greenData, BIG Data

By Liora Alschuler, Lantana Consulting Group, Chief Executive Officer

Today's fledgling exchange networks are starved for data. Everything we know about networks tells us that we won't begin to realize the return on investment for the network or its nodes until we put massive amounts of data through the pipes. Putting massive amounts of clinical information onto health information exchange networks supports continuity of care and data reuse.

The Lantana **Trifolia Toolkit** uses the greenCDA methodology to open up the flow of Big Data within the enterprise and onto our health information networks. A comprehensive view of clinical information coupled with the tools to make compliance cost-effective can make good on the promise of the electronic health record.

In the *HIMSS12 Health Story Showcase*, the **Trifolia Toolkit** works with two sources: narrative from dictation and a semi-structured note. It enriches the source data through a variety of techniques to create coded, validated CDA documents ready for the *Showcase* transition of care scenario. The **Trifolia Toolkit** shown here is an operational prototype with production slated for Q1 2012.

> *The Lantana **Trifolia Toolkit** uses the greenCDA methodology to open up the flow of Big Data within the enterprise and onto our health information networks.*

Unlocking the Power of Big Data: 1,200,000,000 Electronic Documents

Health information technology (HIT) manages the information surrounding patient care delivery, reimbursement, monitoring, research, and reporting and is critical to all aspects of modern care. Most hospitals today are wired only for basic administrative functions—the *who*, *what*, *where*, and *how* much does it cost of healthcare. Clinical data, the patient record itself, is only now coming online in electronic lab systems, radiology information systems, practice management, and electronic health records. The government, state and federal, would like to accelerate this and has invested substantially through the HITECH and Affordable Care Acts in incentives to induce greater use of electronic medical records and greater (responsible) sharing of clinical information. [1]

Yet, with all the advances in HIT, the bulk of the record—approximately 1.2B notes each year in the US—remains locked out of the party. The most advanced all-electronic record systems move massive numbers of clinical documents, yet the largely narrative information in these documents is not accessible to nor integrated into the EHR; it is not available for transitions of care, decision support, public health, or quality reporting. Much reimbursement workflow is tied to information in these records, which currently must be identified and coded manually or semi-automatically. Today, this effort provides no value back to the clinical record. The advent of EHRs for structured data entry has not lessened the reliance on narrative in clinical records despite predictions of its imminent demise. The Health Story Project suggests that far from being an impediment to EHR adoption, narrative notes can, should, and will complement and support the EHR.

Background: Health Story, CDA, the Consolidation Project, Big Data, and HIMSS12

Health Story: An industry consortium dedicated to making the full clinical record available for care delivery and reuse in reporting and research. [2] Health Story uses CDA.

Simple CDA for Exchange

Patient	Adam Everyman		
Date of birth	November 25, 1954	Sex	Male
Contact info	17 Daws Rd. Blue Bell, MA 02368, USA Tel: (781)555-1212	Patient IDs	12345 2.16.840.1.113883.19.5
Document Id	999021 2.16.840.1.113883.19		
Document Created:	March 3, 2005, 17:15:04 +0500		
Performer (primary care physician)	Dr. Henry Seven		
Author	Henry Seven		
Contact info	21 North Ave Burlington, MA 01803, USA Tel: (555)555-1002		
Encounter Id	9937012 2.16.840.1.113883.19	Encounter Type	Evalua
Encounter Date	From March 29, 2005 to March 29, 2005		
Legal authenticator	Henry Seven signed at March 29, 2005, 22:44:11 +0500		
Contact info	21 North Ave Burlington, MA 01803, USA Tel: (555)555-1002		
Document maintained by	Good Health Clinic		
Contact info	21 North Ave Burlington, MA 01803, USA Tel: (555)555-1212		

```
<recordTarget>
  <patientRole>
  ...
  <patient>
    <name>
      <given>Adam</given>
      <family>Everyman</family>
    </name>
  </patient>
  </patientRole>
</recordTarget>
```

The Clinical Document Architecture (CDA): An exchange specification for clinical documents from Health Level Seven (HL7). CDA for human readers is simple; CDA for global semantic interoperability across all healthcare use cases is complex. [3]

The Consolidation Project: A joint effort of Health Story, HL7, Integrating the Healthcare Enterprise (IHE), and the Office of the National Coordinator (ONC) Standards and Interoperability Framework. Consolidation took requirements for the Continuity of Care Document (CCD) and the eight Health Story Implementation Guides published through HL7 (History & Physical, Consult Note, Discharge Summary, etc.) and made them consistent with related IHE work and the ONC requirements for Meaningful Use. [4]

Big Data: Wikipedia says that the term applies to "data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time." We think this applies in the current HIT environment because the vast majority of mission-critical data still lie outside the capabilities of databases, warehouses and analytic engines. This paper, the greenCDA method, and the **Trifolia Toolkit** profiled here are all about how to get Big Data moving over healthcare networks and into our EHRs and applications.

