SURVEY OF ISRAEL

MINISTRY OF CONSTRUCTION AND HOUSING

BUILDING NATIONAL DATABASES USING EXPERT SOURCING

Prof. Yaron Felus
Chief Scientist

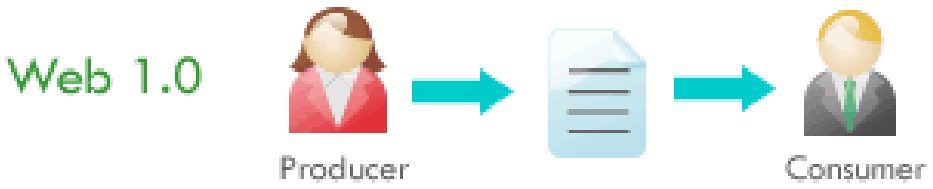
20 November, 2018

Contents

- **Background on internet Eras**
- What is a data lossless nation?
  - Data products in the construction process
  - National mapping standard
  - The Inter-agency GIS committee
- On-line QC tool
- The data integration tool
- Conclusions

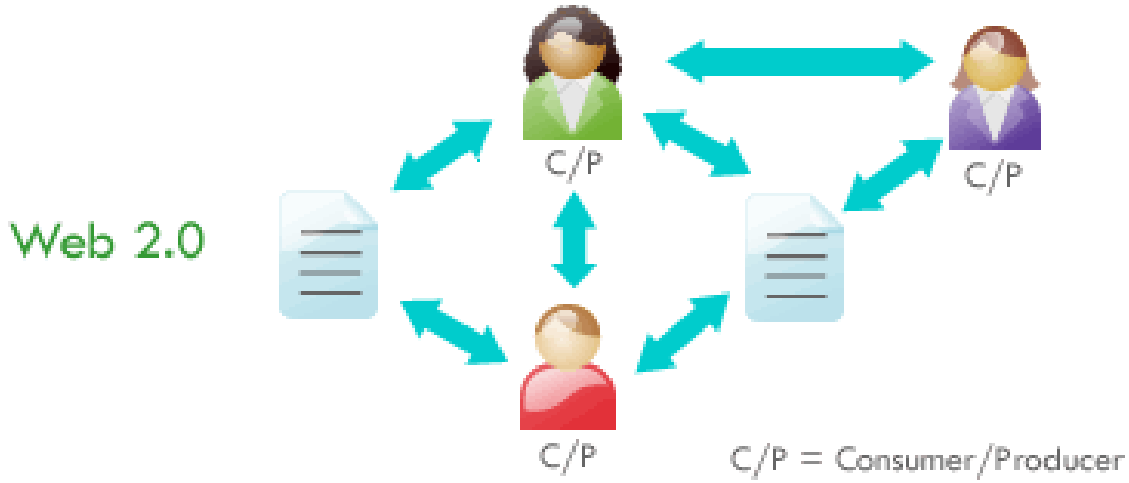
Internet eras

One organization providing information (Read)

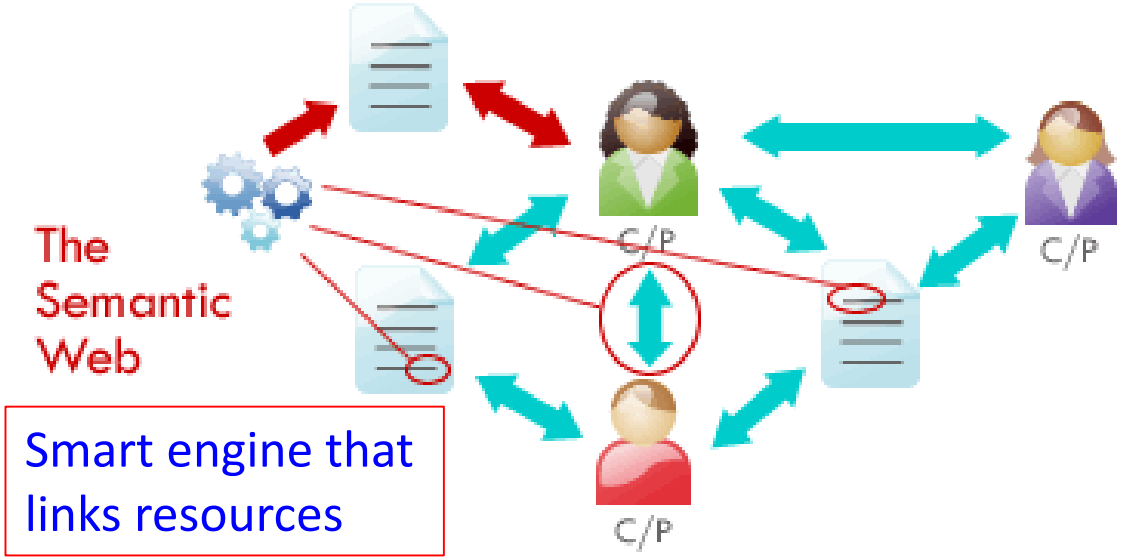


Collaboration with the public (Read and write)

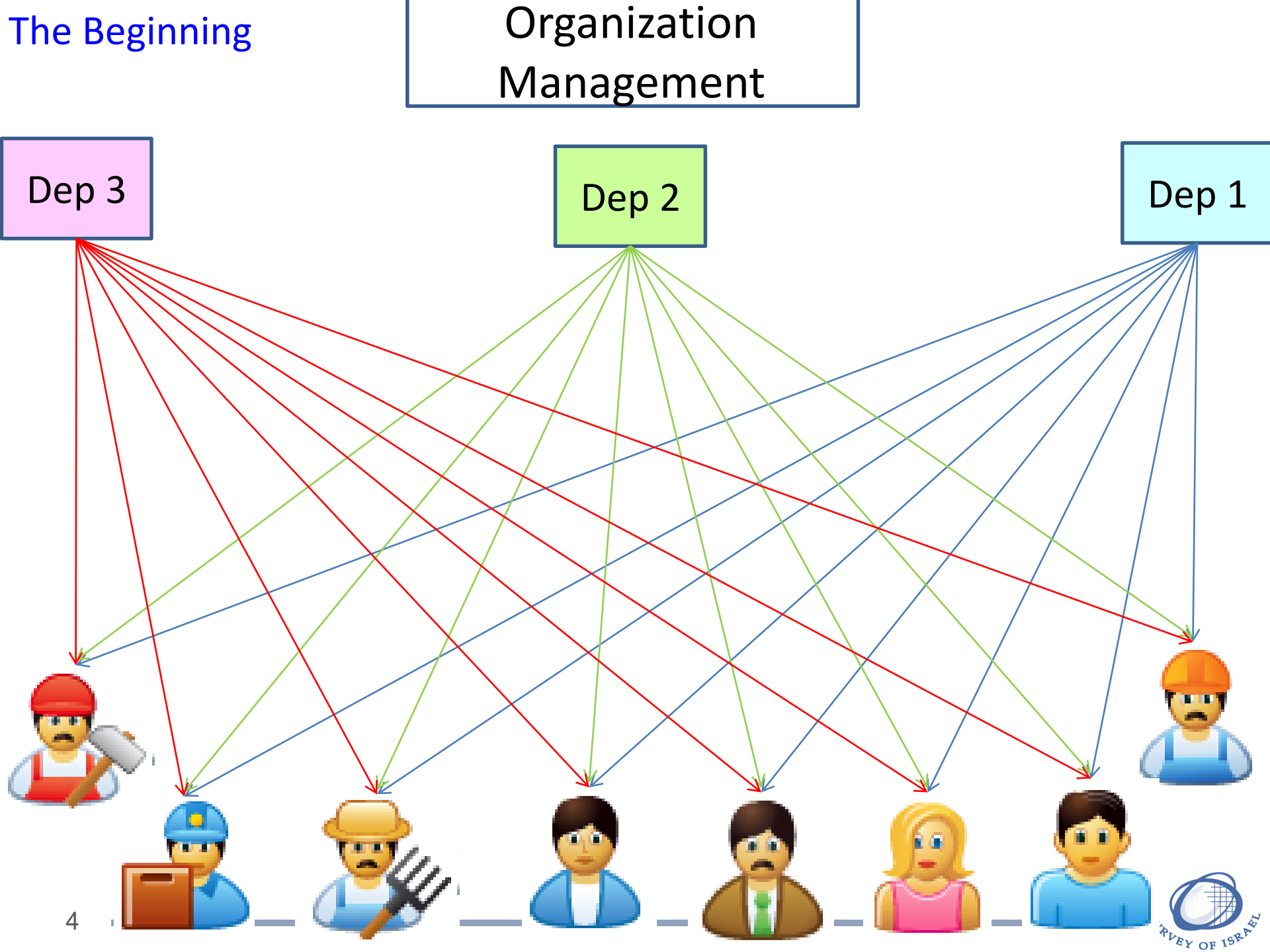
Wiki, facebook, youtube, waze



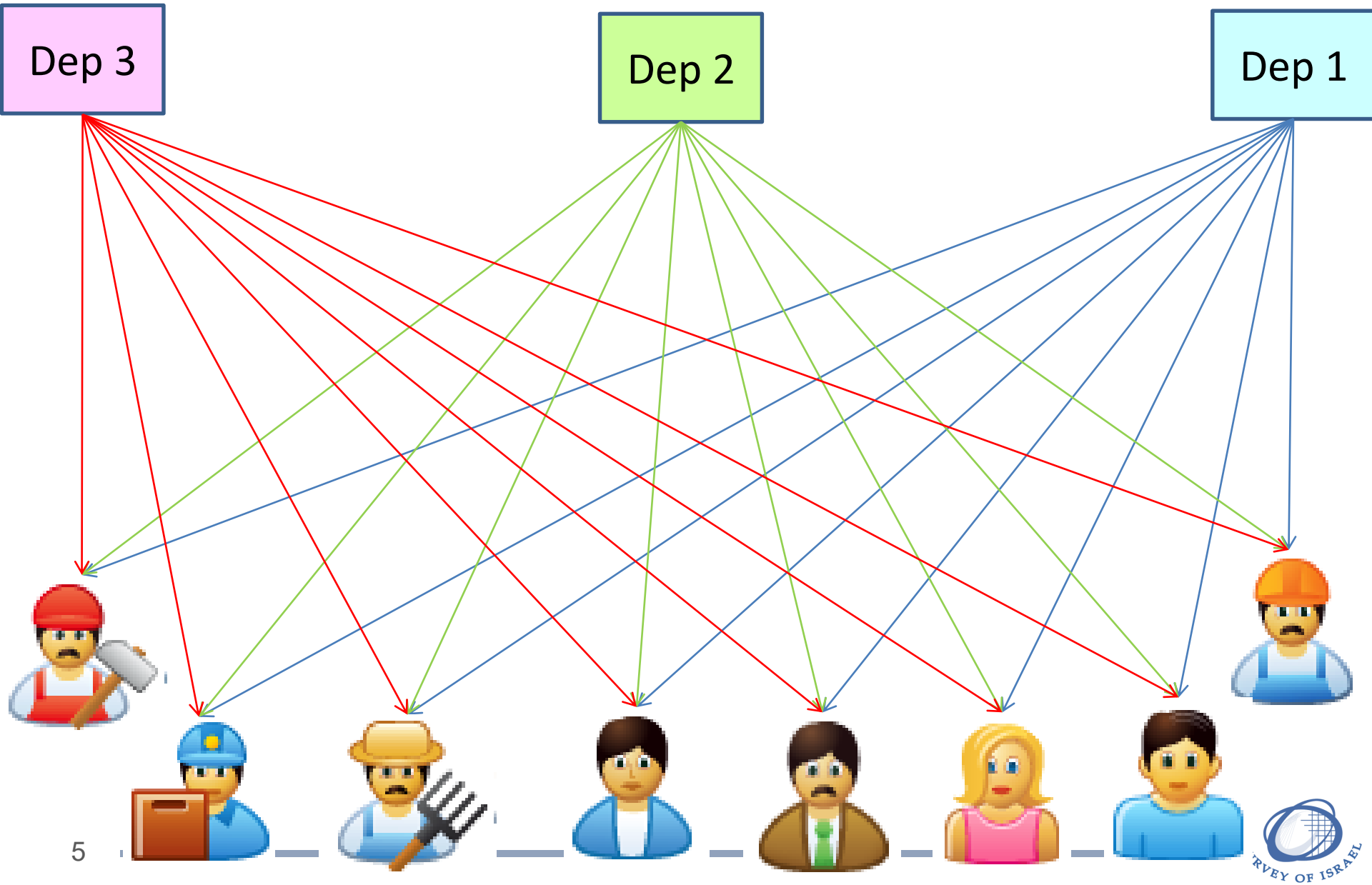
Web 3.0: Smart and connected system (Read, write and understand)



