Encoding application profiles in a computational model of the crosswalk
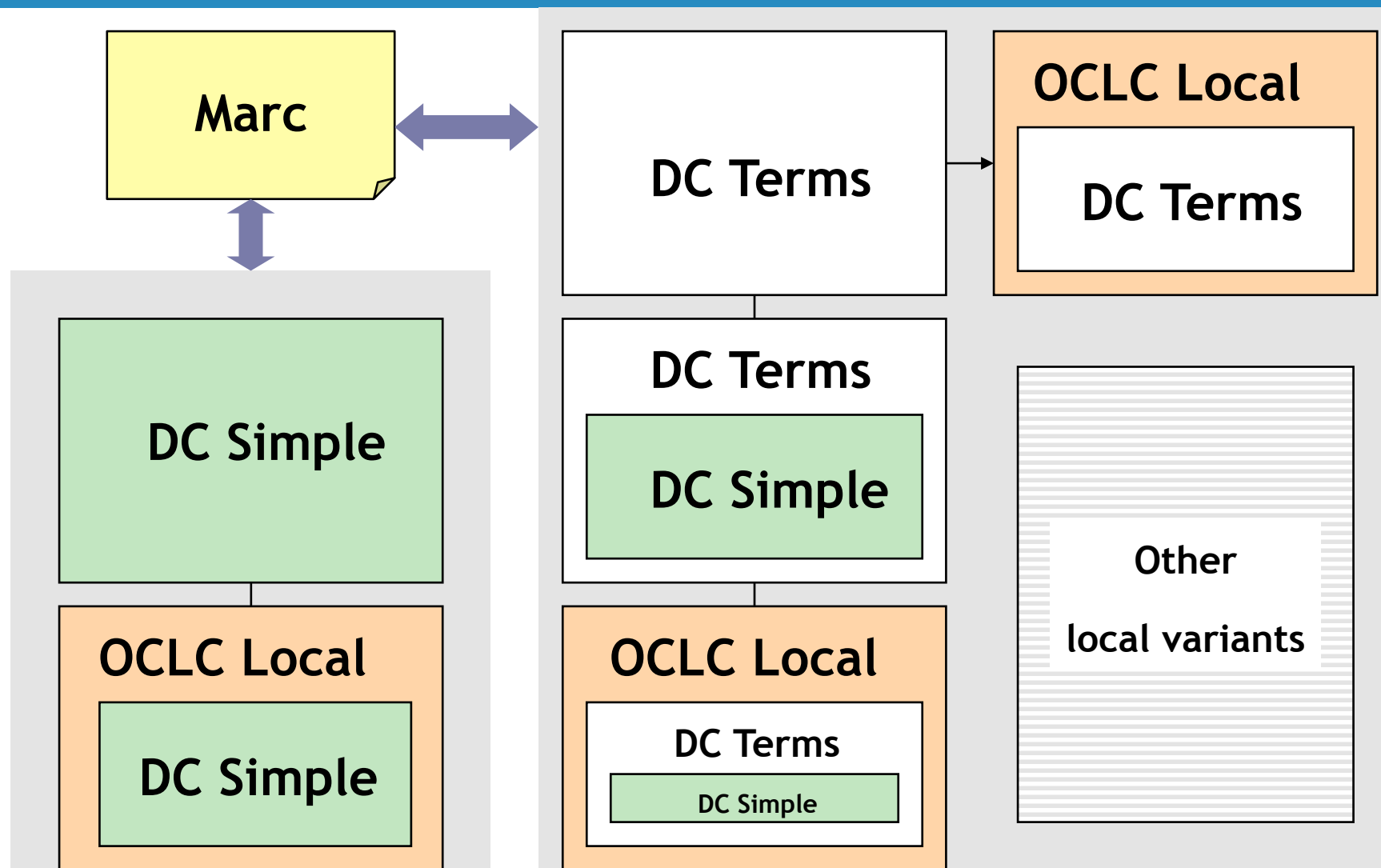
Berlin, Germany

DC-08

23 September2008

Jean Godby, Research scientist

Eric Childress, Consulting product manager

Devon Smith, Consulting software engineer

OCLC™ The world's libraries. Connected.