HIMSS12: The *HIMSS12 Health Story Showcase* puts all these into play in a Transition of Care scenario designed by the American College of Physicians and deployed by Lantana, our partners and colleagues in Health Story, and a group of dedicated volunteers. [5]

greenCDA is about Big Data

greenCDA lives to get Big Data flowing over health information networks. It walks on both sides of the complexity divide over how to do health information interoperability. On the side of simplicity, it provides simple, use-case specific XML schemas that use business names (not abstractions) and which can be understood and implemented by developers without lengthy training. At the same time, these simple-to-create instances can be transformed automatically into canonical CDA for conformance with the most general, global specification meeting the highest criteria for semantic interoperability. Getting to that level of interoperability directly is tough; getting there with the greenCDA method is easy. [6]

Equivalent greenCDA vs. Normative XML

Normative CDA
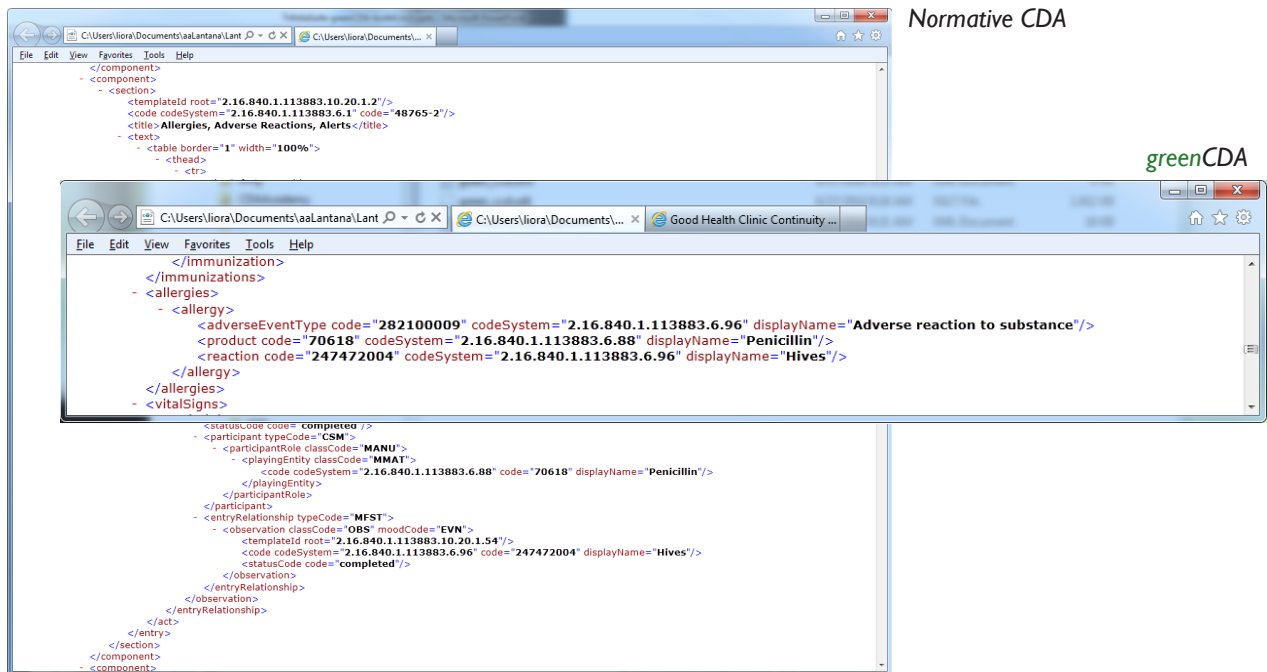
greenCDA

Where canonical CDA might require nested tags for `<component>`, `<section>`, `<code>`, `<value>`, greenCDA can express the equivalent with one developer-friendly tag such as `<allergies>`. The former give you global interoperability on a single framework; the latter gives you ready-to-implement interfaces.

