Department of Veterans Affairs
Veteran Health Administration
Knowledge Based Systems

Informatics Architecture Support Services

"Data Element" Modeling and Relationship to SOLOR

Walter Sujansky
Sujansky & Associates, LLC

Veterans Health Administration, Office of Informatics & Analytics and Health Informatics Informatics Architecture Support Services Contract VA701-16-C-0039 CLIN 0012, 5.4.1A. Bi Weekly Web Based Workshops

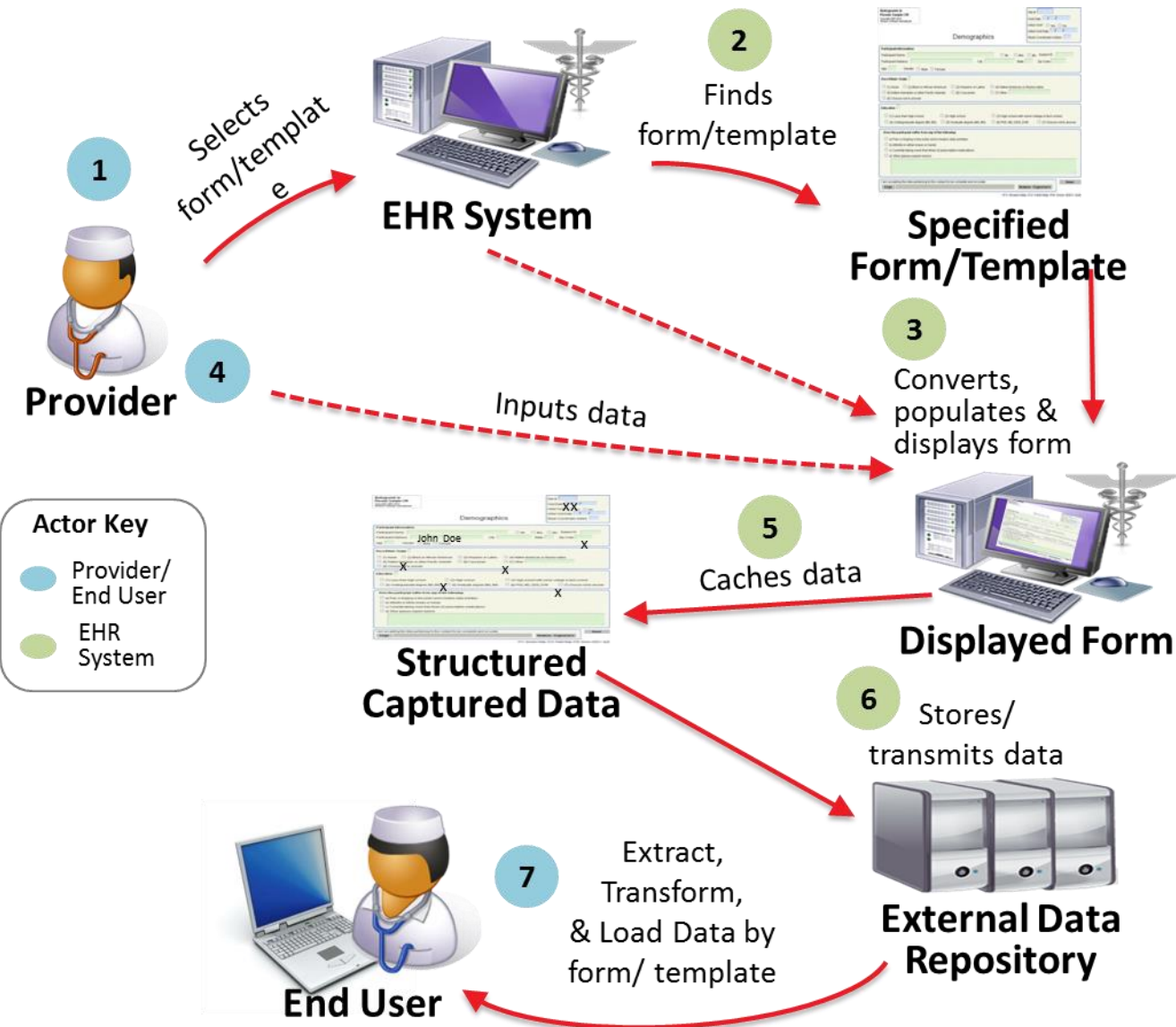www.BookZurman.com

Outline

- What are "Data Elements"
  - Origins
  - Model and Uses
- Examples of Data Elements
- Issues with Data Elements
- Relationship to SOLOR

Origins

- ONC S&I Structured Data Capture (SDC) Initiative
  - Initiated 2013
  - Goal:
    - "Enabling the collection of structured data within EHRs to supplement data collected for other purposes, including clinical research, adverse event reporting, and public health reporting."

