
Karl Morteo, Project Manager, IT Division, UN-FAO  Karl.Morteo@fao.org
Topics

- Animal Production & Health
- Economics & Policy
- Education & Extension
- Engineering, Technology & Research
- Farming Practices & Systems
- Fisheries & Aquaculture
- Food Security
- Food safety & Human nutrition
- Forestry
- Natural Resources & Environment
- Plant Production & Protection
- Rural & Social Development
- Climate, Energy and Tenure
- Land and Water
- Agriculture
Recommendations from Evaluations, Review and Consultations

- **Unifying** the fragmented and “linear” information storage and dissemination in the area of the **technical systems**
- **Consolidating** the fragmented and duplicated Information Systems and Information Management **environments** presently distributed across many organizational units
- A top priority for the Information Technology Division is **Consistent, Integrated, Authoritative and Easy to Access Corporate Data Repository**
Uniting Our Data

http://data.fao.org is a one-stop shop that aggregates, integrates, and catalogues data from multiple sources within the Organization. These entries cover topics related to nutrition, food and agriculture and include data such as statistics, maps, pictures, documents and more.
Overview of Deliverables

- **Web Site**: [http://data.fao.org](http://data.fao.org) for data consumers, managers and information systems integrators
- **Data Warehouse** Repository (supporting statistics, maps and digital assets)
- Tools for **data and metadata management** including support for data hotels
- A platform for integration and exchange
- **Integration** of existing **information systems**
- **Harmonisation** and **de-duplication** of data
- **Widgets** to embed services in websites, including charts, maps and content
- Advocacy and communication
- Project management and administration
**Mantra**

1. **Uniting our data** – data.fao.org tagline

2. **Engage not just disseminate** (harness the crowd) #FAOdata

3. **Mobile First** – J. Boye – Morgan Stanley

4. **Serve the data in the most convenient format** for all users - Linked and Open Data

5. **Eat your own dog food** -
   
The Website

Should be the home screen
Website - Use cases

1. Browse and share
2. Find and use (e.g. journalists preparing reports)
3. Query and analyze
4. Just give me the data
Site structure

- Full text search
- 6 official languages
- Login
- Help

Navigation by Topics, Concepts, Tags, Places, Catalogue

Structured search

Filter by object types
So you're more of a Stats person, right?

Commodo vitae, ornare sit amet, wisi. Aenean fermentum, elit eget tincidunt condimentum, eros ipsum rutrum orci, sagittis tempor lacus enim ac dui. Donec non enim in turpis pulvinar facilisis. Ut felis. Praesent dapibus, neque id cursus faucibus, tortor neque egestas augue, eu vulputate magna eros eu erat. Aliquam erat volutpat. Nam dui mi, tincidunt quis, accumsan porttitor, facilisis luctus, metus

Maps

Documents

Pictures
Landing Page

Algeria

COUNTRY AREA
238 174 [1000 Ha]
FAOSTAT Data reported on country official publications or web sites (Official) or trade country files, 2009.

LAND AREA
238 174 [1000 Ha]
FAOSTAT Data reported on country official publications or web sites (Official) or trade country files, 2009.

AGRICULTURAL AREA
41 325 [1000 Ha]

POPULATION - EST. & PROJ.
36 486 [1000]

GDP
159 426
[millions of US dollars]

HUMAN DEVELOPMENT INDEX
0.6
UNDP - HDRO Not available, 2010.

Country Codes
ISO3 : DZA
ISO2 : DZ
UN : 12
GAUL : 4
UNDP : ALG
FAOSTAT : 4

This country is part of ... 
Geographic and Economic groups (3) »
northern Africa, Africa, Arab Maghreb Union,
ITS RELATED TO ...
Land border countries (6) »
Libya, Mali, Mauritania, Morocco, Niger, Tunisia,
IT CONTAINS ...

07/06/2012
data.fao.org - European Data Forum 2012
Engage, not just disseminate

Food security situation in East Africa September 2011

Geographic bounding box

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North bound latitude</td>
<td>11.979166</td>
</tr>
</tbody>
</table>

Private resource
- 311 likes
- 2K views
- 25 comments
- 25 downloads
- 67 tags
- 11 cites

Tagged Under
- CROP YIELDS (1)
- CROPS (4)
- MAIZE (1)
- RICE (1)
- WHEAT (2)

This map is part of ...
Topics (1)
Engage, not just disseminate

Food security situation in East Africa

Geographic bounding box

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North bound latitude</td>
<td>11.979166</td>
</tr>
</tbody>
</table>

Show me the data

Link To It! In 3 Easy Ways

This map is part of ...
Engage, not just disseminate

Data in the most convenient for you

RDF, JSON, XML, CSV, Excel, RoamBi, GeoTIFF, WMS
Engage, not just disseminate

Login and share with popular Social Networks:
- Google+
- Facebook
- Twitter

Expertise at FAO:
- Curate
- Create
- Link, Tag, Share
- Comment
### Crops, National Production