Smart engine that links resources

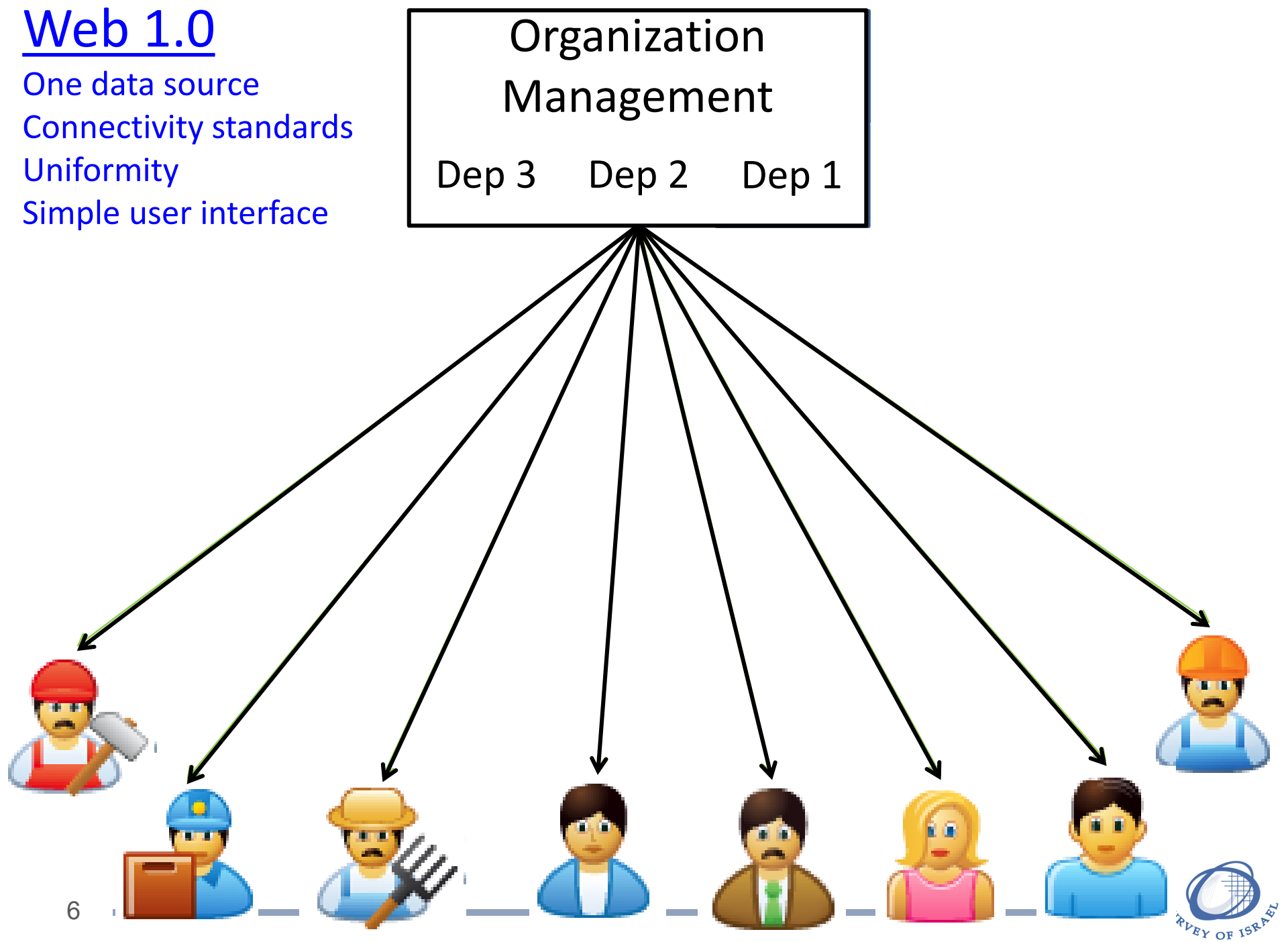

Organization Management
Dep 3
Dep 2
Dep 1

Organization Management



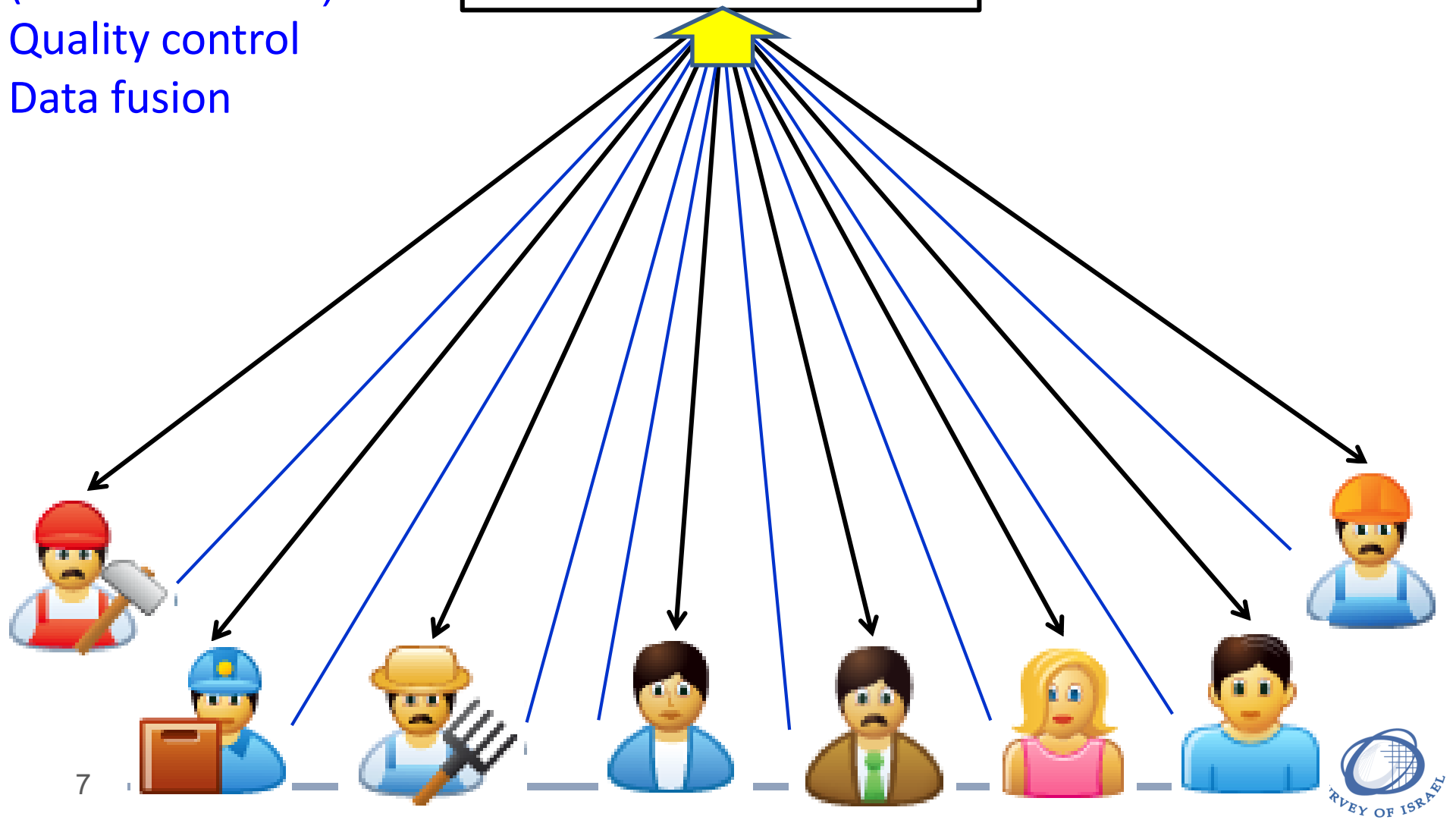
Web 1.0

One data source
Connectivity standards
Uniformity
Simple user interface

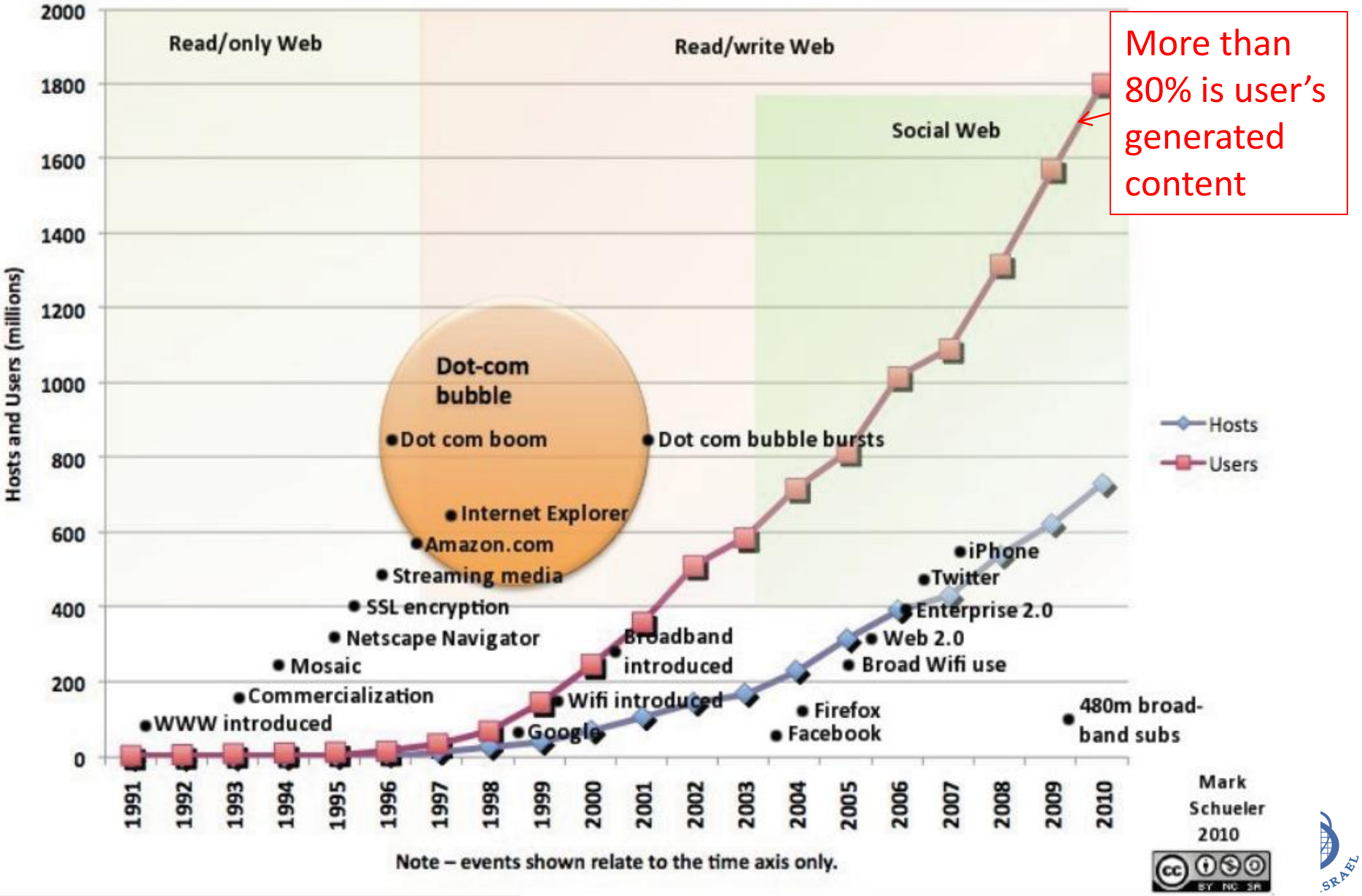


Web 2.0

Collaboration with the public
(Read and write)
Quality control
Data fusion



The world is ready for that…

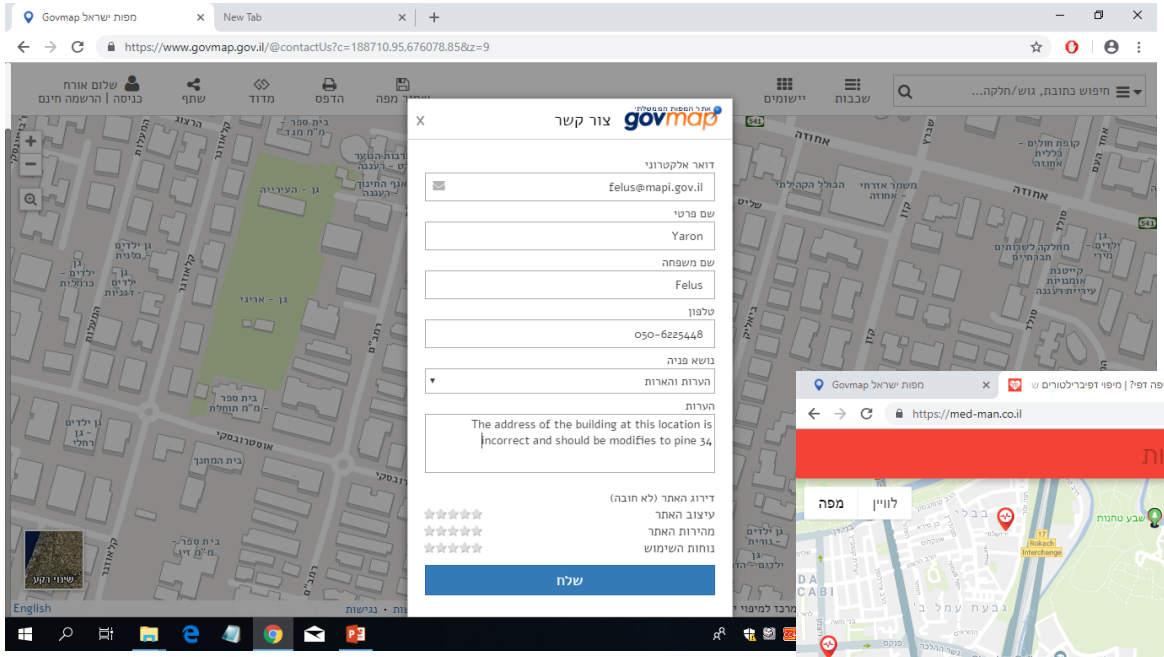


Crowdsourcing Vs wisdom of the crowd Vs Expert Sourcing

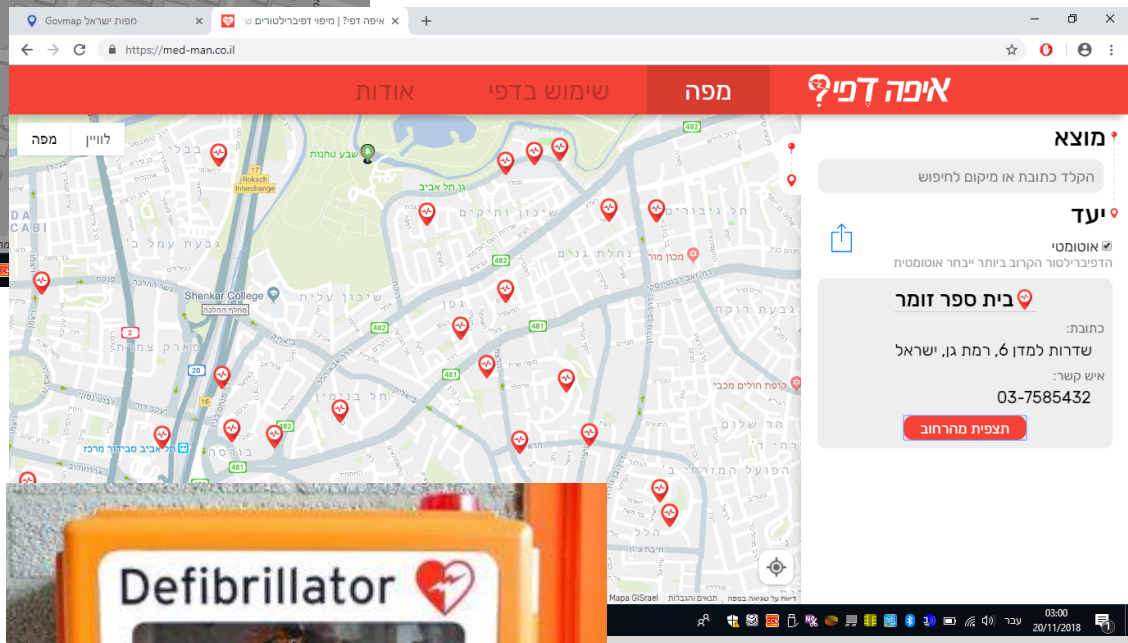
- **Crowdsourcing:** the act of obtaining information from a large group of people
- **Wisdom of the crowd** is the collective opinion of a group of individuals