S&I SDC Initiative

EHR System

1 — Selects form/template

2 — Finds form/template

Specified Form/Template

3 — Converts, populates & displays form

Provider

4 — Inputs data

Actor Key
- Provider/End User
- EHR System

Structured Captured Data

John Doe

5 — Caches data

Displayed Form

6 — Stores/transmits data

7 — Extract, Transform, & Load Data by form/template

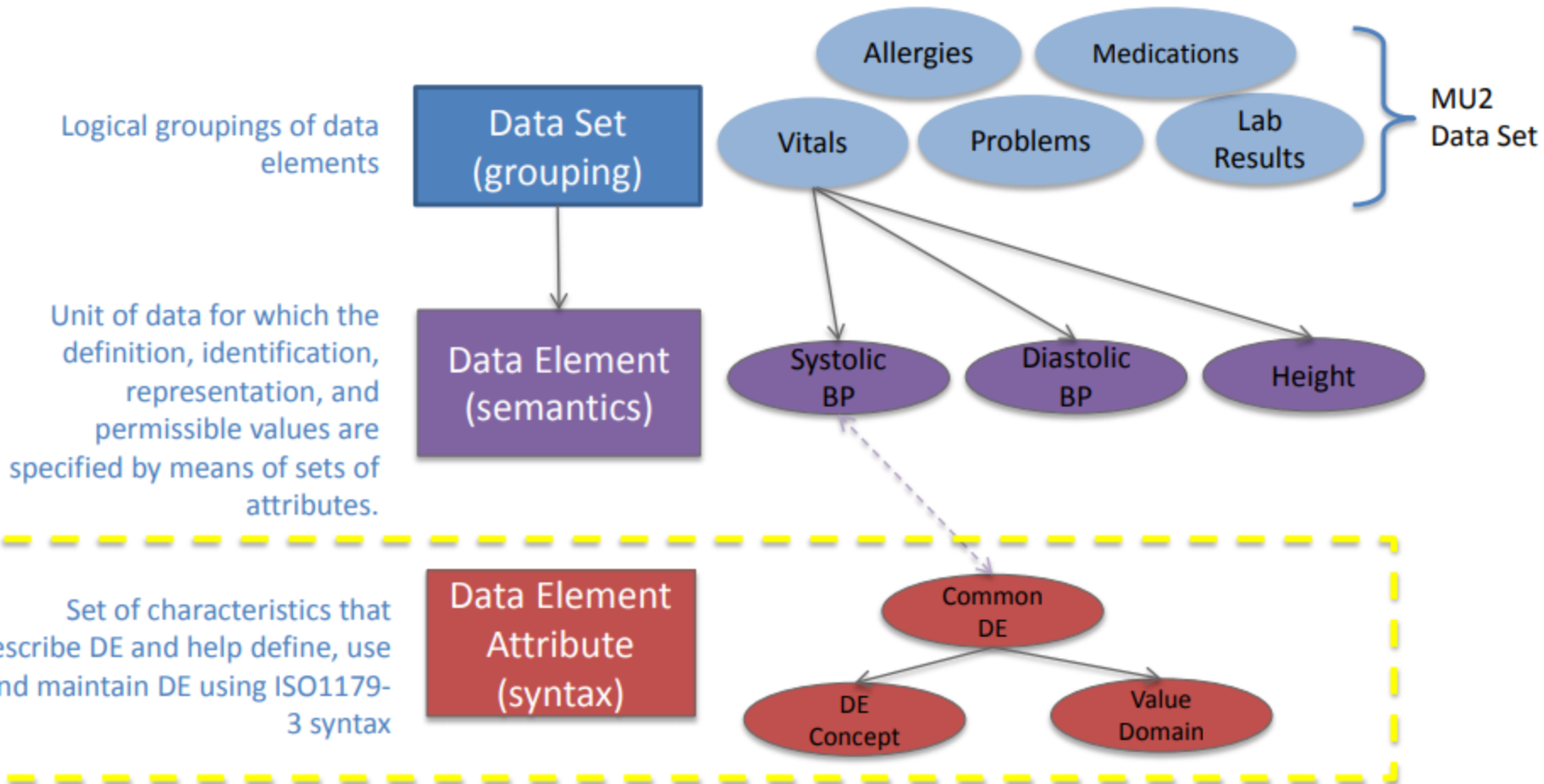**External Data Repository**

End User

4

Origins

- ONC S&I Structured Data Capture (SDC) Initiative
  - Initiated 2013
  - Goal:
    - "Enabling the collection of structured data within EHRs to supplement data collected for other purposes, including clinical research, adverse event reporting, and public health reporting."
  - Standards specifications for:
    1. Structure/definition of the *data elements* that will be used to fill the specified forms or templates
    2. Structure or design of the form
    3. How EHRs interact with the form or template
    4. Guidance to pre-populate and auto-populate form or template with existing patient data

SDC Data Elements

Data Element (DE)	A data element is a unit of data for which the definition, identification, representation, and permissible values are specified by a set of attributes and considered in context to be indivisible.

