



BUILDING NATIONAL DATABASES USING EXPERT SOURCING

Prof. Yaron Felus
Chief Scientist

20 November, 2018

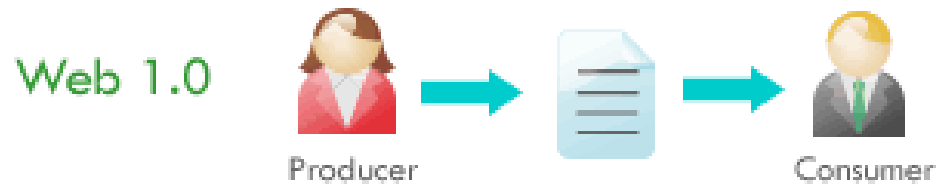
Contents

- **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**



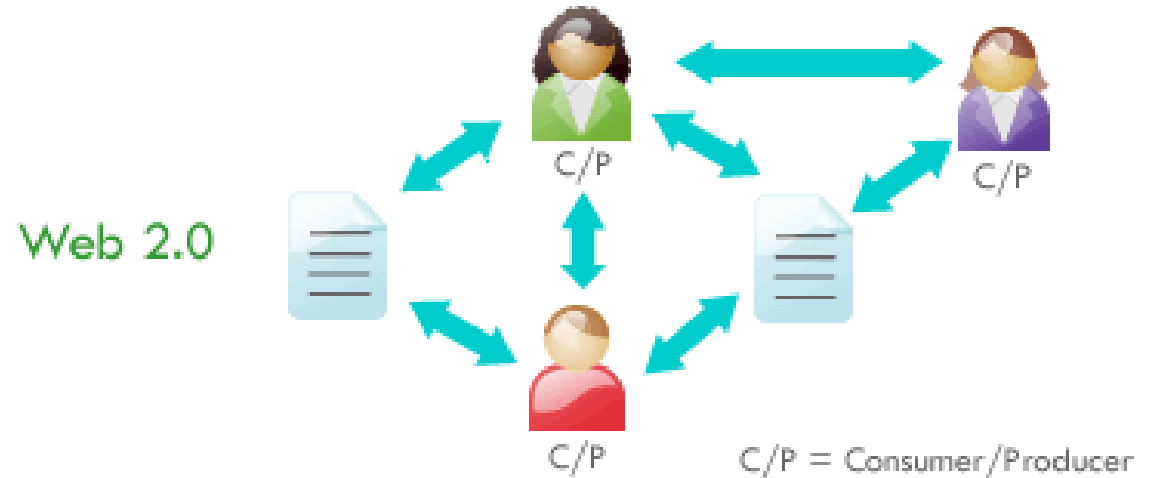
Internet eras

One organization providing information (Read)

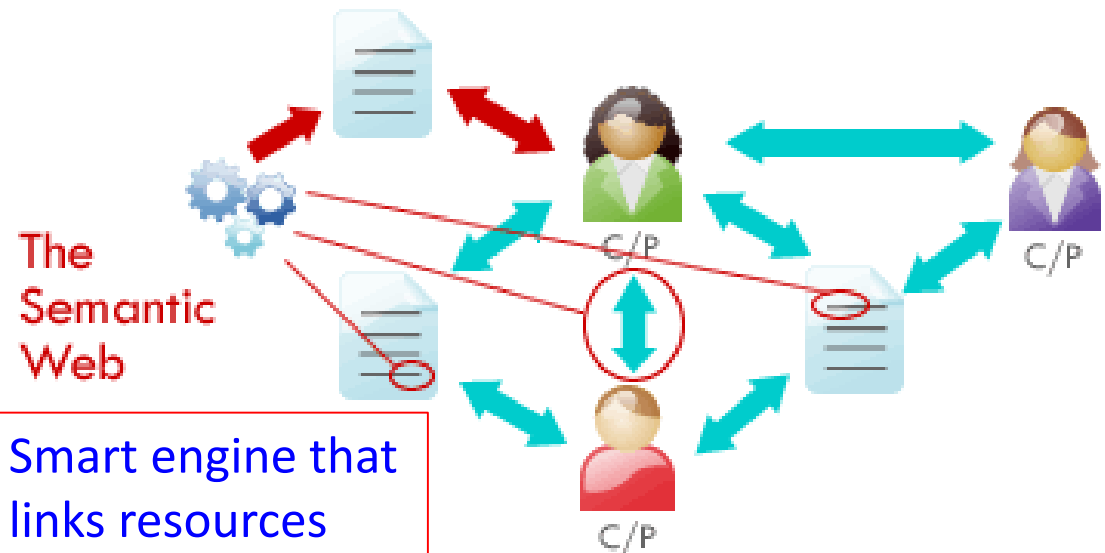


Collaboration with the public (Read and write)