- **Expert-sourcing:** the act of obtaining information from a large group of experts

SURVEY OF ISRAEL

Crowdsourcing applications



Feedback and notifications on problems in Topographical DB at the National Geospatial Portal



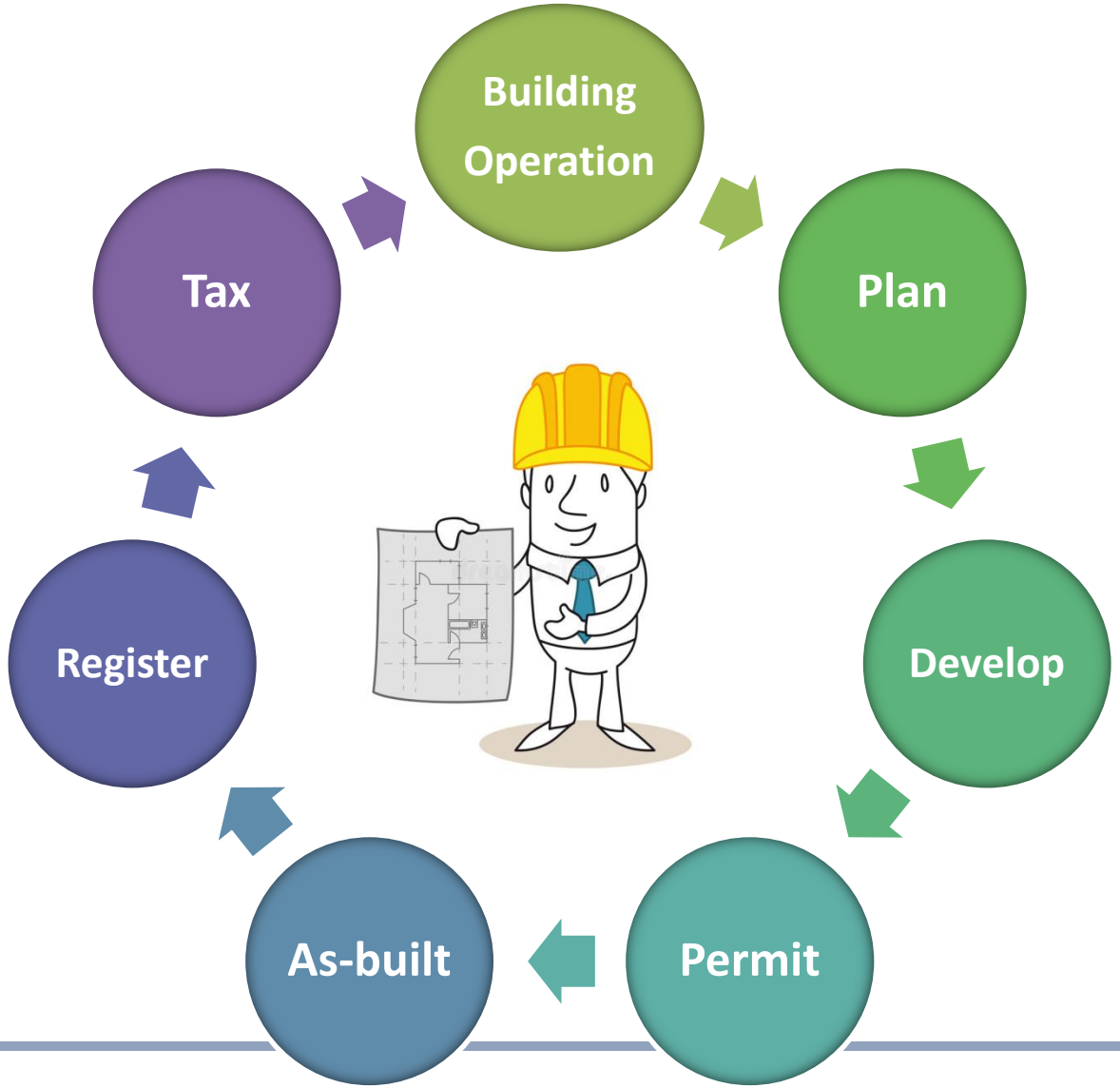
Crowdsourcing of defibrillator locations

Contents

- Background on internet Eras
- What is a data lossless nation?
  - Data products in the construction process
  - The Inter-agency GIS committee
  - National mapping standard
- On-line QC tool
- The data integration tool
- Conclusions

What is data lossless culture?

Every piece of data, creating by any professional is QC-ed, integrated and **stored in the national database**.