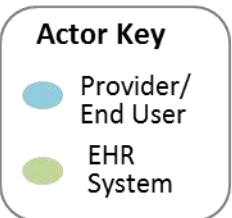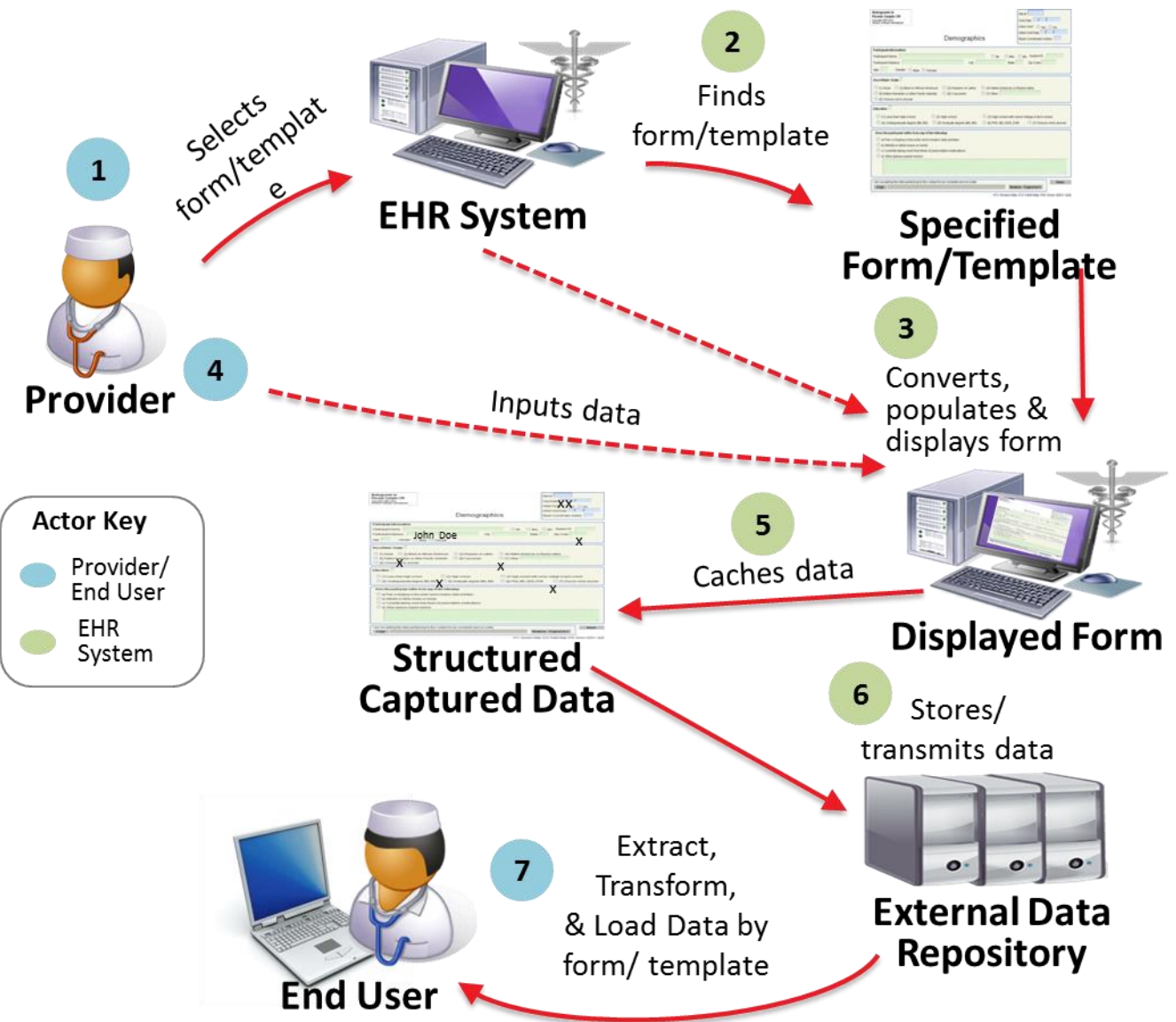|	
Common Data Elements (CDEs)	Standardized data element descriptions for collection and exchange of data of common interest to a particular community, and thus the community has agreed to share their definition, management and use. CDEs share a common set of attributes which facilitates their reuse in different settings, and are intended to aide in interoperability and data reuse.

SDC Data Elements

Common Data Elements (CDEs) are those DEs that are developed, maintained and used based on commonly agreed upon principles by the user community. CDEs are reusable across a variety of clinical and non-clinical domains.