The MARC-Dublin Core relationship

Three principles in the application profile

OCLC™

1. Draw on one or more existing namespaces.

2. Refine standard definitions—but only by making them narrower, not broader.

3. Introduce no new data elements.

The Crosswalk Web Service at OCLC

- Enables OCLC to translate from one *metadata format* to another.

  - A "metadata format" is a triple that consists of a metadata schema, a structural encoding, and a character encoding.

  - Supported standards are bibliographic, but the software can handle other types of data.

- Can be called from any product or service that processes metadata.

Data flow through the Crosswalk Web Service

OCLC™

MARC input

↓

Convert to input structure

↓

Translate to DC Terms

↓

Convert to output structure

↓

DC Terms output

```
<?xml version="1.0" encoding="UTF-8"?>
<qualifieddc xmlns
      dcterms='purl.org;dc/terms' >
  <dctermsset>
    <dcterms:spatial>
            northwest
    </dcterms:spatial>
  </dctermsset>
</qualifieddc>
```

Input to the translation

DC Simple	DC Terms	MARC tag	MARC subfields	
dc: subject	dcterms: subject	650	a	...
	dcterms: Audience	521	a	
dc: coverage	dcterms: Spatial	522	a	

From crosswalk to translation

OCLC™

```
<translation>
    <header>
        <sourceschema=“marc21” namespace=“uri:marc:21”/>
        <targetschema=“dcSimple” namespace=“uri:dcSimple”/>
    </header>
    <map id=“dcSimple:1”>
        <source>
            <mainpath>
                <branch><step name=“522”/><step name=“a”/></branch>
            </mainpath>
        </source>
        <target>
            <mainpath>
                <branch><step name=“coverage”/></branch>
            </mainpath>
        </target>
    </map>
</translation>
```

A MARC to DC Terms translation

```
<translation>
    <header>
        <sourceschema="marc21" namespace="uri:marc:21"/>
        <targetschema="dcTerms" namespace="uri:dcTerms"/>
    </header>
    <map id="dcTerms:1" override="dcSimple:1">
        <source>
            <mainpath>
                <branch><step name="522"/> <step name="a"/></branch>
            </mainpath>
        </source>
        <target>
            <mainpath>
                <branch><step name="spatial"/> </branch>
            </mainpath>
        </target>    </map></translation>
```

Encoding application profiles in translations

OCLC™

MARC to DC Simple

```
<translation>
<header>
  <sourceschema name='dcSimple namespace='uri:dc:simple'/>
  <targetschema name='marc namespace='uri:marc:21'/>
</header>
<map id='dcsimple:1'>
  <source>
      <mainpath> <branch><step name='650'/><step name='a'/>…
  </source>
  <target>
      <mainpath>
         <branch><step name='subject'/></branch>
      </mainpath>
   </target>
</map>
<map id='dcsimple:2'>
  <source>
      <mainpath> <branch><step name='522'/><step name='a'/>…
  </source>
  <target>
      <mainpath><branch><step name='coverage'/></branch>
      </mainpath>
   </target>
</map>
</translation>
```

```
<translation>
<header>
    <sourceschema name='dcTerms' namespace='uri:dc:terms'/>
    <targetschema name='marc' namespace='uri:marc:21'/>
</header>
<import file='dcSimple2Marc.seel'/>
<map id='dcterms:1'>
  <source>
     <mainpath>
         <branch><step name='521'/><step name='a'/><branch>
     </mainpath>
  </source>
  <target>
     <mainpath>
        <branch><step name='audience'/> </branch>
     </mainpath>
   </target>
</map>
<map id='dcterms:2' override='dcsimple:2>
  <source>
     <mainpath>
         <branch><step name='52l'/>step name='a'/><branch>
     </mainpath>
  </source>
  <target>
     <mainpath>
        <branch><step name='spatial'/></branch></mainpath>
   </target>
</map>
</translation>
```