Wiki, facebook, youtube, waze

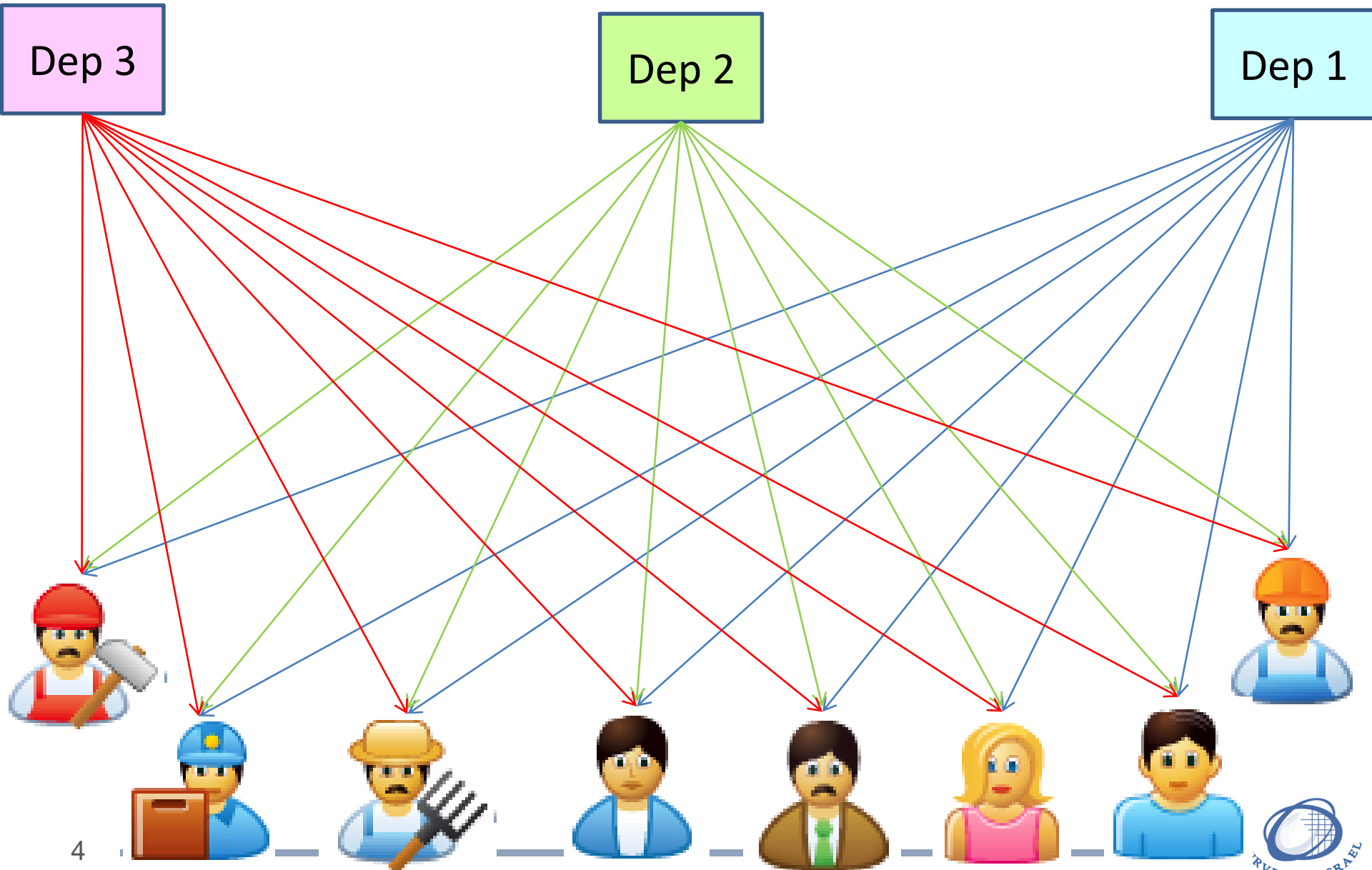


Web 3.0: Smart and connected system (Read, write and understand)

