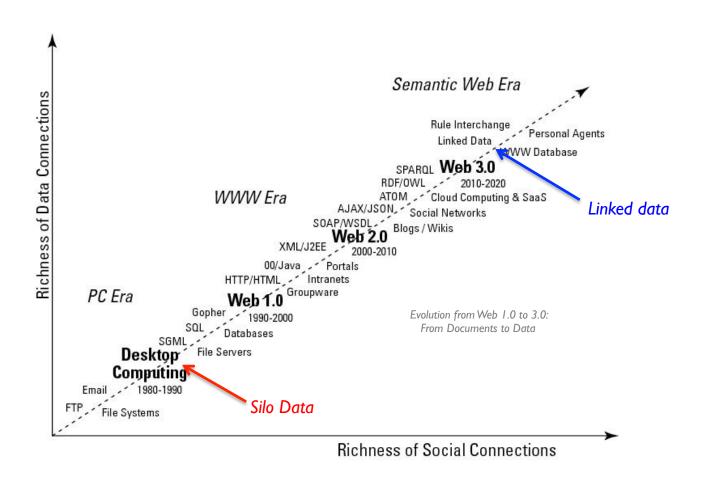# *Reference*

Linked Data Summary
Linked Data Links

Linked Data

- *Evolution of Computing: Increased connectivity*
- *What is it Linked Data?*
- *What problems does it solve?*
- *Who uses Linked Data?*
- *Health Data: Many diverse models*
- *Linked Data: Accommodates model diversity*
- *Health Data: PCAST Recommendation*

Semantic Web Era

Rule Interchange

Linked Data

Personal Agents

WWW Database

SPARQL **Web 3.0**

RDF/OWL 2010-2020

ATOM Cloud Computing & SaaS

AJAX/JSON Social Networks

SOAP/WSDL Blogs / Wikis

Web 2.0

XML/J2EE 2000-2010

OO/Java Portals

HTTP/HTML Intranets

Groupware

Web 1.0

Gopher 1990-2000

SQL Databases

SGML File Servers

Desktop Computing 1980-1990

Email

FTP File Systems

WWW Era

PC Era

Evolution from Web 1.0 to 3.0: From Documents to Data

Linked data (blue arrow)

Silo Data (red arrow)

Richness of Data Connections (y-axis)

Richness of Social Connections (x-axis)

DHCP was officially launched in 1982 and is the foundation of all VA VISTA and DoD CHCS systems.
This was in the early PC Era, before the age of inter-networked data ("the internet"), when databases were isolated inside institutional systems (Silo Data: red arrow). Note that the TCP/IP internet protocol had not even been invented at that time. To bring any VISTA system into today's WWW Era, allowing data to be linked to all VISTA systems within VA enterprise (or to any CHCS system in DoD), VISTA needs be web data-standards capable. This would transform VISTA to a WWW Database that could manage Linked Data (blue arrow).

http://en.wikipedia.org/wiki/History_of_the_Internet
http://paanchiweb.blogspot.com/2012/12/getting-essence-of-semantic-web.html

Web 1.0: Document Web (HTML)
Linked Documents
Read-only web (**humans** only)

Web 2.0: Social Web
Linked People
Read-write web (**humans** only)

Web 3.0: Semantic Web (RDF)
Linked Data
Read-write web (**machine processable**)

http://tomgruber.org/writing/CollectiveKnowledgeSystems.htm

Linked Data: *What is it?*

The World Wide Web (W3C) Standard for semantic information integration

HTML (hypertext markup language)
For **humans** to exchange information

enables →

Linked Documents
(Document Web)

RDF (resource description framework)
For **computers** to exchange information

enables →

Linked Data
(Semantic Web)

"The Semantic Web [Linked Data] provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries."
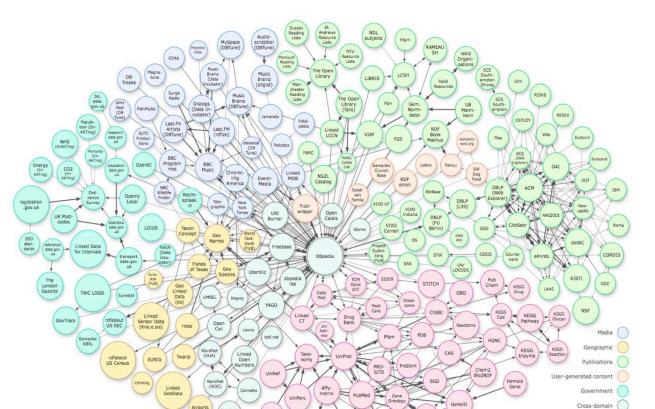
Tim Berners-Lee, MIT Professor and Inventor of the World Wide Web

As a W3C standard this supports Internet-scale data integration.