MARC to DC Terms

File   Cataloging   Authorities   Edit   Action   Batch   View   Tools   Window   Help

Text Strings ▾   User Tools ▾

OCLC   48893831   No holdings in OCL - 6 other holdings

Books ▾   Rec stat   c   Entered   20020130   Replaced   20070315095236.1

Type	a	ELvl	I		Srce	d	Audn		Ctrl		Lang	eng
BLvl	m	Form			Conf	1	Biog		MRec		Ctry	ja
		Cont	b t		GPub		LitF	0	Indx	1		
Desc	a	Ills	a		Fest	0	DtSt	s	Dates	2001	,	

040		OCC ‡c OCC ‡d OCL ‡d OCLCQ ‡d TEF ‡d OCL
020		4924600989
020		9784924600980
041	0	eng ‡a jpn
082	0 4	025.3/4 ‡2 22
090		Z699.35.M28 ‡b I58 2001
092		‡b
049		OCLC
111	2	International Conference on [
245	**1 0**	Proceedings of the Internation Informatics ... [et al. ; edited b
246	1	‡i At head of title: ‡a DC-2001
260		Tokyo : ‡b National Institute o
300		xi, 297 p. : ‡b ill. ; ‡c 30 cm.
546		English and Japanese.
536		Material support for the DC-2 and Information Science (ULI

```xml
<?xml version="1.0" encoding="UTF-8" ?>
- <dctermset>
- <qualifieddc xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.o
    xmlns:dcmitype="http://purl.org/dc/dcmitype/" xmlns:xsi="http://www.w3.org/2001
    xsi:noNamespaceSchemaLocation="http://dublincore.org/schemas/xmls/qdc/2006/0
    <dc:contributor>Oyama, Keizo.</dc:contributor>
    <dc:contributor>Gotoda, Hironobu.</dc:contributor>
    <dc:contributor>Kokuritsu Jōhōgaku Kenkyūjo.</dc:contributor>
    <dc:contributor>Dublin Core Metadata Initiative.</dc:contributor>
    <dc:creator>International Conference on Dublin Core and Metadata Applications (200
    <dcterms:issued>[2001]</dcterms:issued>
    <dc:description>Includes bibliographical references.</dc:description>
    <dc:description>Material support for the DC-2001 conference provided by: National In