- At design time, a highly-trained specification developer uses requirements to develop the CDA templates along with the green schemas and transforms.
- In implementation, any XML developer can use greenCDA applications to create documents that are automatically transformed to the canonical CDA. Senders get expedited implementation; recipients get the flexibility and full semantic interoperability of CDA, making it possible to integrate and reuse information across a wide spectrum of use cases.

While still early days for greenCDA, indications are that it can introduce time savings of up to 90% for the sender, which pave the way for large scale information sharing. [7] Leveraging this method can be cost-effective for the lion's share of the clinical record, making it available for cost-reducing and quality-improving measures. The **Trifolia Toolkit** deployed in the *Showcase* does just that—it uses the greenCDA method to open up Big Data for EHR consumption and reuse.

Simple-to-create instances can be transformed automatically into canonical CDA for semantic interoperability. Getting to that level of interoperability directly is tough; getting there with the greenCDA method is easy.

The Lantana Trifolia Toolkit for greenCDA

The **Trifolia Toolkit** [8] makes the full health story available for use in clinical care, reimbursement, and reporting. It imports structured or narrative clinical records and exports industry-standard, coded documents that can be consumed by an EHR or reporting tool to support care coordination, quality and public health reporting, reimbursement, and decision support.

The **Trifolia Toolkit** supports:

- Rapid development of standards-compliant electronic documents on a tight timeline for cost-effective CDA implementation
- Any CDA template compliant format, such as simple narrative and/or coded documents
- C32-compliance for Meaningful Use Stage One
- Consolidation Project documents for quality reporting, transitions of care

The **Trifolia Toolkit** assembles, transforms, and validates clinical source data and documents in a single service, enriching the information with coded terms and producing valid CDA that meets requirements for interoperability and reuse. It can work from multiple sources—narrative, structured or semi-structured documents—and can enrich them using natural language processing (NLP), computer assisted coding (CAC) and terminology services.

The *HIMSS12 Health Story Showcase* demonstrates these features in two ways:

- Level 3 Referral Note: Transformation of a narrative referral note to coded CDA with Consolidation Project templates using natural language processing and terminology services. Partners: BayScribe and IMO
- Level 3 History & Physical Note: Generation of a History & Physical note from semi-structured (CDA Level 2) enriched through computer-assisted coding. Partners: ChartLogic (Stella EHR) and OptumInsight (CAC)

Lantana is working today with partners in business and government to deploy the **Trifolia Toolkit** in two pilot projects for reporting of quality data. The **Trifolia Toolkit** will make it easy to reuse existing EHR data, reducing error and redundant data entry. The pilots are scheduled for Q2-Q3 2012. Please contact Lantana for more information.

RESOURCES

[1] HITECH: ONC/CMS Final Rules: https://www.cms.gov/EHRIncentivePrograms and http://healthit.hhs.gov/media/MU/n508/MU_SCC_CombinedGrid.pdf

[2] Health Story: www.healthstory.com

[3] CDA: http://www.hl7.org/implement/standards/cda.cfm
CDA Academy: www.cdaacademy.com

[4] Consolidation Project: http://wiki.siframework.org/CDA+Consolidation+Project+Charter and http://www.hl7.org/documentcenter/public/standards/dstu/CDAR2_IG_IHE_CONSOL_R1_DSTU_2011DEC.zip

[5] See the Health Story website for more information on the collaboration with ACP and development of the *HIMSS12 Health Story Showcase*

[6] greenCDA: http://www.hl7.org/implement/standards/cda.cfm

[7] Use of greenCDA: http://www.hl7.org/documentcenter/public_temp_F404B3B0-1C23-BA17-0CD6918A4C0CB3C9/plenary/2011/3_Daniel%20Pollock%20-%20Panel%20Presentation%20-%20HL7%20Plenary%20and%20WGM%2009.12.2011.pdf

[8] Lantana **Trifolia Toolkit**: http://www.lantanagroup.com/what-we-do/application-development/

Lantana Consulting Group

714.232.8679

www.lantanagroup.com

ABOUT LANTANA

Lantana Consulting Group provides services and software for standards-based health information exchange. Lantana's client list includes Fortune 100 companies and organizations in the public and private sectors, including the Centers for Disease Control and Prevention (CDC), the Centers for Medicare & Medicaid Services, the National Cancer Institute and the National Quality Forum.