Common Data Element - Example

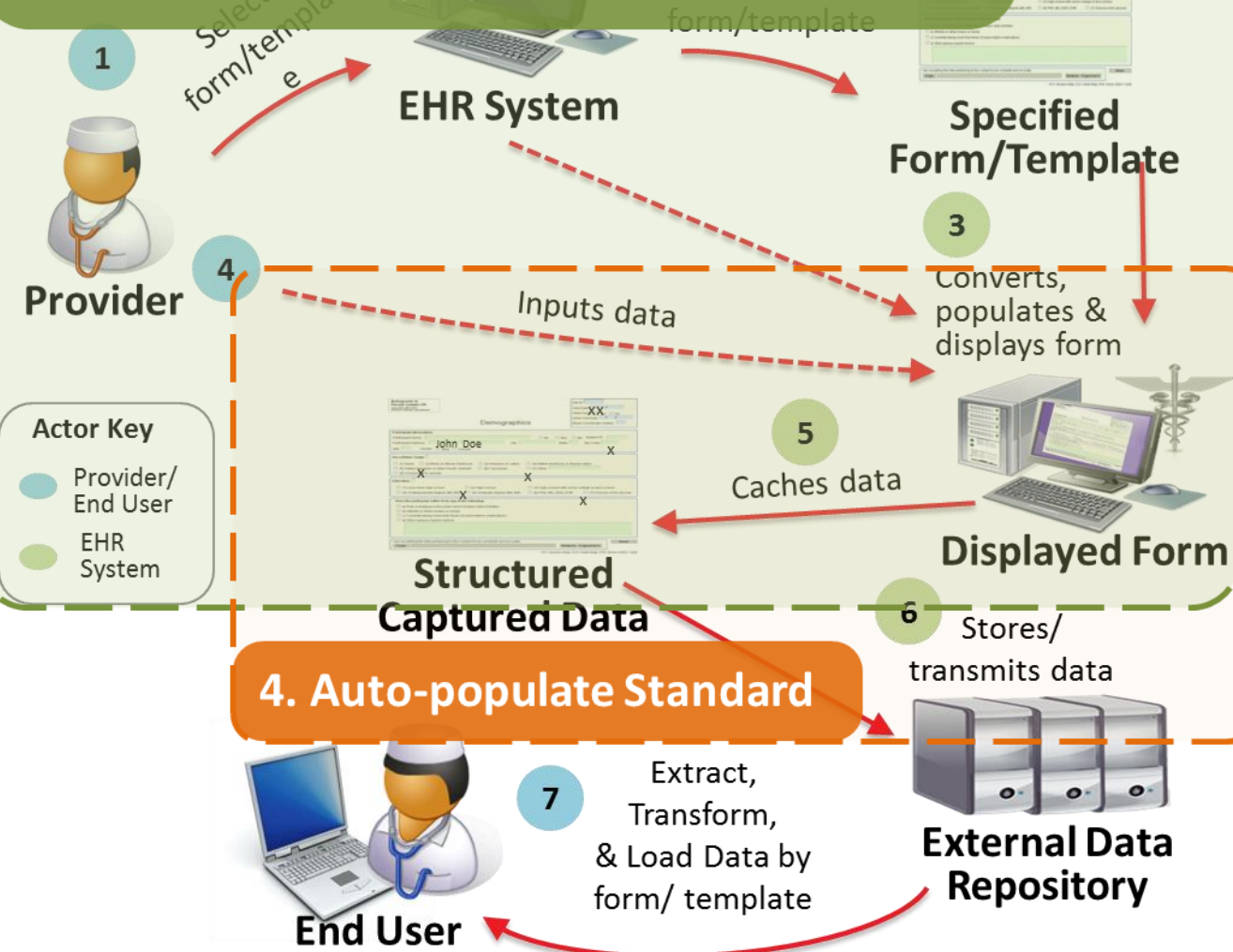| SDC PSE/AE SWG Data Element	name (M)	URI (O) **	permitted value (M) *	text (M) *	value meaning description (O) *	value meaning concept code (O) *	code system id (C) *	code system name (C) *	code system version (C) *
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Life-threatening	Life_Threatening_ Adverse_Event	An adverse event, and/or its immediate sequelae, which is associated with an imminent risk of death or which is associated with physical or mental disabilities that affect or limit the ability of a person to perform activities of daily living (eating, ambulation, toileting, etc.) See NCI Common Terminology Criteria for Adverse Events.	C41337	2.16.840.1. 113883.3.2 6.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Hospitalization - initial or prolonged	Hospitalization	The condition of being treated as a patient in a hospital. An indication or description that an adverse event is associated with or prolongs hospitalization.	C25179 C83052	2.16.840.1. 113883.3.2 6.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Disability or Permanent Damage	Disablity	Any physical, mental, or functional impairment that limits a major activity. It may be partial or complete. (Taber's)	C21007	2.16.840.1. 113883.3.2 6.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Congenital Anomaly/Birth Defect	Congenital_Abnor mality	Any abnormality, anatomical or biochemical, evident at birth or during the neonatal period.	C2849	2.16.840.1. 113883.3.2 6.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Death	Death	The absence of life or state of being dead.	C28554	2.16.840.1. 113883.3.2 6.1.1	NCI Thesaurus	14.02d

S&I SDC Initiative

- ONC S&I Structured Data Capture (SDC) Initiative
  - Initiated 2013
  - Goal:
    - "Enabling the collection of structured data within EHRs to supplement data collected for other purposes, including clinical research, adverse event reporting, and public health reporting."
  - Standards specifications for:
    1. Structure/definition of the *data elements* that will be used to fill the specified forms or templates
    2. Structure or design of the form
    3. How EHRs interact with the form or template
    4. Guidance to pre-populate and auto-populate form or template with existing patient data