        Dublin Core Metadata Initiative (DCMI), Japan Science and Technology Corporatior
        Information Science (ULIS), Communications Research Laboratory (CRL), National
    </dc:description>
    <dcterms:extent>xi, 297 p. : ill. ; 30 cm.</dcterms:extent>
    <dc:identifier>4924600989</dc:identifier>
    <dc:identifier>9784924600980</dc:identifier>
    <dc:language>English and Japanese.</dc:language>
    <dc:language xsi:type="http://purl.org/dc/terms/ISO639-2">eng</dc:language>
    <dc:publisher>National Institute of Informatics</dc:publisher>
    <dc:relation>Papers also available in electronic version.</dc:relation>
    <dc:relation xsi:type="http://purl.org/dc/terms/URI">http://purl.org/dcpapers</dc:
```
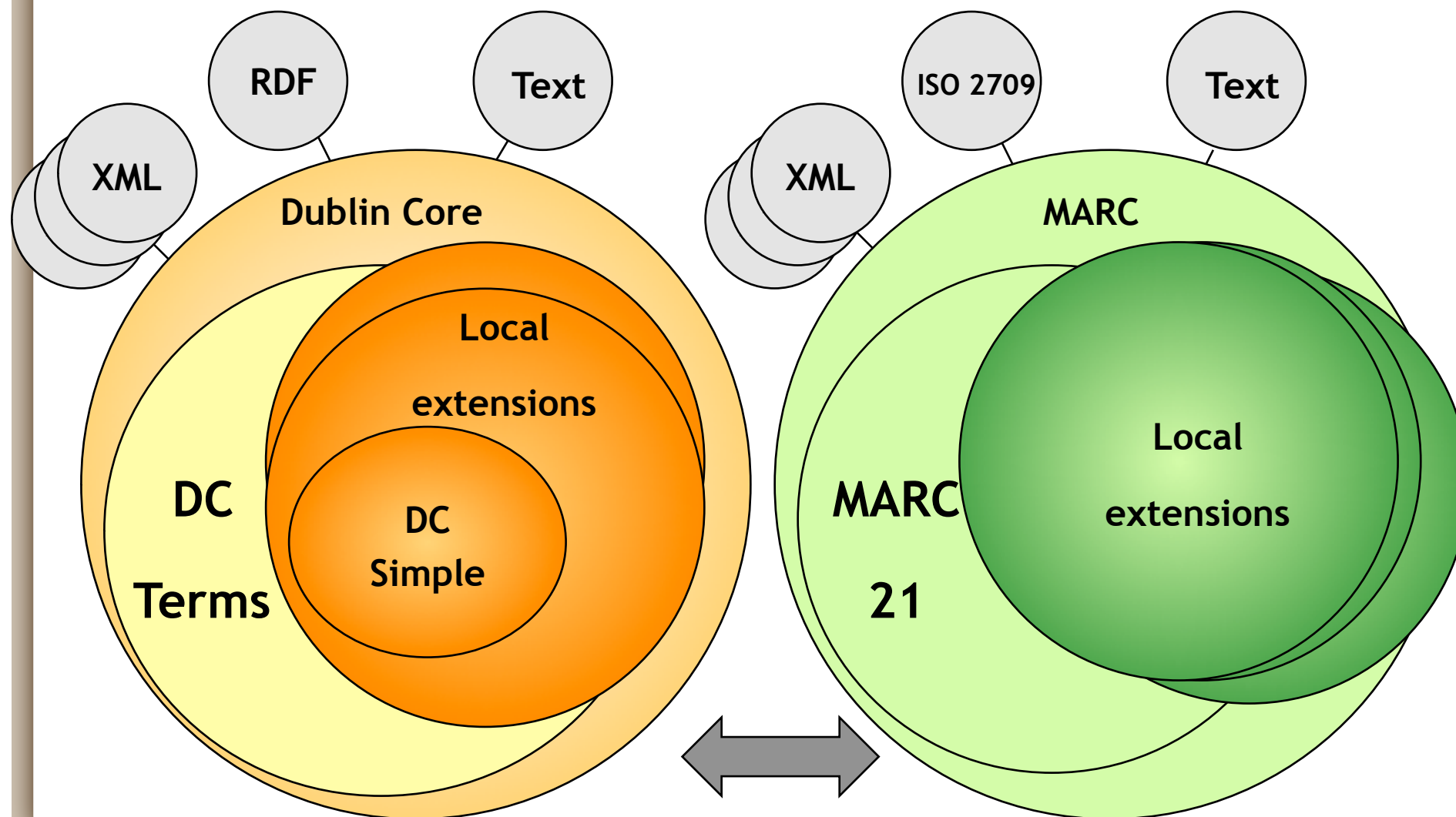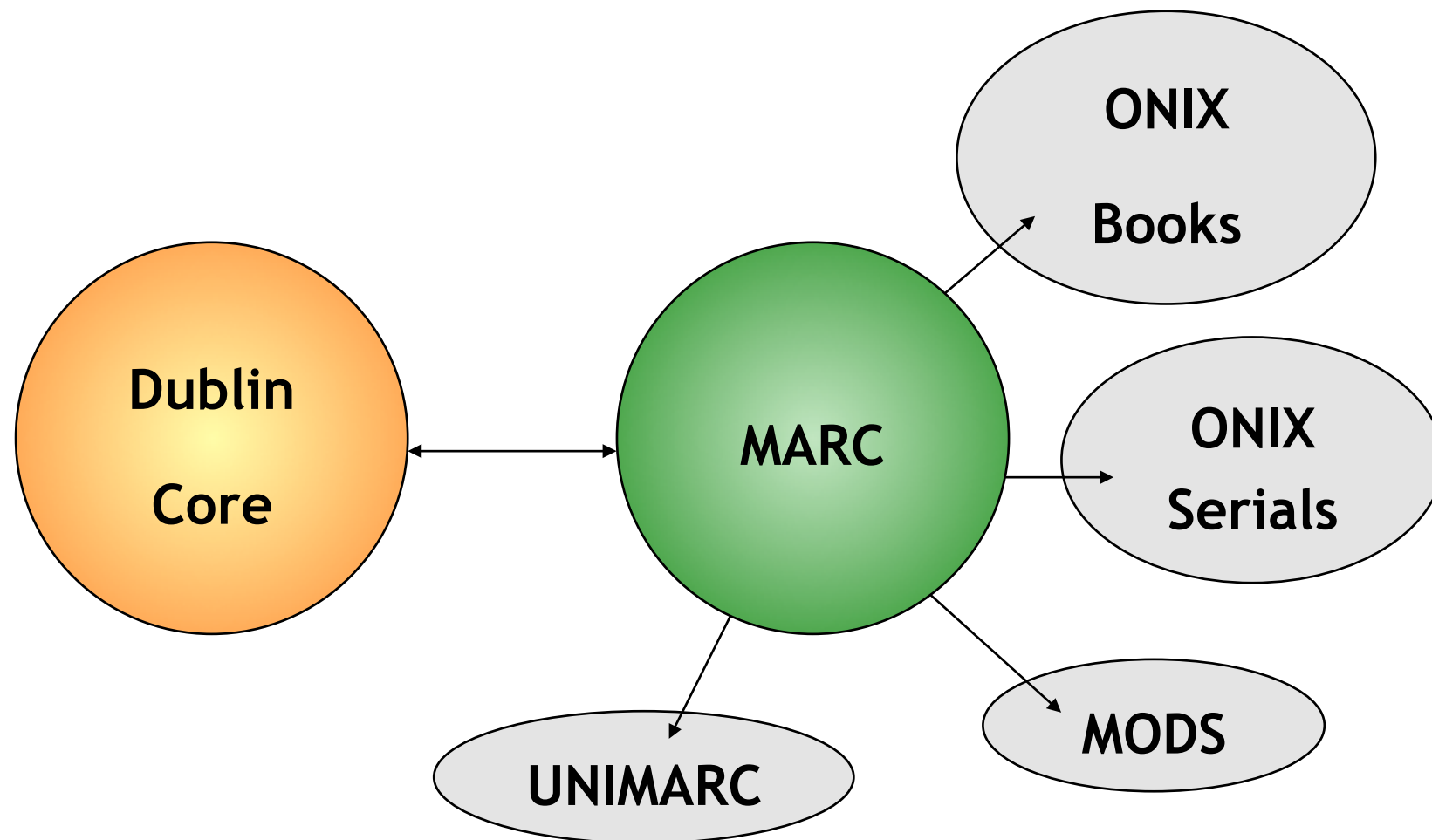
Translations and conversions, expanded

