Global Agricultural Concept Scheme

GACS and Agrisemantics

Tom Baker, Caterina Caracciolo, Anton Doroszenko, Lori Finch, Osma Suominen, Sujata Suri

DC-2016, Copenhagen
October 13, 2016

http://agrisemantics.org/gacs/
Who

- Food and Agriculture Organization of the UN
- CAB International (UK)
- National Agricultural Library (US)
### Three overlapping thesauri – *rice, aquaculture, plant pests...*

<table>
<thead>
<tr>
<th>Bibliographic database</th>
<th>Indexed records</th>
<th>Number of concepts</th>
<th>Coverage</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CABI Thesaurus (CABT)</strong></td>
<td>CAB Abstracts</td>
<td>8.3 million</td>
<td>264,500 terms with about 206,400 plant, animal and microorganism names Multilingual in 10 languages</td>
<td>Pure and applied life sciences, technology and social sciences</td>
</tr>
<tr>
<td><strong>NALT</strong></td>
<td>Agricola</td>
<td>5.2 million</td>
<td>&gt; 120,610 terms English and Spanish</td>
<td>In depth coverage of agriculture, biology and related disciplines. Linked Open Data: XML, RDF-SKOS, PDF, MARC and DOC</td>
</tr>
<tr>
<td><strong>AGROVOC</strong></td>
<td>AGRIS</td>
<td>8 million</td>
<td>&gt; 32,000 concepts available in 23 languages</td>
<td>All areas of interest of FAO, including food, nutrition, agriculture, fisheries, forestry, environment etc. Linked Open Data, SKOS-XL</td>
</tr>
</tbody>
</table>
AGROVOC

English, Spanish, Portuguese, German, Czech, Persian, Polish, Hindi, French, Italian, Russian, Japanese, Hungarian, Chinese, Slovak, Thai, Lao, Turkish, Korean, Arabic, Telugu ...

CAB Thesaurus

English, Spanish, Portuguese, Dutch + others

NAL Thesaurus

English, Spanish
Milestones

- 2013-04  G8 - Conference on Open Data in Agriculture & Nutrition
- 2013-10+ GODAN - Global Open Data for Agriculture & Nutrition
- **GACS**
  - 2014-02+ Phase 1: Feasibility Study
  - 2014-06+ Phase 2: Proof of Concept (GACS Alpha)
  - 2015-07+ Phase 3: GACS Beta
  - 2016-05  **GACS Beta 3.1 Soft Launch**
Why

1. To improve the semantic interoperability of thesauri maintained by FAO, CABI, and NAL
2. To provide core concepts broadly supported across the three thesauri
3. To achieve efficiencies of scale by maintaining the core concepts in cooperation
4. “Open data needs common semantics for linking diverse information” (G8)
1. Take the 10,000 most-used concepts from each.
How

2. Automatically map them to each other

Awkwardly mapped clusters were anticipated
3. Verify mappings, one by one

Resolve “lumps” on telecons
Six concepts

Energy intake c:43268

Energy consumption c:43252

Fuel consumption c:49752

Energy use and consumption n:28693

Energy consumption a:c_16121

C1647

Example Lump

c: CABT
n: NALT
a: AGROVOC
Example Lump

About natural resources.

About organisms.

Mappings fixed
Used to create

gacs:

energy intake
nalt:21416

cabt:43268

energy intake
nalt:21416

cabt:43268

gacs:

energy use and consumption
nalt:28693

energy consumption
agrovoc:c_16121

cabt:43252

energy consumption
cabt:43252

energy use and consumption
nalt:28693

cabt:49752

fuel consumption
cabt:49752
...three concepts in GACS Core

- **energy intake**
  - gacs:c5381

- **energy consumption**
  - gacs:c4653

- **fuel consumption**
  - gacs:c20300

BT
...which map back to their sources
GACS – Global Agricultural Concept Scheme

May 2016

GACS Core Beta 3.1
- 15,000+ concepts
- 350,000+ labels in 29 languages
GACS Core Beta

- Concepts considered **reasonably stable**
- **URIs** not *expected* to change, but may
- Some **quality issues** fixed but work remains
- Next: design and **fix a coherent hierarchy**
- **Testing** is encouraged
- **Feedback** welcome!

http://agrisemantics.org/gacs/
GACS objectives

- To support applications of subject description, natural language processing, retrieval, automated indexing, spelling and translation
- To serve as a hub for mapping and integrating related or specialized vocabularies for semantic interoperability
- To improve discoverability of information and data to the benefit of farmers, researchers, and policy makers
GACS in Agrisemantics
Agrisemantics vision

• **GACS as a hub** linking user-oriented thesauri with semantically more precise domain ontologies.

• **Domain ontologies link to datasets** about food and agriculture, in order to make that data more interoperable and reusable.