#### About

#### Data

**YOU SELECTED**

- **COLUMNS**
  - YEAR (2001 - 2011)

- **ROWS**
  - ALL AREA
  - WHEAT
  - AREA HARVESTED [HA]

**Filter by when**

#### Show 20 entries

<table>
<thead>
<tr>
<th>Area</th>
<th>Crops, National Production</th>
<th>Measures</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>Wheat</td>
<td>Area Harvested [Ha]</td>
<td>25 730 600.00</td>
<td>26 344 700.00</td>
<td>25 195 700.00</td>
</tr>
<tr>
<td>China</td>
<td>Wheat</td>
<td>Area Harvested [Ha]</td>
<td>24 664 068.00</td>
<td>23 908 072.00</td>
<td>21 997 075.00</td>
</tr>
<tr>
<td>United States of America</td>
<td>Wheat</td>
<td>Area Harvested [Ha]</td>
<td>19 616 500.00</td>
<td>18 544 500.00</td>
<td>21 474 100.00</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>Wheat</td>
<td>Area Harvested [Ha]</td>
<td>22 832 700.00</td>
<td>24 477 700.00</td>
<td>20 005 600.00</td>
</tr>
<tr>
<td>Australia</td>
<td>Wheat</td>
<td>Area Harvested [Ha]</td>
<td>11 529 000.00</td>
<td>11 170 000.00</td>
<td>13 067 000.00</td>
</tr>
</tbody>
</table>
Mobile First - Responsive Design
Data Loading

10 of 16 knowledge divisions

356,000,000 Statistical values
734,000 Geo Layers
1,500,000 Statistical Maps

435 Documents
90 Pictures

40+ Terabytes (2 Stats, 8 Maps online)
25 Information Systems
finding data will be this easy...

Release One

12 December 2012

For more information, visit us at data.fao.org
Guiding Principles

- **Custodian only** providing the receptacle, the data ownership, rights and responsibilities stay with the contributor
- Will not replace existing thematic ‘brands’
- **Open** to all FAO contributors
- **One stop shop** single point of access and reference for all data
Data Flow and Integration

Data Source

Ingest

Harmonise
Integrate
Enrich

Publish
data.fao.org - The Big Picture

Website

Services and Widgets

Specialised application(s) consume/provide

Orchestration and Integration

Search
Full text
Structured

Catalogue
Identity
Metadata
Linked Data ...

Statistics
Statistical Data
Warehouse
Time Series
Indicators
Observations

Maps
Geospatial
Raster
Vector
Point

Content
Documents
Pictures
Video
Multimedia
Pages

Infrastructure
Logging
Caching
Security
Audit ...

07/06/2012  data.fao.org - European Data Forum 2012
Catalogue, UUIDs and APIs

Every asset is registered in the Catalogue with a Universally Unique Identifier (UUID).

UUID: 11ff9540-be95-46c6-b2bc-c5874d21d4d1

With a corresponding URI http://ref.data.fao.org/UUID

Assets can also be accessible and managed through an Application Programming Interfaces (API)

http://api.data.fao.org/resources/images/countryprofiles/flag-arg
Types of Metadata

- **Structural**, typically derived from physical data, organization and/or structure. Examples include the dimensions of a data cube, number of facts in a data set or image file format.

- **Reference**, describes the data and is the information used to understand the data. Examples include title, author, subjects, keywords, publisher, method of collection and processing, data source and references.

- **Administrative**, technical information regarding the data storage and rights management including: ownership, permissions, visibility, security, rights back up, audit, and preservation.

- **Social**, describes how the community consume, use and categories the data. Examples include: number of citations, download counts, user rating, comments and tagging.
Types of Metadata

- **Structural**, typically derived from physical data, organization and/or structure. Examples include the dimensions of a data cube, number of facts in a data set.

- **Reference**, describes the data and is the information used to understand the data. Examples include title, author, subjects, keywords, publisher, method of collection and processing, data source and references.

- **Administrative**, technical information regarding the data storage and rights management including: ownership, permissions, visibility, security, rights backup, audit, and preservation.

- **Social**, describes how the community consume, use and categorize the data. Examples include: number of citations, download counts, user rating, comments and tagging.