OCLC™

Next Step 2: A graphical user interface

OCLC™

Inputs

```
<map>
Source: MARC 245 $a
Target: DC Title
</map>
```

```
<map>
Source: MARC 650 $a
Target: DC Subject
</map>
```

Editing interface

Map database

Outputs

Search interface

Standard translation
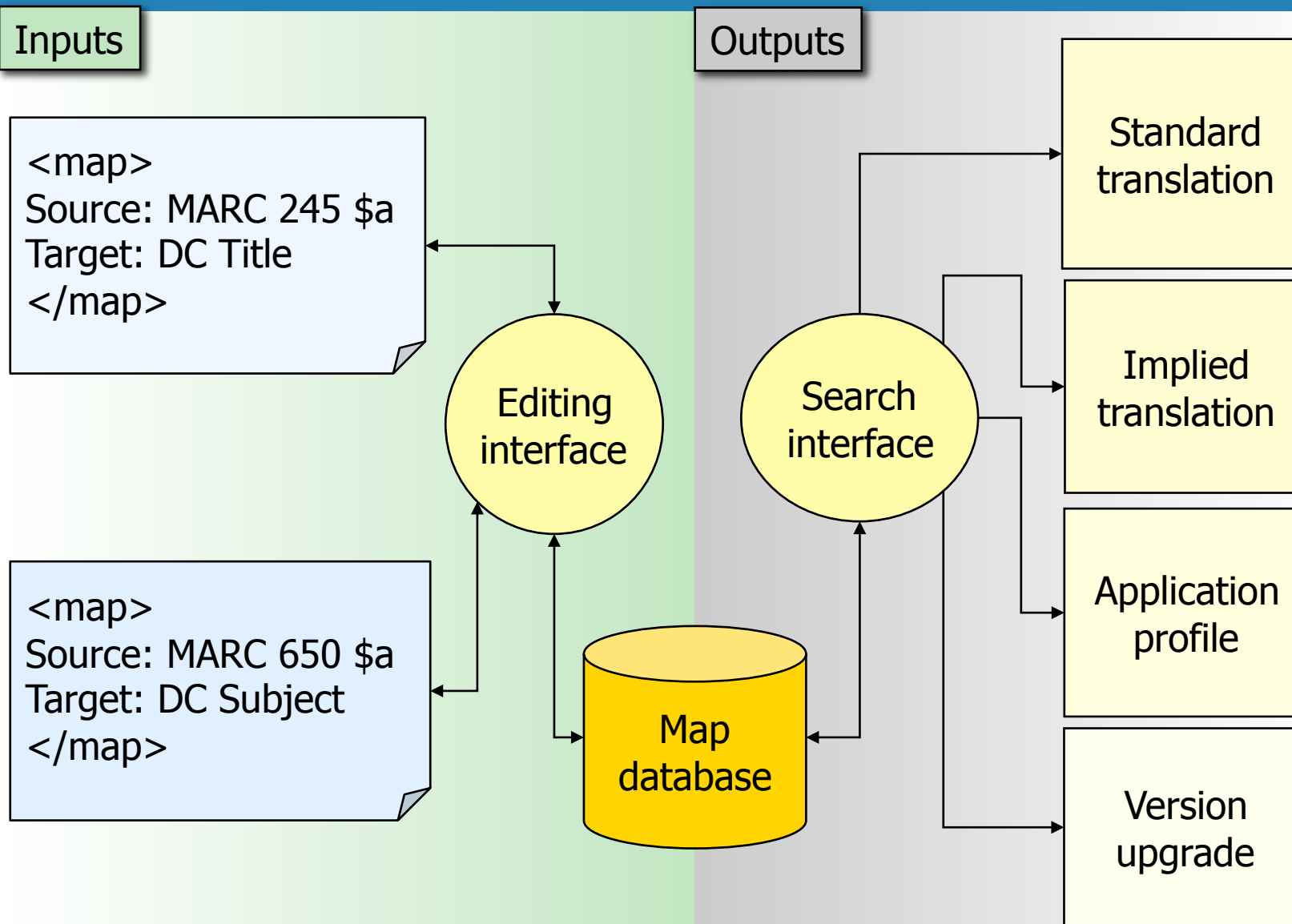
Implied translation

Application profile

Version upgrade

Questions?

For more information

The Crosswalk Web Service public demo. OCLC ResearchWorks.

C.J.Godby, D.Smith, and E.Childress (2008). **"Toward element-level interoperability in bibliographic metadata"** **Code4Lib Journal**