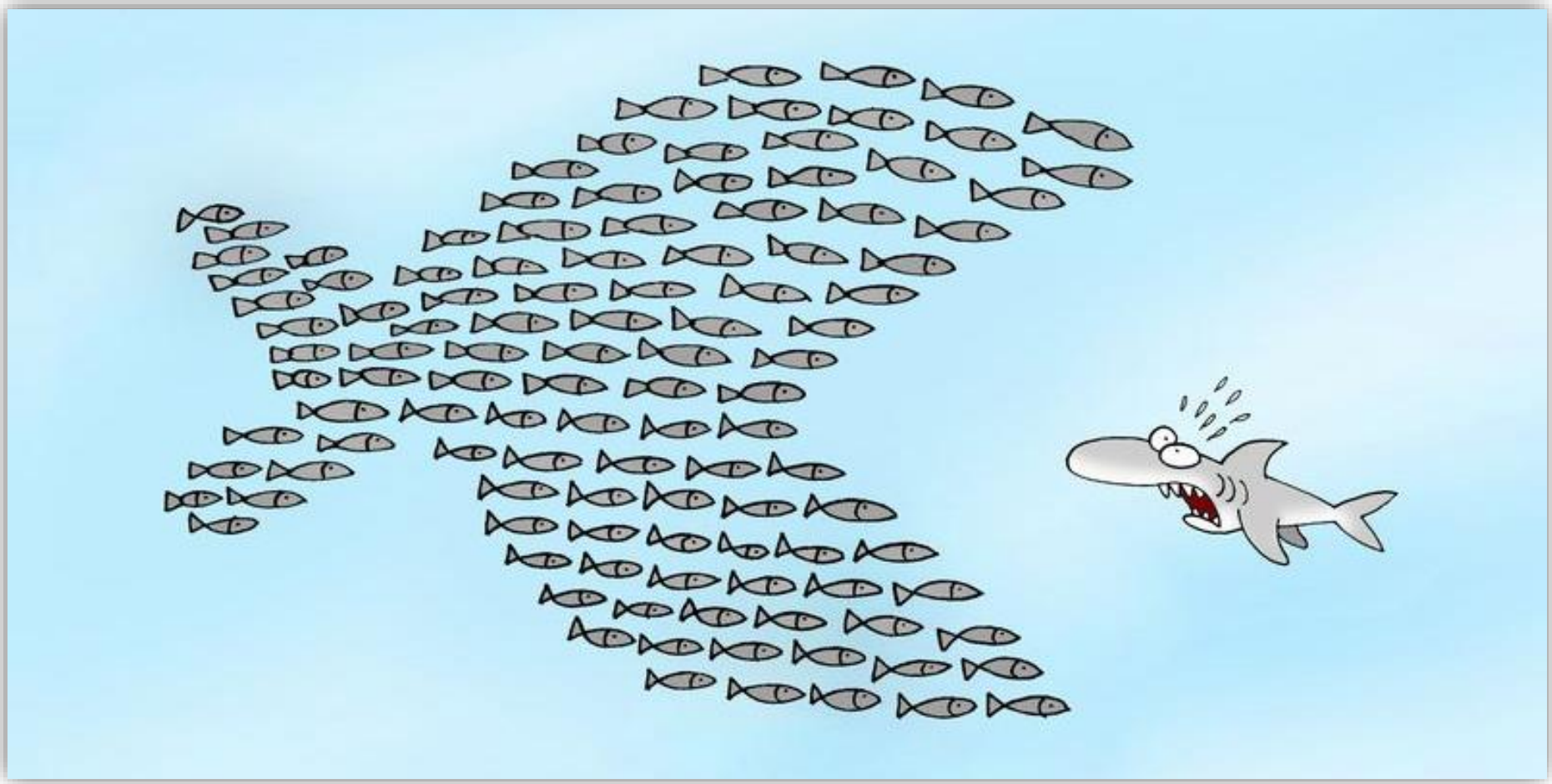


SURVEY OF ISRAEL

How to create a data lossless culture?

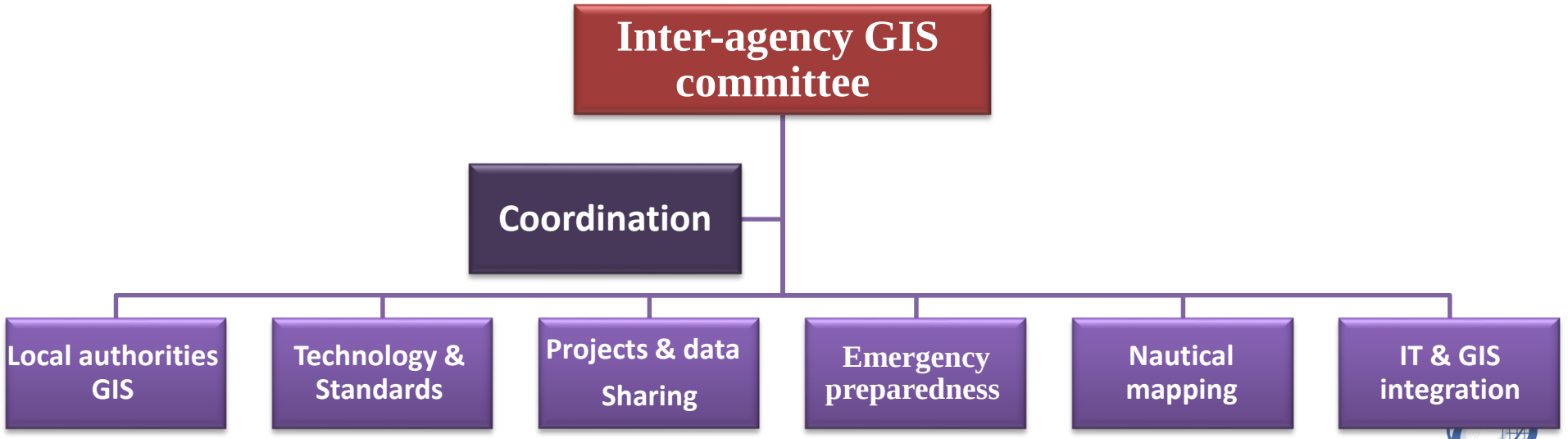
1. How to identify all the data generating parties?
2. How to convince organizations and people to share data? Despite the many barriers (legal, security, privacy, copyrights, tradition)
3. How to deal with different data format, structures and types which makes it hard to reuse data? Large amount of information unstructured.
4. How can we perform quality control on so many datasets?
5. How can we integrate all the data into a uniform database?

The power of collaboration



How to identify all the data generating parties?

- The inter-agency GIS committee was established in 1997 by a prime-minister order. It is Chaired by the SOI Director General.
- The committee has delegates from 45 governmental offices and 256 local authorities. This year the committee teams had 22 meetings with 176 participants.
- The key objective of the committee was to promote efficient use of GIS in the national level through data and knowledge sharing, and standards and working procedure development.

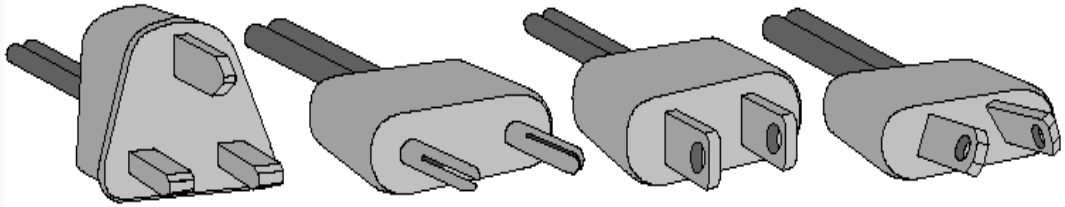


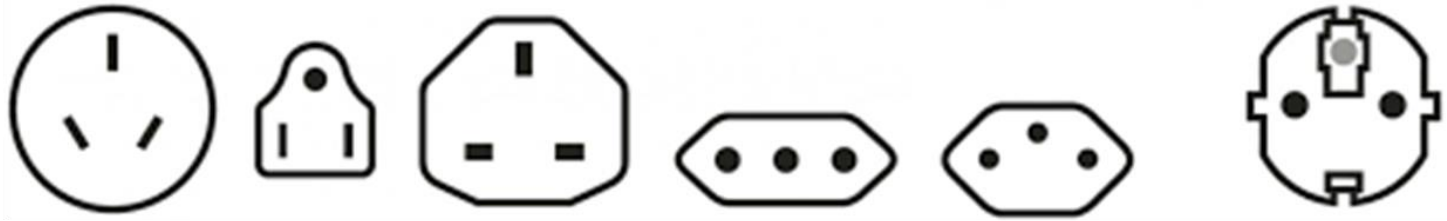
How to create a data lossless culture?

1. How to identify all the data generating parties?
2. How to convince organizations and people to share data? Despite the many barriers (legal, security, privacy, copyrights, tradition)
3. **How to deal with different data format, structures and types which makes it hard to reuse data? Large amount of information unstructured.**
4. How can we perform quality control on so many datasets?
5. How can we integrate all the data into a uniform database?

Why standards?

Standards that link hardware and software



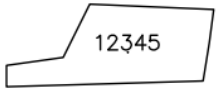
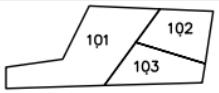
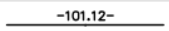


Standards as a common language

SURVEY OF ISRAEL

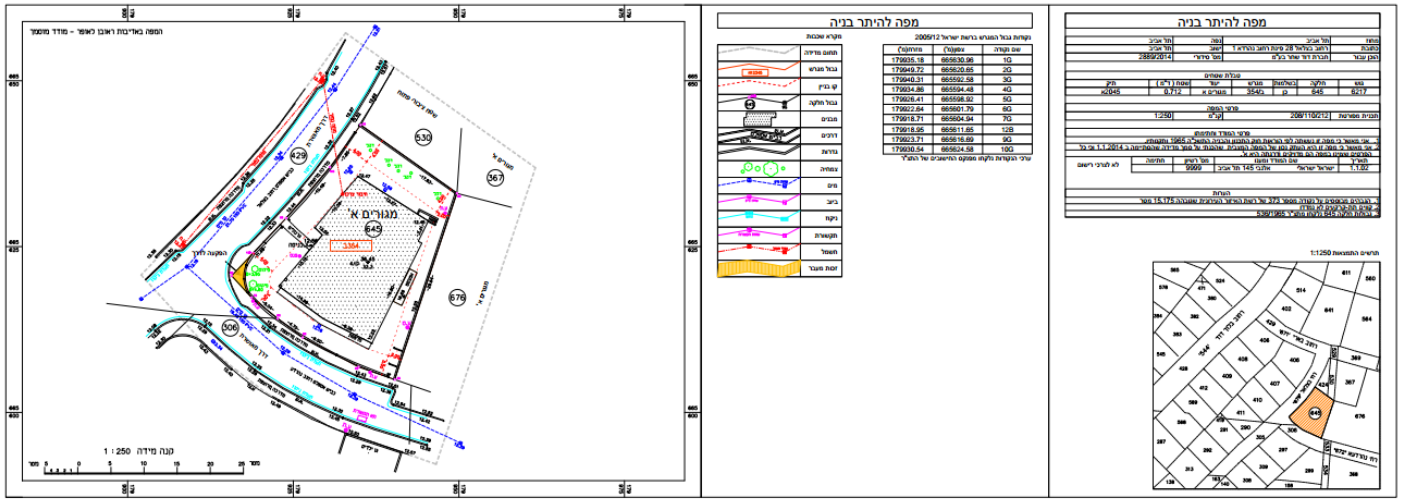
National mapping standard (published 2016)

Define data format and layers

C1600	פוליגון גוש	
C1601	נתוני גוש	
C1602	פוליגון חלקה	
C1603	נתוני חלקה	
C1609	מידת חזית	
C1610	נקודת גבול קיימת	
C1611	נקודת גבול חדשה	
C1642	מספר אליפסה	

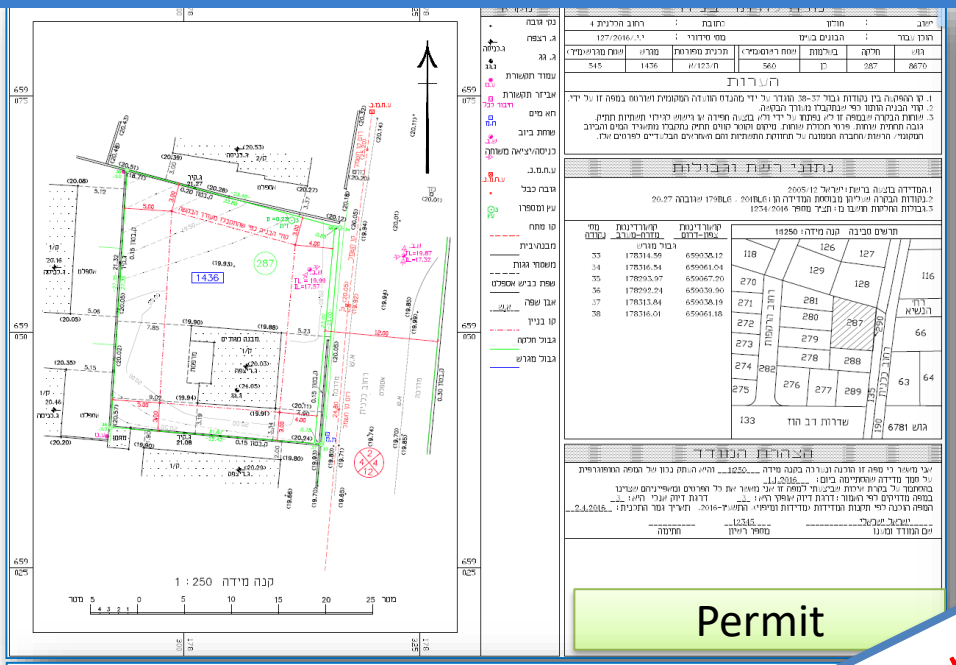
M4800	תווית תיאור	
M4801	קו ביוב תת קרקעי	
M4802	קו סניקה	
M4804	שוחת ביוב	
M4805	משאבה	
M4807	שסתום אויר/ נשם	
M4819	כיוון זרימה	
M4820	אגן חמצון	
M4850	אביזר ביוב כללי	
M4851	נתוני גובה במהלך קו ביוב	