S&I SDC Initiative

3. EHR Interaction
Standard: Find, display, cache, store/transmit

1. CDE Structure Standard

- Clinical Research CDEs
- AHRQ CDEs [Common Formats]
- Other Domain CDEs

Form Library

2. Form/Template Structure Standard

- Other domain-specified Forms

Template Library

- Domain-specified Templates

(1) Selects form/template

EHR System

(2) form/template

Specified Form/Template

(3) Converts, populates & displays form

Provider

(4) Inputs data

Actor Key
- Provider/ End User
- EHR System

Structured Captured Data

(5) Caches data

Displayed Form

4. Auto-populate Standard

(6) Stores/ transmits data

(7) Extract, Transform, & Load Data by form/ template

External Data Repository

End User

11

Common Data Element - Example

SDC PSE/AE SWG Data Element	name (M)	URI (O) **	permitted value (M) *	text (M) *	value meaning description (O) *	value meaning concept code (O) *	code system id (C) *	code system name (C) *	code system version (C) *
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Life-threatening	Life_Threatening_Adverse_Event	An adverse event, and/or its immediate sequelae, which is associated with an imminent risk of death or which is associated with physical or mental disabilities that affect or limit the ability of a person to perform activities of daily living (eating, ambulation, toileting, etc.) See NCI Common Terminology Criteria for Adverse Events.	C41337	2.16.840.1.113883.3.26.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Hospitalization - initial or prolonged	Hospitalization	The condition of being treated as a patient in a hospital. An indication or description that an adverse event is associated with or prolongs hospitalization.	C25179 C83052	2.16.840.1.113883.3.26.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Disability or Permanent Damage	Disablity	Any physical, mental, or functional impairment that limits a major activity. It may be partial or complete. (Taber's)	C21007	2.16.840.1.113883.3.26.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Congenital Anomaly/Birth Defect	Congenital_Abnormality	Any abnormality, anatomical or biochemical, evident at birth or during the neonatal period.	C2849	2.16.840.1.113883.3.26.1.1	NCI Thesaurus	14.02d
Outcomes attributed to adverse event	Adverse Event Outcome ICSR Terminology	http://ncit: C54583	Death	Death	The absence of life or state of being dead.	C28554	2.16.840.1.113883.3.26.1.1	NCI Thesaurus	14.02d

CDE Structure Standard

Attribute Name	Cardinality	Optionality	Definition	Constraints	Example	IHE DEX Attribute Name
scoped identifier	1..1	Mandatory	A globally unique identifier for a metadata item within a specific namespace. It is usually composed of the DE Identifier, DE Version, and either the full_expansion OR shorthand_expansion associated with the namespace. If a Data Element repository does not use DE Verstion, then the scoped_identifier may not include a version number. It is up to the owner of the data element to determine what the scoped_identifier value should be.		http://nci.nih.gov/xml/owl/cadsr/data_element_identifier#2006410v3.0 46552-6_LOINC	id
version	1..1	Mandatory	The unique version identifier of the metadata item.		3	version
name	1..*	Mandatory	The designation of an object by a linguistic expression; The primary means of identification of objects and concepts. A name by which a metadata item is known within a specific context.	SHOULD be "survey_quest_text" selected from LOINCDB, bound DYNAMIC, binding strength CWE LOINC OID: 2.16.840.1.113883.6.1	Person Body Mass Index Value Treated for urinary tract infection in past 14 days [OASIS]	
preferred question text	1..1	Mandatory	The preferred text used to express the data element as a question.	SHOULD be "survey_quest_text" selected from LOINCDB, bound DYNAMIC, binding strength CWE LOINC OID: 2.16.840.1.113883.6.1	Body Mass Index (BMI) Has this patient been treated for a Urinary Tract Infection in the past 14 days?	displayName
alternative question text	1..*	Optional	Alternative expressions of the question text that have been approved for use this is data element.	SHOULD be "survey_quest_text" selected from LOINCDB, bound DYNAMIC, binding strength CWE LOINC OID: 2.16.840.1.113883.6.1	Body Mass Index	
definition	1..*	Mandatory	Representation of a concept by a descriptive statement which serves to differentiate it from related concepts. A statement that expresses		A calculated numerical quantity that represents an individual's	definition

From:
https://oncprojectracking.healthit.gov/wiki/download/attachments/16123327/SDC%20SWG%20Data%20Element%20Mapping%20Templatev4%20%282%29.xlsx?version=1&modificationDate=1489605858000&api=v2
(SDC Data Element Attributes tab)