*The cure for information overload is...* more information!
### Metadata Standards

<table>
<thead>
<tr>
<th>Asset</th>
<th>Standards / Formats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics</td>
<td>SDMX, DDI</td>
</tr>
<tr>
<td>Maps</td>
<td>OGC, ISO19115</td>
</tr>
<tr>
<td>Documents</td>
<td>Dublin Core, MODS, FRBR</td>
</tr>
<tr>
<td>Pictures</td>
<td>IPTC, XMP</td>
</tr>
<tr>
<td>Tags</td>
<td>more</td>
</tr>
</tbody>
</table>
UN-FAO corporate, multilingual thesaurus of topics related to agriculture, forestry, fisheries, environment and more. Available in more than 20 languages, with an average of 40,000 terms. Published as Linked Data and aligned with twelve vocabularies. http://aims.fao.org/agrovoc/lo
AGROVOC

UN-FAO corporate, multilingual thesaurus of topics related to agriculture, forestry, fisheries, environment and more. Available in more than 20 languages, with an average of 40,000 terms. Published as Linked Data and aligned with twelve vocabularies: [http://aims.fao.org/agrovoc/lo](http://aims.fao.org/agrovoc/lo)

Information Backbone
Linking inside and outside
Exact/Broader/Narrower/Related
LOD2: Publink Consultation Program

- Experience, Guidance and Best Practices
- Linked Data Model to expose social metadata
- Alignment and LOD-ification

Central Product Classification in RDF integrated in the data catalogue of data.fao.org
UN-FAO geopolitical ontology. Available in 6+1 languages. Includes FAO, UN and internationally recognized data sources. Records historical changes.

Published as Linked Data, RDF and OWL

FAO Geopolitical Ontology

Names (short, official), Codes,

Area types, Groups membership

Supporting data: coordinates, Land borders, Currency names and codes, Adjectives of nationality,

Basic statistical data: country area, land area, agricultural land area, population and GDP
APIs, Widgets and Embedded Content

GDP 12 015
[millions of US dollars]

http://review-data.fao.org/developers

07/06/2012
data.fao.org - European Data Forum 2012
What’s a widget?

Program that performs simple function and can be embedded into a webpage (e.g. weather report or stock quote)

http://review-data.fao.org/developers
Widgets – Country Profiles

Country Profiles = FAO site + Widgets
Country Information

- FAO Representations websites (Sub-regional Office for the Pacific Islands)
- International Plant Protection Convention Country Profiles [www.ippc.int](http://www.ippc.int)
Web Application Example - Key Water Indicator Portal
Web Application Example - Key Water Indicator Portal

http://www.unwater.org/statistics_KWIP.html
Global Agro-Ecological Zones

Information at 3 inputs levels (high, intermediate, low) on:

- agro-ecological zones
- agro-climatically attainable yields
- yield constraints
- crop calendars
- agro-ecological suitability and productivity assessment
- potential production estimates
- actual yield and production
- yield and production gaps
Global Agro-Ecological Zones

- **THEMES**
  - 5 thematic areas (> 300,000 global datasets at mainly 5 arc-minutes, also core layers at 30 arc-seconds)

- **CROPS**
  - 11 crop groups, 49 crops, 92 crop types and 280 Crop/LUTs
  - downscaling of 23 crops/commodities are available for 23 crop/commodities
  - yield and production gap analysis for 17 crops/commodities

- **WATER SUPPLY**
  - 5 water supply types (rain-fed, irrigation, gravity, sprinkler, drip)

- **INPUT LEVELS**
  - 4 Input levels (High, Intermediate, Low, Mixed)

- **TIME PERIOD**
  - Historical 1961-2000, 30 year average (1961-1990) and Future, 2020s, 2050s, 2080s
http://gaez.fao.org
data.fao.org - The Technology

Orchestration and Integration

Search
- Full text
- Structured

Catalogue
- Identity
- Metadata
- Linked Data
...

Statistics
- Statistical Data Warehouse
- Time Series
- Indicators
- Observations

Maps
- Geospatial
- Raster
- Vector
- Point

Content
- Documents
- Pictures
- Video
- Multimedia
- Pages

Infrastructure
- Logging
- Caching
- Security
- Audit
...

Specialised application(s) consume/provide
Technologies – Just 1 or 2 ;-) 


Technologies – Just 1 or 2 ;-)  


Some Big Bits

- GeoNetwork
- ESB and BPM
- Alfresco
- Java
- JBoss
- Ontotext
- GeoServer
- Fedora Commons

Clustered and Load Balanced

PostgreSQL ~2TB

07/06/2012 data.fao.org - European Data Forum 2012
Infrastructure

- Fine-Grained monitoring and alert – RHQ + Nagios
- Backup: VM images and DB redundancy
- 14* Physical Servers, 40+ Production VMs
- 100 TB SAN, 24 port 1 Gigabit switch

**DevOps:** Developers and SysAdmins working together as one to deliver agile, dependable and responsive solutions:

*Changing the way we do IT*
Build Lifecycle – Continuous Integration

- Development (Java, Maven2)
  - Subversioning (SVN)
  - Bug & Issue tracking (JIRA)
- Continuous Integration (Jenkins)
  - Build Repository (Artifactory)
- Unit Testing
- Integration Testing (Cactus)
- Production Server
Thank you for listening ... 

Questions?