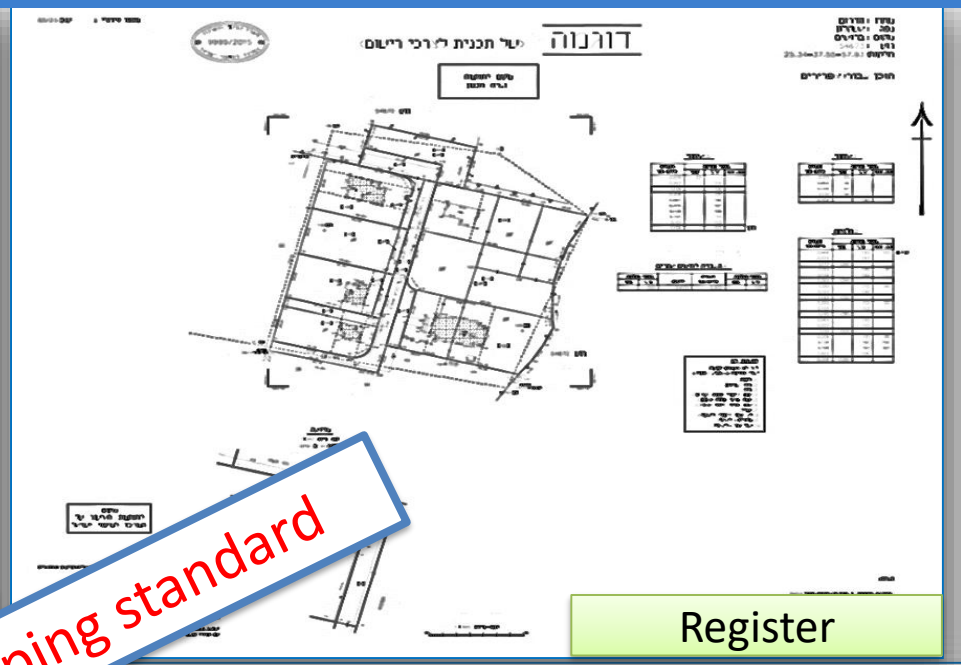
Define the content and the cartography



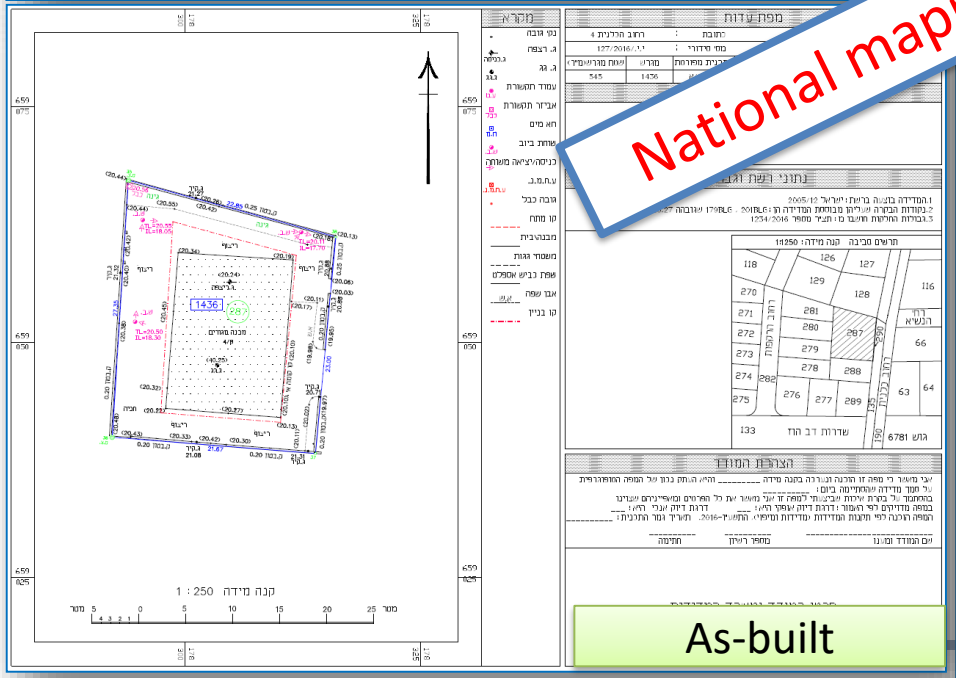
SURVEY OF ISRAEL

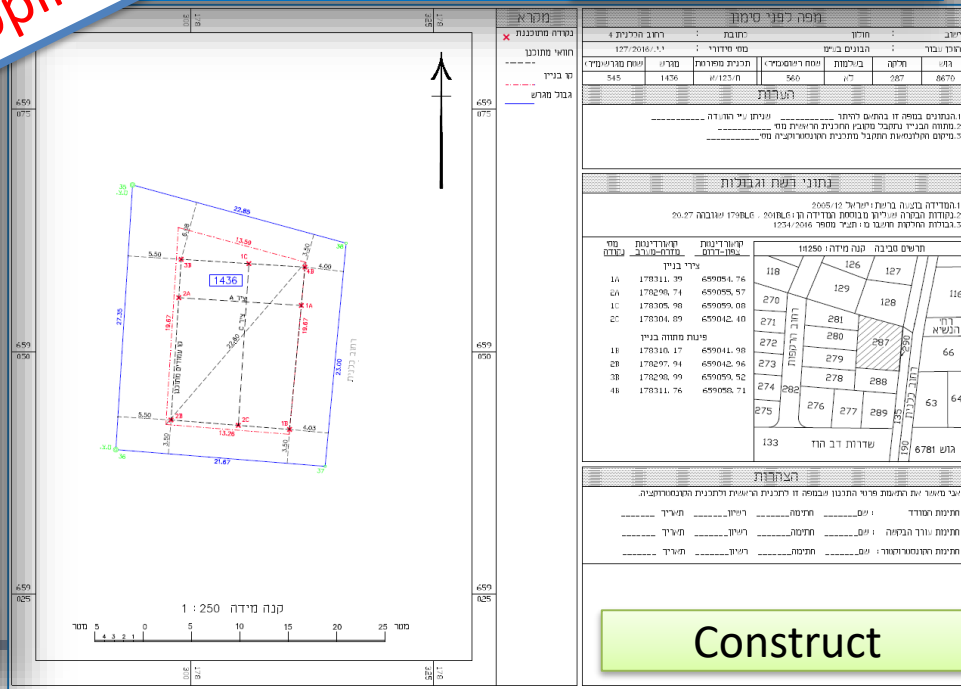
Uniform mapping





National mapping standard

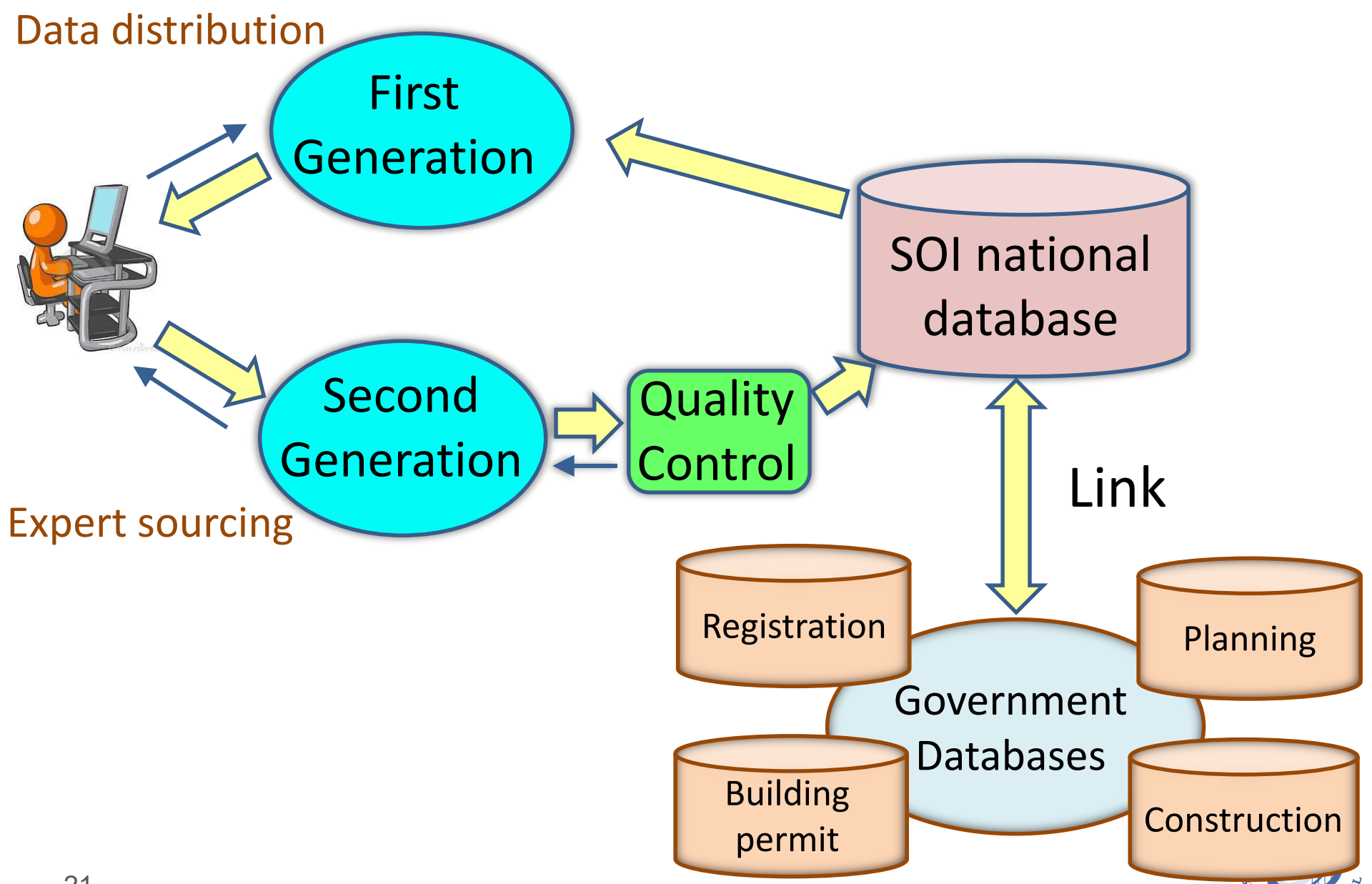




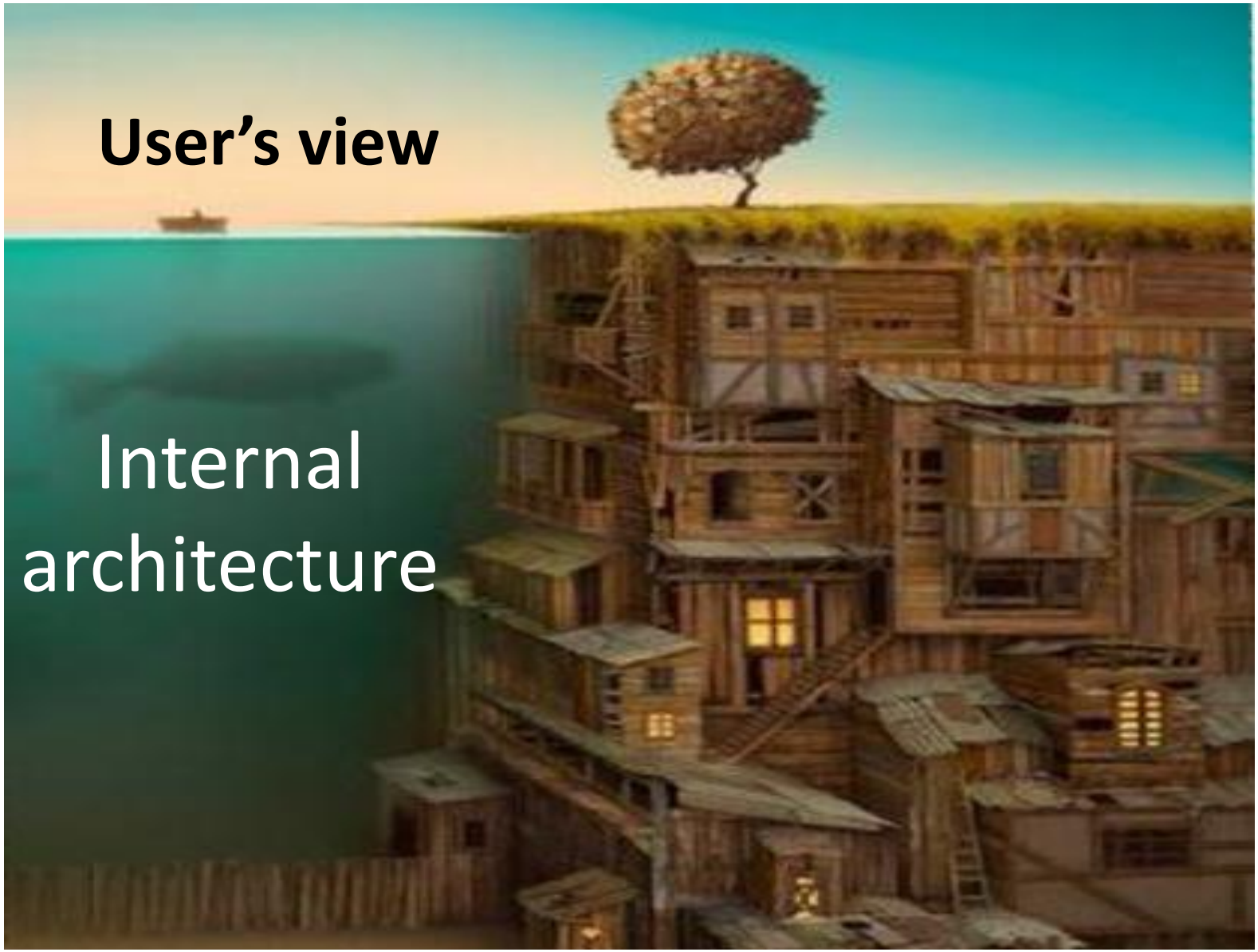
How to create a data lossless culture?

1. How to identify all the data generating parties?
2. How to convince organizations and people to share data? Despite the many barriers (legal, security, privacy, copyrights, tradition)