CDE Structure Standard

Attribute Name	Cardinality	Optionality	Definition	Constraints	Example
format	0..1	Optional	template for the structure of the presentation of the value(s)		YYYY-MM-DD for a date
maximum character	0..1	Optional	Maximum number of characters available to represent the Data_Element value		4
unit of measure	1..1	Conditional	The name of a unit of measure.	If datatype is Physical Quantity type then "unit of measure" attribute SHALL be supplied. SHALL be UCUM code from value set below bound	mmHg
low value	0..1	Optional	Low end of data range	If value set is non-enumerated, then low value MAY be provided. Enumerated value domains SHALL NOT have low value	80
high value	0..1	Optional	High end of data range	If value set is non-enumerated, then high value MAY be provided. Enumerated value domains SHALL NOT have high value	120
value set	1..1	Conditional	If a value domain draws its values from a value set, then value set must be associated with the value domain, and value set members must be made available as permissible values.		
value meaning concept code	0..1	Optional	The preffered standard terminology code for the value set member.	SHOULD be CONCEPT ID selected from SNOMED CT, bound DYNAMIC, binding strength CWE SNOMED CT OID: 2.16.840.1.113883.6.96	73211009, 502372015
value meaning description	0..1	Optional	The exact text for the meaning of the permitted value.	SHOULD be 'term' selected from SNOMED CT, bound DYNAMIC, binding strength CWE SNOMED CT OID: 2.16.840.1.113883.6.96	The substance inside red blood cells that binds to oxygen in the lungs and carries it the tissues.
code system id	1..1	Conditional	unique identifier for the code system	SHALL be OID If value member codes originate from specific coding	2.16.840.1.113883.6.1
code system name	1..1	Conditional	name of the code system	If value member codes originate from specific coding system then coding system name SHALL be provided	LOINC
code system version	1..1	Conditional	version of the code system	If value member codes originate from specific coding system then coding system version SHALL be provided	2.43
content Model	0..1	Optional	The Content Model that the Data Element is interrelated with.		
mapping type	0..1	Optional	Type of the mappingSpecification. The type shall be selected from Mapping Specification Type Value Set (1.3.6.1.4.1.19376.1.7.3.1.1.22.1).	Value SHALL be selected from IHE DEX Profile, Mapping Specification Type Value Set. OID: 1.3.6.1.4.1.19376.1.7.3.1.1.22.1	XPATH Scripts, SPARQL or SQL queries
mapping script	0..1	Optional	The exact specification to locate the Data Element in a Content Model (Example: XPATH Scripts, SPARQL or SQL query).		select * from table_name
id	0..1	Optional	Identifier for the contentModel. This is the OID of the contentModel.		2.16.840.1.113883.10.20.1 (for ASTM/HL7 CCD)
name	0..1	Optional	Name of the contentModel.		ASTM/HL7 CCD

14

Results of S&I SDC Initiative

1. IHE SDC Profile
   - Utilizes the IHE Retrieve Form for Data Capture (RFD) Profile
     - For retrieving, displaying, completing, and submitting forms
   - Utilizes the ISO/IEC 19763-13 Meta-model for Framework Interoperability (MFI)
     - For specifying the content of forms
   - Defines a set of XML schemas for defining SDC forms
     - Top level "<FormDesign>" element
     - Others: "<Section>", "<Question>", "<ListItem>", …
     - See https://github.com/esacinc/sdc-schema-package/tree/master/schema/sdc
   - Published 2014

1. IHE SDC Profile

Integrating the Healthcare Enterprise

IHE

5 **IHE Quality, Research, and Public Health Technical Framework Supplement**

10 **Structured Data Capture (SDC)**

15 **Rev. 2.1 – Trial Implementation**

See
http://ihe.net/uploadedFiles/Documents/QRPH/IHE_QRPH_Suppl_SDC.pdf

See
https://github.com/esacinc/sdc-schema-package/blob/master/example/form/CAP_Breast_Res.xml

```xml
- <Question title="Primary Tumor Site: Invasive Carcinoma (Note C)" ID="16250.100004300" minCard="0"
  mustImplement="false">
    <OtherText val="Primary Tumor Site: Invasive Carcinoma" type="reportText" name="Tum_Site_113"/>
  - <ListField>
    - <List>
        <ListItem title="Upper outer quadrant" ID="16251.100004300" mustImplement="false"/>
        <ListItem title="Lower outer quadrant" ID="16252.100004300" mustImplement="false"/>
        <ListItem title="Upper inner quadrant" ID="16253.100004300" mustImplement="false"/>
        <ListItem title="Lower inner quadrant" ID="16254.100004300" mustImplement="false"/>
        <ListItem title="Central" ID="16255.100004300" mustImplement="false"/>
        <ListItem title="Nipple" ID="16256.100004300" mustImplement="false"/>
      + <ListItem title="Other (specify)" ID="16257.100004300" mustImplement="false">
        <ListItem title="Not specified" ID="16258.100004300" mustImplement="false"/>
    </List>
  </ListField>
```

1. IHE SDC Profile

2. S&I Pilot Implementations
   - Workstreams
     - Public Health
     - Patient Safety Reporting and Adverse Event Reporting
     - PCOR (research)
   - Goals
     - Real world evaluation of SDC IHE Profile & FHIR Profile
     - Harmonize 'content' data elements and forms

Results of S&I SDC Initiative

1. IHE SDC Profile

2. S&I Pilot Implementations

SDC Initiative Deliverables

Expected Deliverable/Artifact:	Status:
Use Case Document	Complete
SDC Candidate Standards List SDC Solution Plan SDC Standards Evaluation	Complete
SDC SOAP/SAML Implementation Guide SDC Data Element Attribute and Structure Mapping Template	Complete
SDC PSE/AE Subworking Group Workflows PSE/AE SubWorking Group Final Presentation SDC Engagement Matrix and Environmental Scan	Complete
FHIR SDC DE Profile FHIR Profile	Complete
IHE SDC Profile	Complete
NLM CDE Repository	Complete
GitHub Link UCSF Pilot Report Out Slide Deck (11/17/2017) CAP & DCG Pilot Report Out Slide Deck (12/08/2017)	Complete