• GACS is seen as the first step towards improving the *coherence and interoperability* of agricultural data.
Mappings within agrisemantics.org

Hypothetical “Pesticides” vocabulary hosted by agrisemantics.org using agrisemantics.org namespace.
Mappings to external vocabularies

Plant Ontology Base URI: "http://purl.obolibrary.org/obo/"
Agrisemantics
‘Grain’ in Agrovoc

Mappings support mash-ups

‘Grain weight’ in Crop Ontology with methods and units of measurement

‘fruit weight’ in Trait Ontology
Agris abstracts indexed using the term ‘grain weight’

Bibliographic abstracts

1. Correlation and path coefficient analysis in bread wheat under drought stress and normal conditions [Jan 2000]
   Subhoni, G.M.
   Abstract:
   This paper describes the association, direct and indirect effects of morphological traits under drought stress and irrigated conditions. Significant mean squares among crosses were obtained for all characters under irrigation as well as drought stress. Positive and significant correlation were observed between grain yield per plant and flag leaf area, specific flag leaf weight, tillers per plant ...

2. Genotypic competition among elite wheat breeding lines under irrigated and rainfed conditions [2012]
   Khan, A.; Mohammad, F.; Hasson, G.; Khalil, I.H.
   Abstract:
   Development of high yielding cultivars is one of the prime objectives of all wheat breeding programs. An experiment was conducted during 2008-09 at Agriculture Research Farm, Khyber Pakhtunkhwa Agricultural University, Peshawar, Pakistan in which fifteen wheat lines were evaluated for yield and yield contributing traits under irrigated and rainfed field conditions. Lines exhibited significant ...

3. Effects of all straw returned to the field on grain number and grain weight at different spikelets and grain positions in winter wheat [May 2011]
Research data

CIMMYT Research Data Dataverse

18TH SEMI-ARID WHEAT YIELD TRIAL
hd1:11528/10170
Version: 2- Released: Mon Sep 14 06:43:43 CDT 2015

CATALOGING INFORMATION

<table>
<thead>
<tr>
<th>Data Citation</th>
</tr>
</thead>
</table>

Citation Format: Print:

Publications

| 18th SAWYT |

Subtitle

| 18th SAWYT Cycle 2010 |

Study Global ID

| hd1:11528/10170 |

Authors

| Singh, Ravi (CIMMYT); Payne, Thomas (CIMMYT) |

Publisher

| International Maize and Wheat Improvement Center (CIMMYT) |
### Data files

<table>
<thead>
<tr>
<th>Trial name</th>
<th>Occ</th>
<th>Loc_no</th>
<th>Country</th>
<th>Loc_desc</th>
<th>Cycle</th>
<th>Cid</th>
<th>Sid</th>
<th>Gen_name</th>
<th>Trait</th>
<th>Trait name</th>
<th>Gen_no</th>
<th>Rep</th>
<th>Sub_block</th>
<th>Plot</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>61665</td>
<td>1</td>
<td>LOCAL CHECK</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>31.5</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>95758</td>
<td>12</td>
<td>DHARWAR DRY</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>27.1</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>24762</td>
<td>30</td>
<td>CHAM 6</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>24.1</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>462733</td>
<td>42</td>
<td>HSB 1313/2*WBL1</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>21</td>
<td>1</td>
<td>3</td>
<td>21</td>
<td>26.5</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>479309</td>
<td>345</td>
<td>SOKOLL/EXCALIBUR</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>23</td>
<td>1</td>
<td>3</td>
<td>23</td>
<td>31.8</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>485682</td>
<td>62</td>
<td>PI 610750/LANG</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>25</td>
<td>1</td>
<td>3</td>
<td>25</td>
<td>25.5</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>485792</td>
<td>55</td>
<td>ALTAR B4/AE.SQUARROSA (219)</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>26</td>
<td>1</td>
<td>3</td>
<td>26</td>
<td>24.5</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>485792</td>
<td>56</td>
<td>ALTAR B4/AE.SQUARROSA (219)</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>27</td>
<td>1</td>
<td>3</td>
<td>27</td>
<td>27.4</td>
<td>g</td>
</tr>
<tr>
<td>18TH SEMI-ARID WHEAT YT</td>
<td>1</td>
<td>61702</td>
<td>SERBIA</td>
<td>RINSKI SANCEVI</td>
<td>2010</td>
<td>485834</td>
<td>26</td>
<td>SOKOLL/SUNCO/2*PASTOR</td>
<td>147</td>
<td>1000_GRAIN_WEIGHT</td>
<td>28</td>
<td>1</td>
<td>3</td>
<td>28</td>
<td>28.5</td>
<td>g</td>
</tr>
</tbody>
</table>
RDA Agrisemantics WG

• Agrisemantics Working Group of the Research Data Alliance
• A forum for knowledge and practices related to semantic interoperability in agriculture
• Convenes stakeholders to design the Agrisemantics platform

https://rd-alliance.org/groups/agrisemantics-wg.html
Agrisemantics objectives

- **Open-access** semantics are easy to re-use.
- They make agricultural data easier to re-use.
- Mapping the semantics promotes cooperation and reduces duplication.
- Coherent semantics benefit research, innovation systems, and value chains.
http://agrisemantics.org/gacs/