3. How to deal with different data format, structures and types which makes it hard to reuse data? Large amount of information unstructured.
4. **How can we perform quality control on so many datasets?**
5. How can we integrate all the data into a uniform database?

Topographical and Cadastral system development



Simplicity – friendly user experience

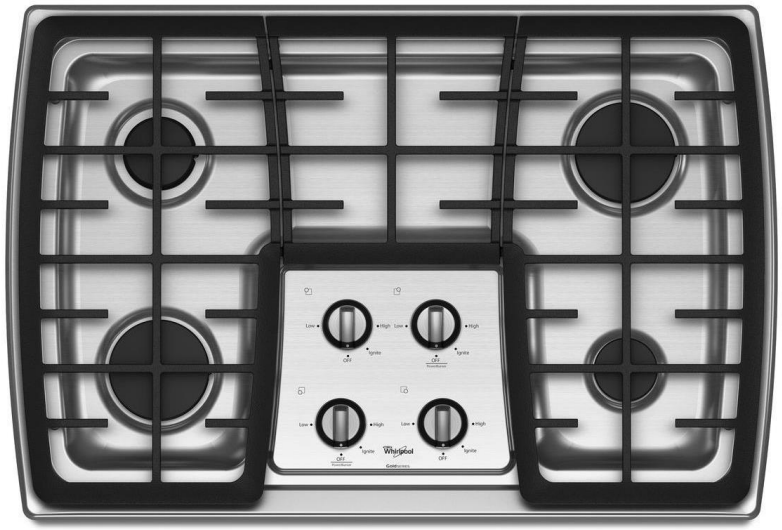


SURVEY OF ISRAEL

Simplicity – friendly user experience

Old design

Gas stove: 4 burners and 5 buttons



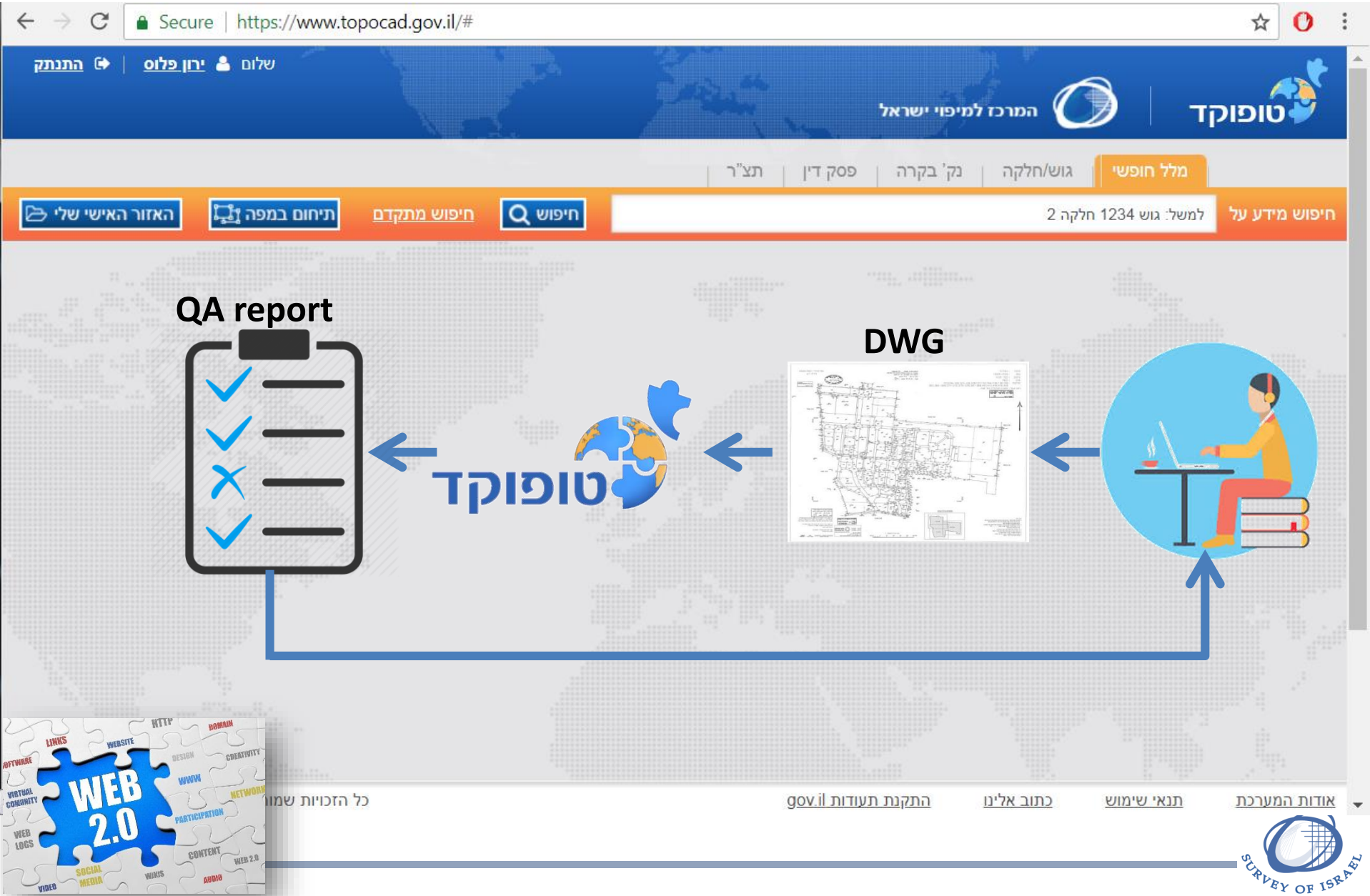
Modern design

No need for user manual or training

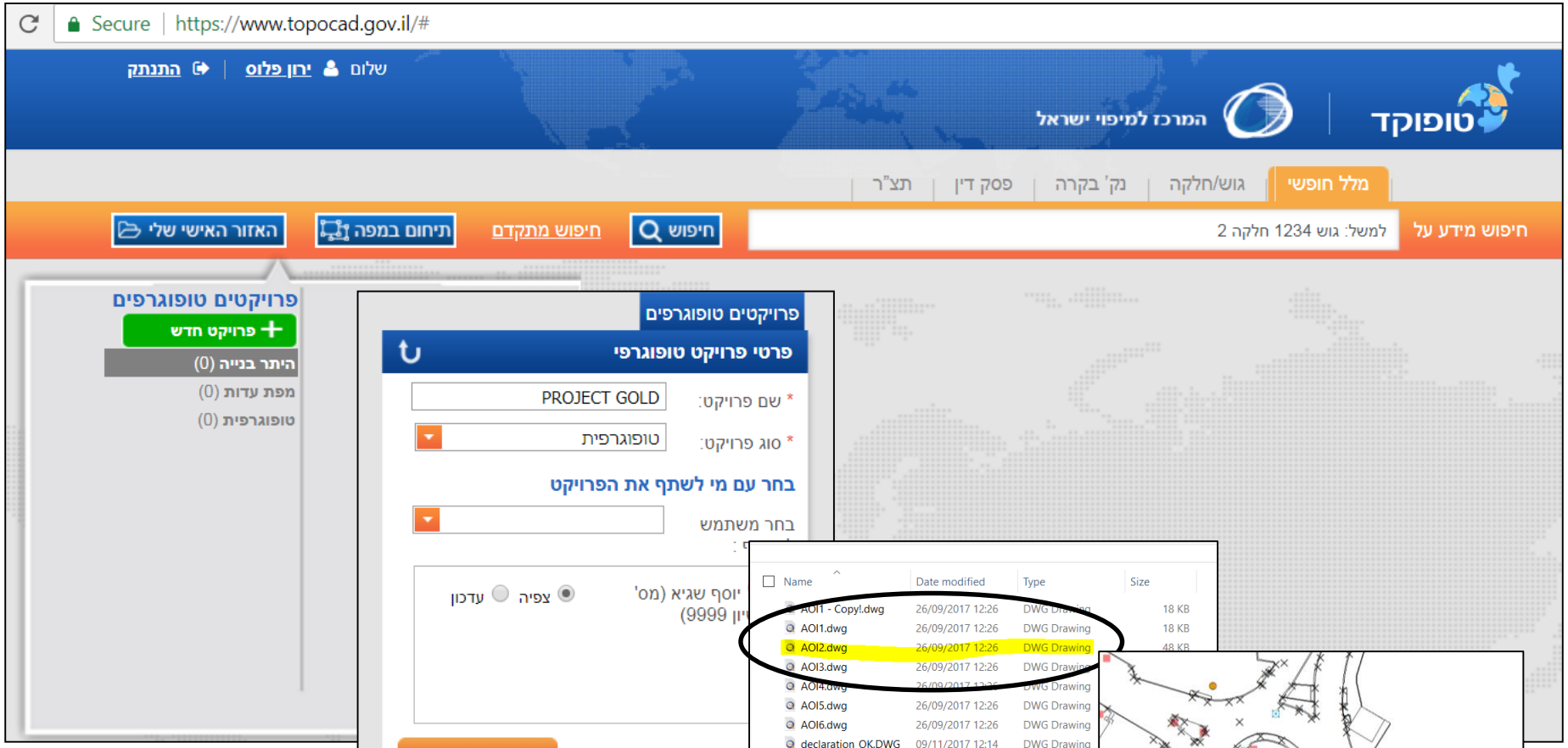
Rules:

1. No need for user manual or training
2. Smallest number of clicks (more is less)
3. Simple and fun to work with

The TopoCad System



Online QC tool



1. Open a project

2. Define project

3. Load data

4. View QC results

Contents

- Background on internet Eras
- What is a data lossless nation?
  - Data products in the construction process
  - The Inter-agency GIS committee
  - National mapping standard
- On-line QC tool
- The data integration tool
- Conclusions

Quality Control Process (ISO 19157)

Spatial data file

Standard compatible — Readable format

Logical consistency — Meet rules

Completeness — No omissions or commissions

Accuracy assessment — Positional accuracy

Thematic accuracy

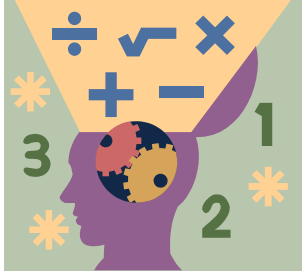
Temporal quality

Incompatible

Compliance (metadata report)

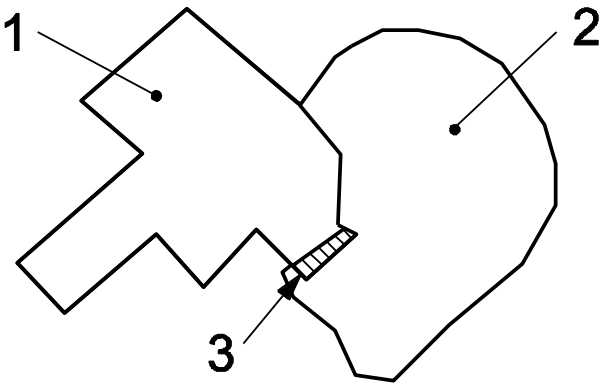
SURVEY OF ISRAEL

Logical consistency



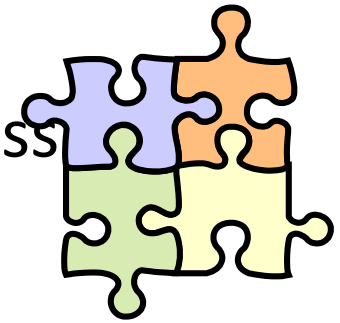
The program finds elements that do not conform to (logical) rules:

- Topology issues (Closed buildings – overshoots & undershoots)
- Topographic elements that appear in an incorrect layer (a road in a layer of buildings)
- Contour lines crossing each other
- A fence corner that appears inside a road
- Parcels overlapping each other (forming parcels smaller than 3 square meters)
- Duplicated lines



Completeness

- Completeness includes missing information (omissions) or information that appear in the file and not in reality (additions)
- The content of the map is defined in the specifications and match the map scale.
- Completeness test are performed interactively by overlaying the data file on top of a orthophotography at the proper scale.
- An error matrix is used to report of completeness



Commission

Omission

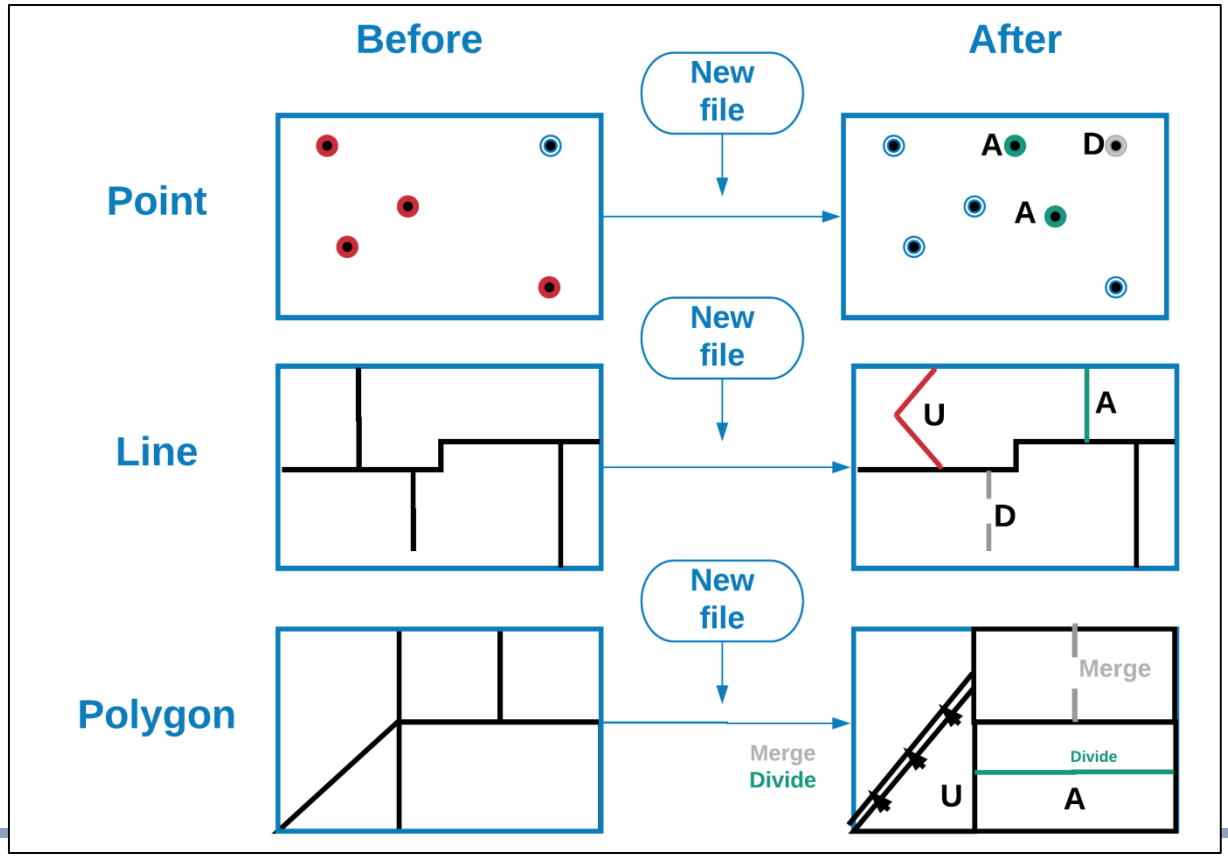
Thematic and positional accuracy

- Tests for thematic and positional accuracy are performed by field checking (of a sample) of objects in the map and comparing the field with the data on the map
- What is the attribute value (type of tree) in the map and what is the correct value in the field.
- Root mean squared error is used to evaluate the positional accuracy. It is calculated from the differences between a value in the map and an accurately measured position in the field.

$$RMSEx = \sqrt{\frac{\sum (dx_i)^2}{n_x}}$$

The data integration process (conflation)

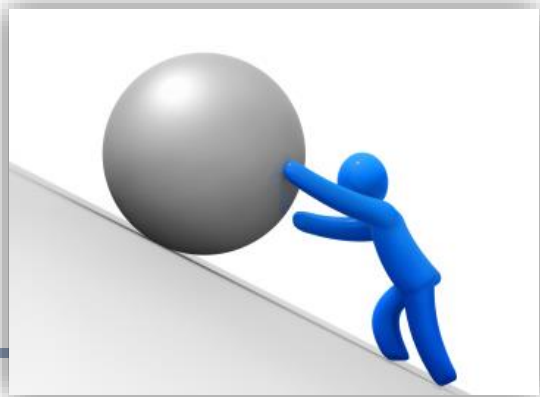
- Data integration is a complex process which is currently done semi-automatically (interactively).
- The computer identifies features to Add, Delete and Updata
- A human operator has to approve the change.