See:
https://oncprojectracking.healthit.gov/wiki/display/TechLabSC/SDC+Pilots

Other "Data Element" Initiatives

See: https://cde.nlm.nih.gov/form/search

Example OWL/SOLOR Concept Expression

60 results for `All Terms` | `NCI ✕` | `All Topics` | `All Statuses` (0.025 secs)

Filter by:

Classification ➕
∨ NCI (60)
  › NCIP (60)

Registration Status
☐ Qualified (60)

Staging, Neuroendocrine Tumor NCI Standard Template Qualified 📌 ⊞
Matched by: **Classification**

21 Questions - The collection of CDEs used in the neuroendocrine tumor staging module based on AJCC Edition 7.

Steward:	NCI
Used By:	NCI

Staging, Cervix Uteri NCI Standard Template Qualified 📌 ⊞
Matched by: **Classification**

10 Questions - The collection of CDEs used in the cervical staging module based on AJCC Edition 7 and FIGO standards.

Steward:	NCI
Used By:	NCI

Staging, Prostate NCI Standard Template Qualified 📌 ⊞
Matched by: **Classification**

10 Questions - The collection of CDEs used in the prostate staging module based on AJCC Edition 7.

Steward:	NCI
Used By:	NCI

Adverse Event/Serious Adverse Event CTCAE v3 NCI Standard Template Qualified 📌 ⊞
Matched by: **Classification**

28 Questions - The collection of CDEs used in the Adverse Events/Serious Adverse Events module.

Steward:	NCI
Used By:	NCI

Other "Data Element" Initiatives

General Details | Form Description | Naming | Classification | Reference Documents

Adverse Event/Serious Adverse Event CTCAE v3 NCI Standard Template

Mandatory Adverse Event Questions

Must be included when this data is collected for reporting.

3. Describe 'Other' Adverse Event

4. Adverse Event Grade

☐ MILD ADVERSE EVENT ☐ MODERATE ADVERSE EVENT ☐ SEVERE ADVERSE EVENT ☐ LIFE-THREATENING OR DISABLING ADVERSE EVENT

☐ DEATH RELATED TO ADVERSE EVENT

5. Adverse Event Name (v3.0)

☐ Acidosis NOS	☐ Capillary leak syndrome	☐ ALLERGY/IMMUNOLOGY OTHER	☐ Dry mouth
☐ Ulcer, GI	☐ Ulcer, GI	☐ Ulcer, GI	☐ Ulcer, GI
☐ Ulcer, GI	☐ Ulcer, GI	☐ Gastric ulcer	☐ Ulceration
☐ Hyperuricaemia	☐ Fanconi syndrome	☐ Dry skin	☐ Urinary frequency
☐ Urinary retention	☐ Urine discolouration	☐ Urticaria NOS	☐ Uveitis
☐ Vaginal discharge (non-infectious)	☐ Vulvovaginal dryness	☐ Vaginal mucositis	☐ Vaginal stenosis/length

6. AE Start Date

☐ Unknown

7. Was patient hospitalized for toxicity?

☐ Yes ☐ No

8. Cycle #

Other "Data Element" Initiatives

Other "Data Element" Initiatives

Adverse event severity grade scale

[+ Add to Quick Board] [⚲ Add to Board] [⤓ Export] [⧉ Copy] [💬 Disc

General Details | **Permissible Values** | Naming | Classification | Concepts | Reference Documents | Properties | Identifiers | Attachments | History | S

Values Allowed

Value Type: Value List
Unit of Measure: N/A
Ids: empty

Permissible Value

Datatype: Text

Value	Code Name	Code	Code System	Code Description
Grade 1	Grade 1			Mild: asymptomatic or mild symtoms, clinical or diagnostic observations only, intervention not indicated.
Grade 2	Grade 2			Moderate: minimal, local or noninvasive intervention indicated, limiting age-appropriate Activities of Daily Living.
Grade 3	Grade 3			Severe or medically significant but not immediately life-threatening: hospitalization or prolongation of hospitalization indicated, disabling,
Grade 4	Grade 4			Life-threatening consequences: urgent intervention indicated
Grade 5	Grade 5			Death related to adverse event