Web-scale semantic integration of data

Linked Data

This figure shows the Linked Open Data (LOD) cloud, which semantically links hundreds of Linked Data sources including Media, Geographic, Government, and Life Sciences databases.

Each circle represents one data source or database. These are semantically linked to other data sources, creating a single virtual federated internet-scale database.

At the center of is DBPedia, the Linked Data version of Wikipedia, which is semantically linked to hundreds of data sources.
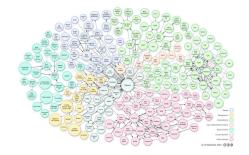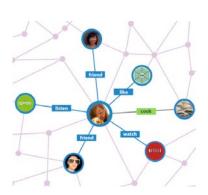
Linked Data: Who Uses It?

*The Linked Data approach to **Internet-scale semantic data integration** is used by the worlds largest organizations such as Google, LinkedIn, Facebook, and IBM Watson.*

Wikipedia (knowledge graph)

IBM Watson (knowledge graph)

LinkedIn (professional graph)

Facebook (social graph)

Google Search (knowledge graph)

Linked Data: Accommodates both Standards and Innovation

Information Models:
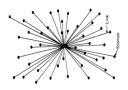An apparent conflict between standardization and innovation

Standards: need to remain static in order not to be disruptive for adopters.

- Static, Brittle
- Centralized
- General (Common Denominator)
- Committee-driven
- Large, "all-or-nothing", disruptive updates

vs.

Innovation: requires continuous evolution of thousands of new information models.

- Adaptive / Evolutionary
- Decentralized
- Highly Specialized, "Best of breed"
- End-user / specialist – driven
- Small, continuous, low-impact updates

What are the options?

Centralized, Model-rigid approach: For exchange of information to occur all models must remain fixed, and data must go through only one central 'broker' model..

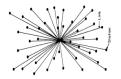Our current approach to healthcare data

Decentralized, Model-flexible approach: Multiple models peacefully co-exist and evolve, mediated by their ability to freely link to any model at all times. In this approach, all models are free to evolve AND are capable of resolving to a common standard model at all times. *The only current technology that supports this granular data-oriented approach is RDF (Linked Data).*

Linked Data supports **both** standardization and innovation

Data Integration: **Legacy vs. Linked Approach**

Architectural Issues	Current Approach	Linked Data (RDF)
Data model characteristics	**Model-rigid.** Only one lowest common denominator model unifies information. Must remain unchanged to orchestrate. **Restrictive expression.**	**Model-flexible.** All data models may independently evolve. **Maximizes expressivity.**
Data model compatibility	**No model diversity permitted** A one-size-fits-all mega-model	**Multiple models peacefully coexist** Data model flexible
Data model evolution	**Costly and difficult to evolve models.** Due to model-rigid architecture.	**Cheaper and easier to evolve models.** Due to model-flexible architecture.
Data access method	Downloading + Aggregating	**Linking + Federating**
Scalability: incremental effort required to add new data sources	**Common model** must be updated	**Individual models** may be independently and incrementally semantically linked.
Primary Function	**Data Syntax and Transport**	**Data Model**
Granularity	**Document-centric** (message-centric)	**Data-centric** (individual data elements)
Semantics	**Weak semantics.** **Extrinsic, separate from the data.** Depends on an external data model.	**Strong semantics.** **Intrinsic, integrated with the data.**

W3C Linked Data Standard
   http://www.w3.org/standards/semanticweb/data
W3C Linked Data Platform
   http://www.w3.org/TR/ldp/
W3C Semantic Web Healthcare and Life sciences
   http://www.w3.org/blog/hcls/
HL7-RDF Healthcare Standards Work Group
   http://wiki.hl7.org/index.php?title=RDF_for_Semantic_Interoperability
Semantic Web
   http://en.wikipedia.org/wiki/Semantic_Web
Linked Data: Tools (>1000)
   http://www.mkbergman.com/sweet-tools
Linked Data: Adaptors (100's)
   http://www.w3.org/wiki/ConverterToRdf
Linked Data: Roadmap for Healthcare Interoperability
   http://yosemiteproject.org

VHA: RDF is an approved technical reference model

http://www.va.gov/TRM/StandardPage.asp?tid=6405#