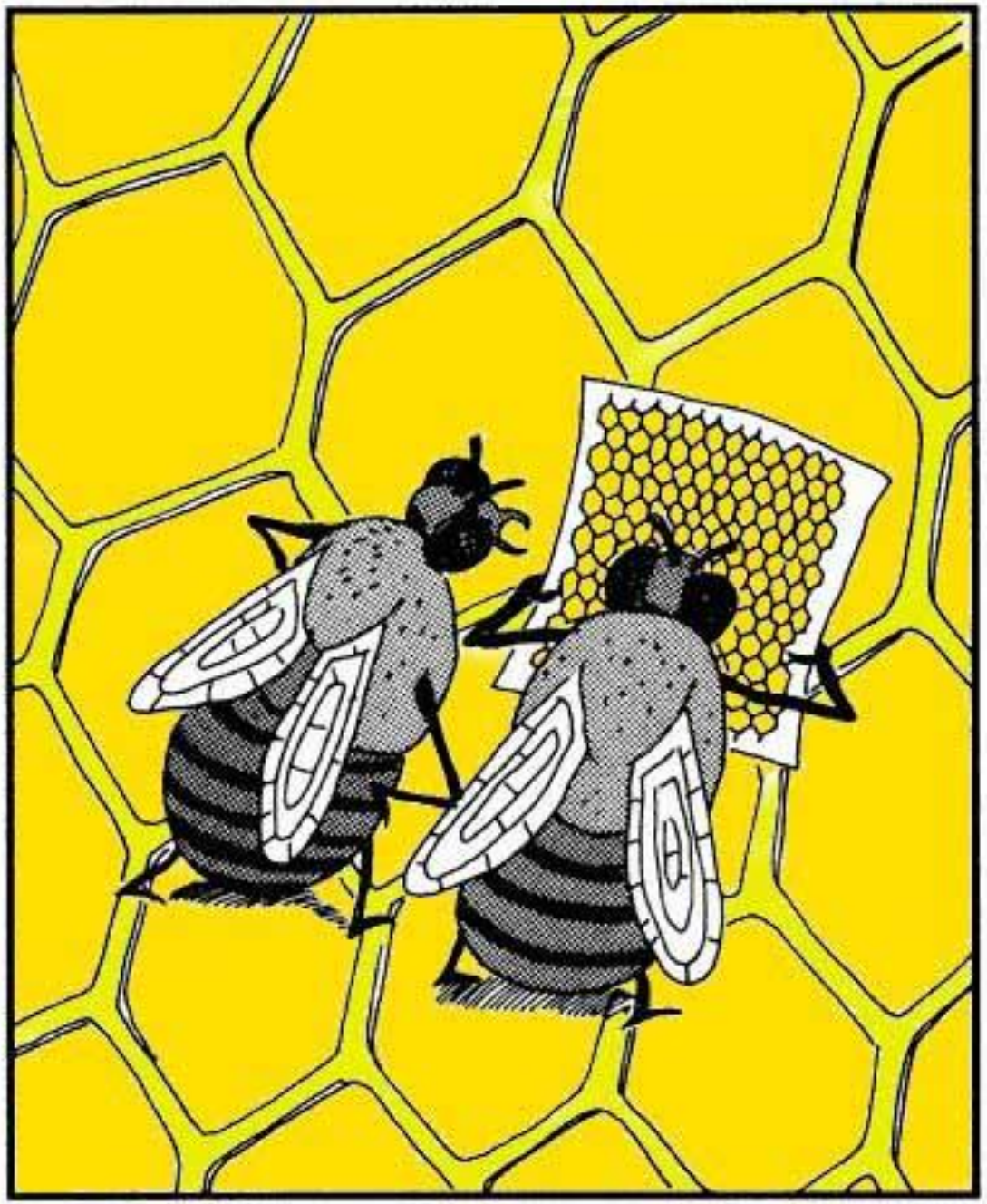
SURVEY OF ISRAEL

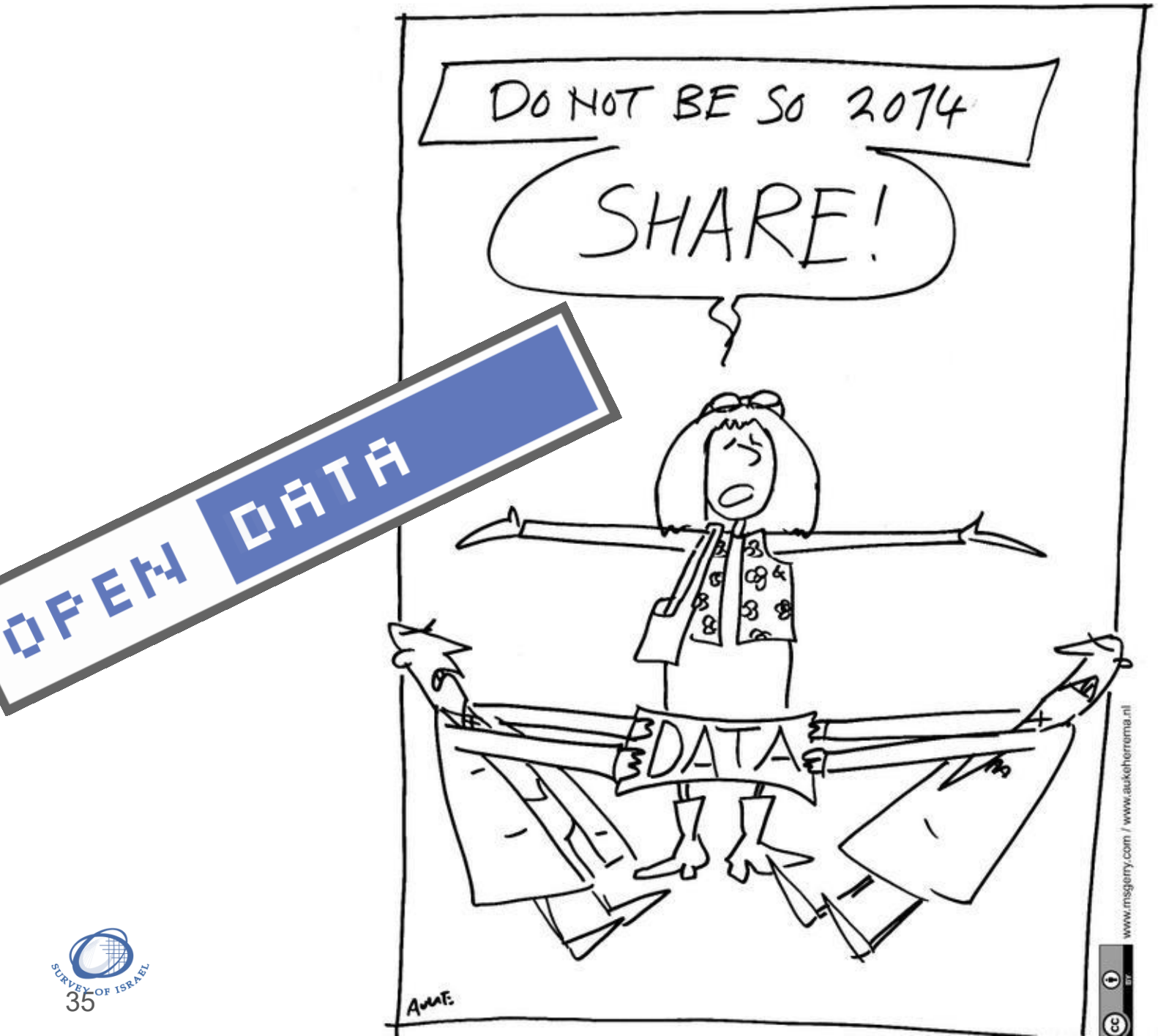
Summary

- Data is generated every steps of governmental and engineering activities
- Expert sourcing is the process of acquiring these data and making a useful database from it.
- We presented a working process.
- Some challenges that will be solved are:
  - A complete automatic process for the quality control and the conflation
  - Comprehensive treatment of the intellectual property rights

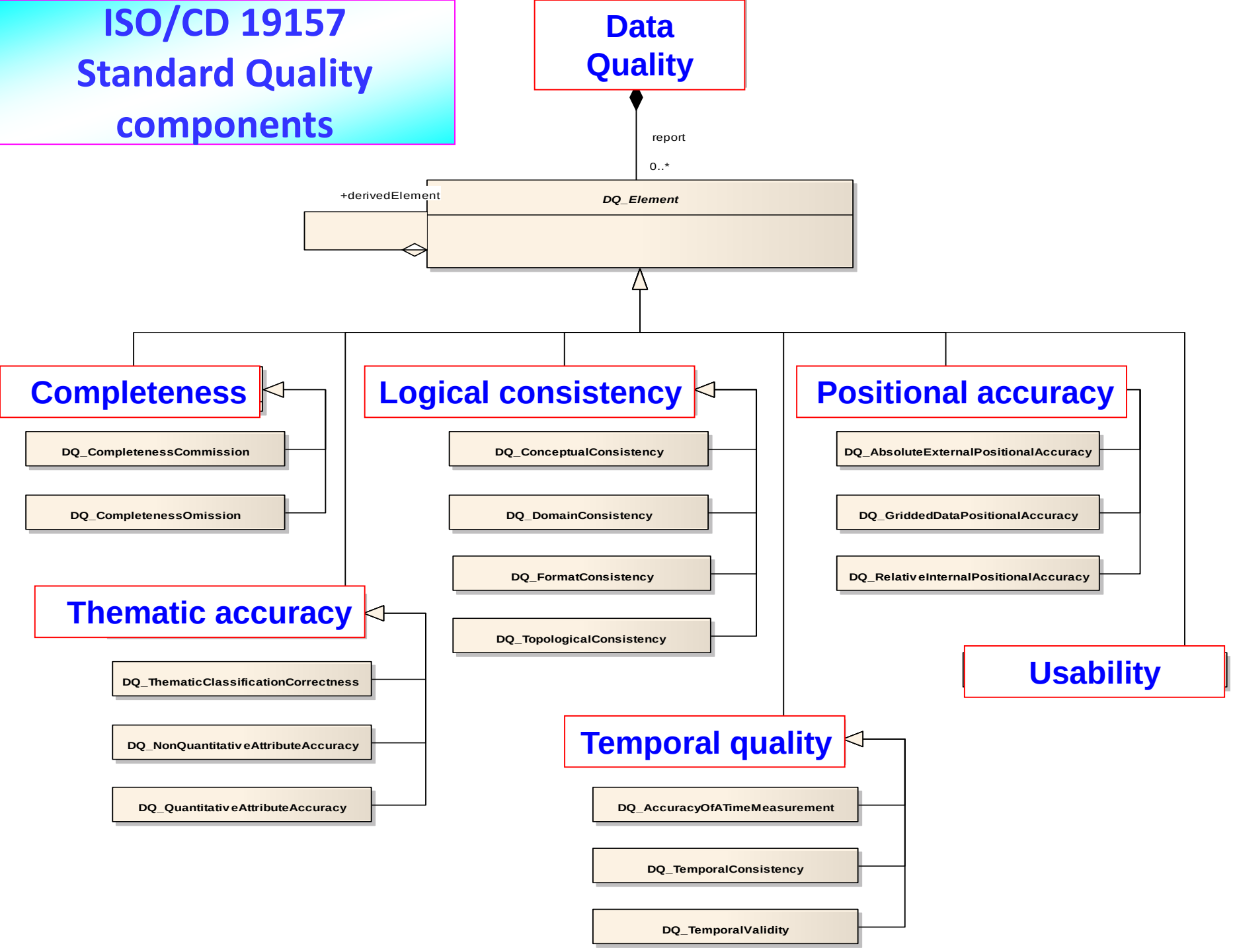


Thanks you



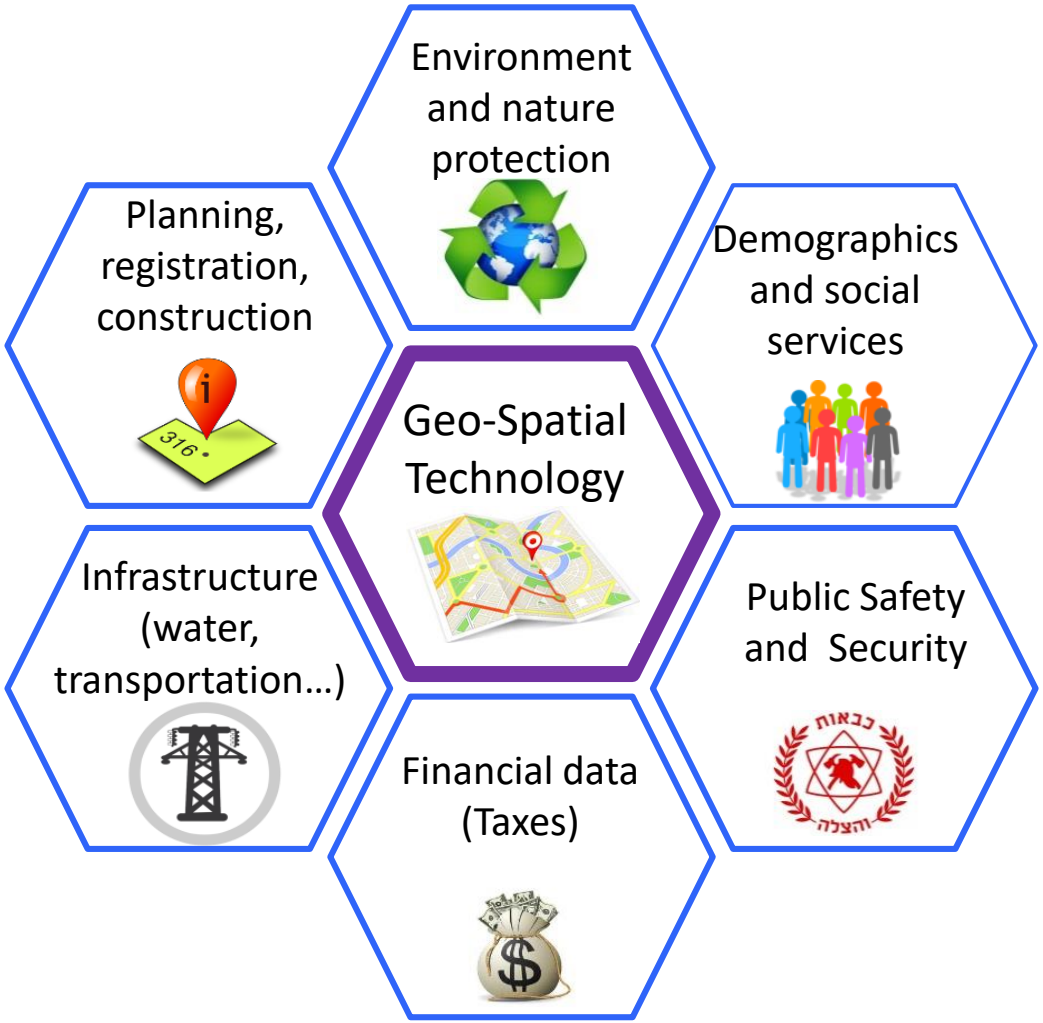

DO NOT BE SO 2074
SHARE!
OPEN DATA
DATA
www.msgerry.com / www.aukeherrema.nl

ISO/CD 19157 Standard Quality components



Why Geo-spatial technologies?

Geo-spatial technology underpins smart nation operations



SURVEY OF ISRAEL