Other "Data Element" Initiatives

- Effectiveness for standardized, structured *data entry*
  - Sharing of forms (good, if occurs)
  - Shared data elements among forms (good, if occurs)
  - All coded concepts must be pre-coordinated (bad)
  - Semantics of data elements may depend on form context
    - E.g., "Cycle #"   (Chemotherapy? IVF?)
    - E.g., "Was patient hospitalized for toxicity" (Drug AE? Env. Exposure?)
  - No semantic model for the *relationships* among the data elements in a form (bad)
    - E.g., "AE Name", "AE Grade", "AE Start Date" all refer to the same AE instance.
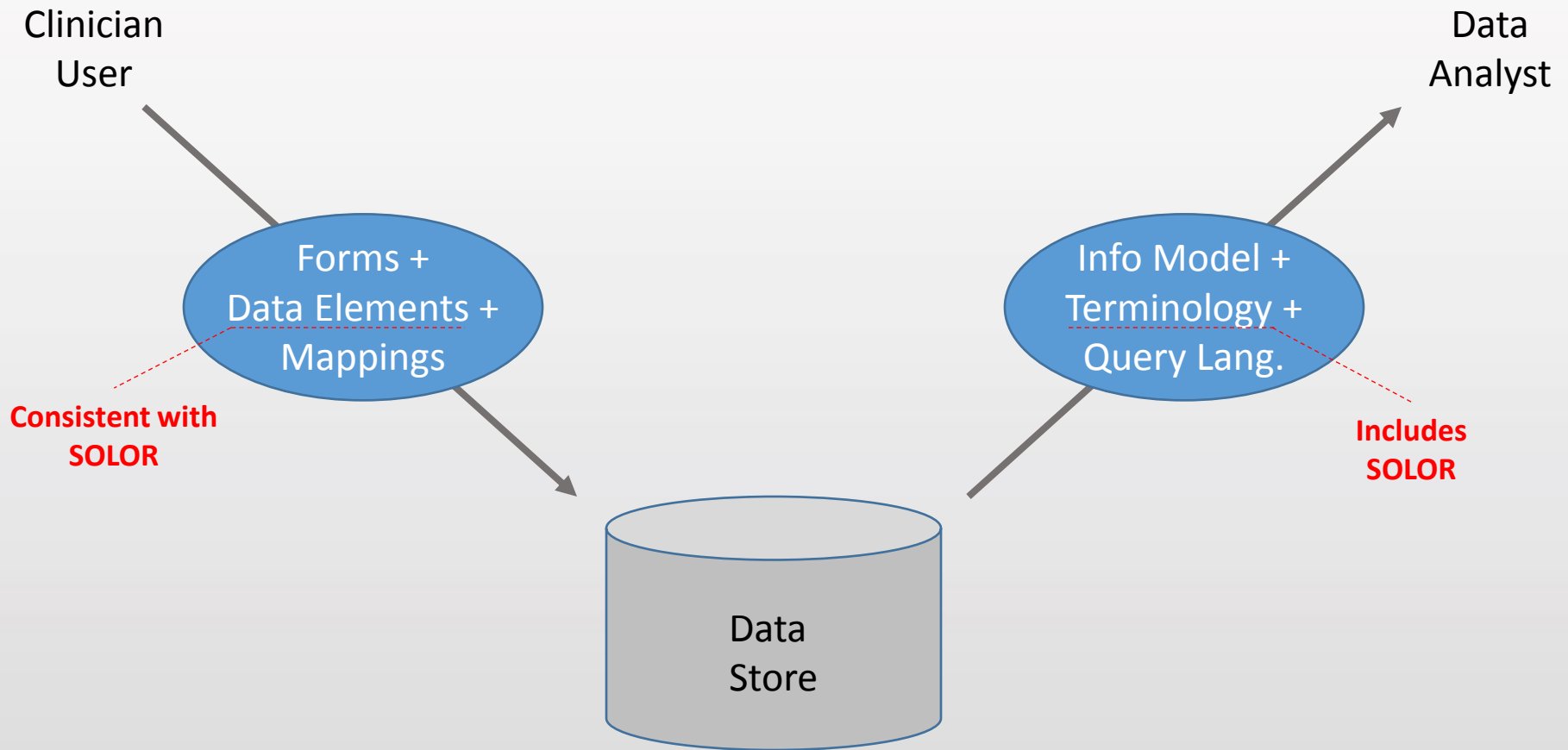
- Effectiveness for *data analysis*
  - No information model independent of individual forms
    - Will data analysts need to understand the content/structure of each form used to collect a data element to understand its full semantics?
    - Will the complete semantics of each data element be the same across each form used to collect it?
    - How will data analysts even know the set of attributes available for querying certain types of objects/events?
      - E.g., Adverse Events (Severity? Manifestation? Outcome? Causative agent? Start Date? End Date? Status? Related treatment protocol? …)
  - Potential for creation of duplicative or overlapping data elements, without means of reconciliation
    - New data elements may be created for each new form, without adequate consideration of existing data elements
    - E.g., "Adverse Event: None, Mild, Moderate, Severe"   vs.   "Severe Adverse Event: Yes, No"

- Create data model that is independent of the structure of the data-entry forms
  - Consisting of information model + terminology
  - Useful for querying and data analysis across varying domains and use cases
- Map the data elements appearing in forms to semantically equivalent representations that are consistent with the data model
  - Data may be translated in real-time, as the forms are submitted
  - Data may be translated later, as data are converted from a forms-based operational data store to a model-based analytical data store

Recommendations

Clinician User

Data Analyst

Forms +
Data Elements +
Mappings

Consistent with SOLOR

Info Model +
Terminology +
Query Lang.

Includes SOLOR

Data Store

Thank you

Questions?