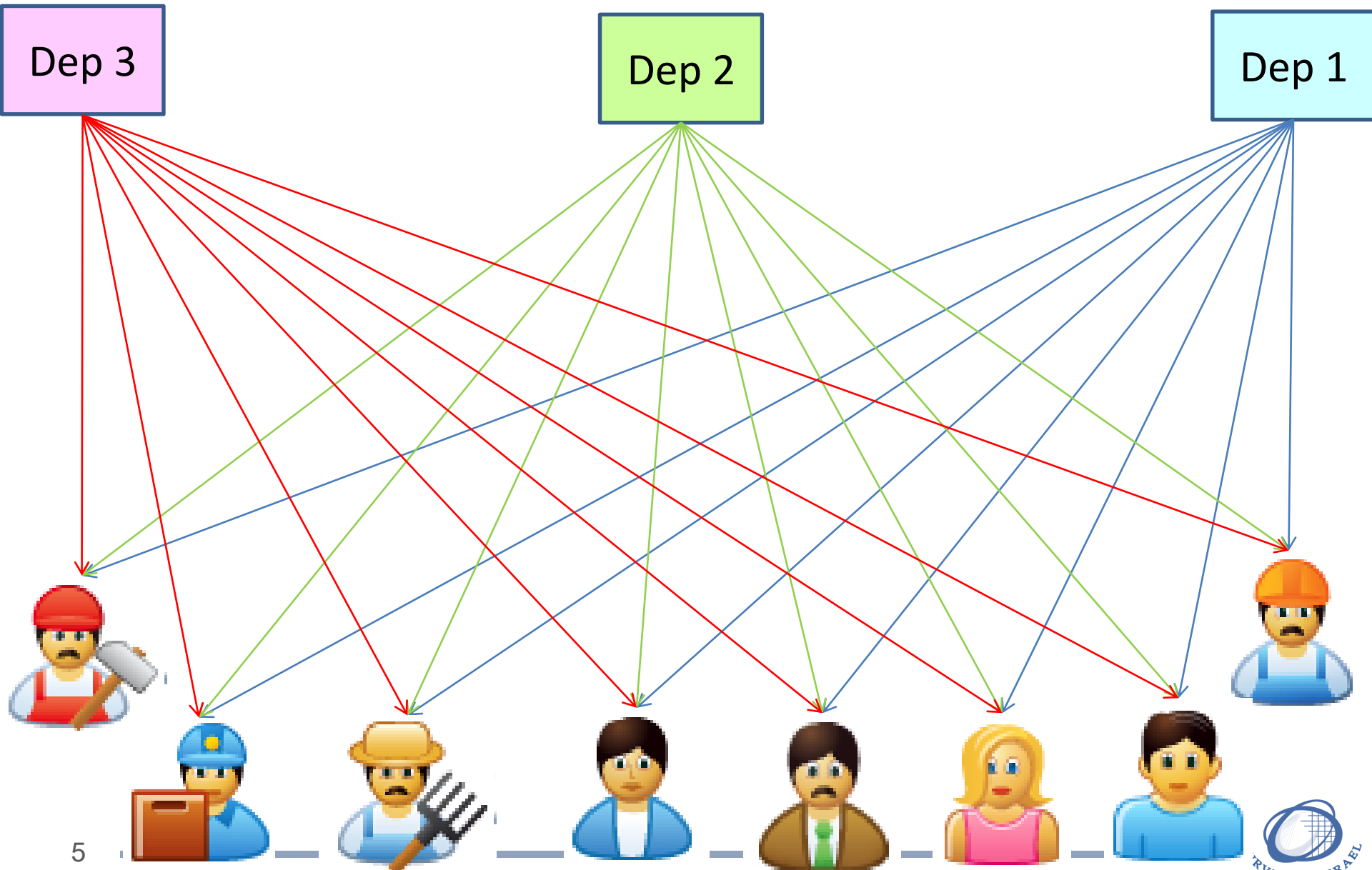


The Beginning

Organization Management

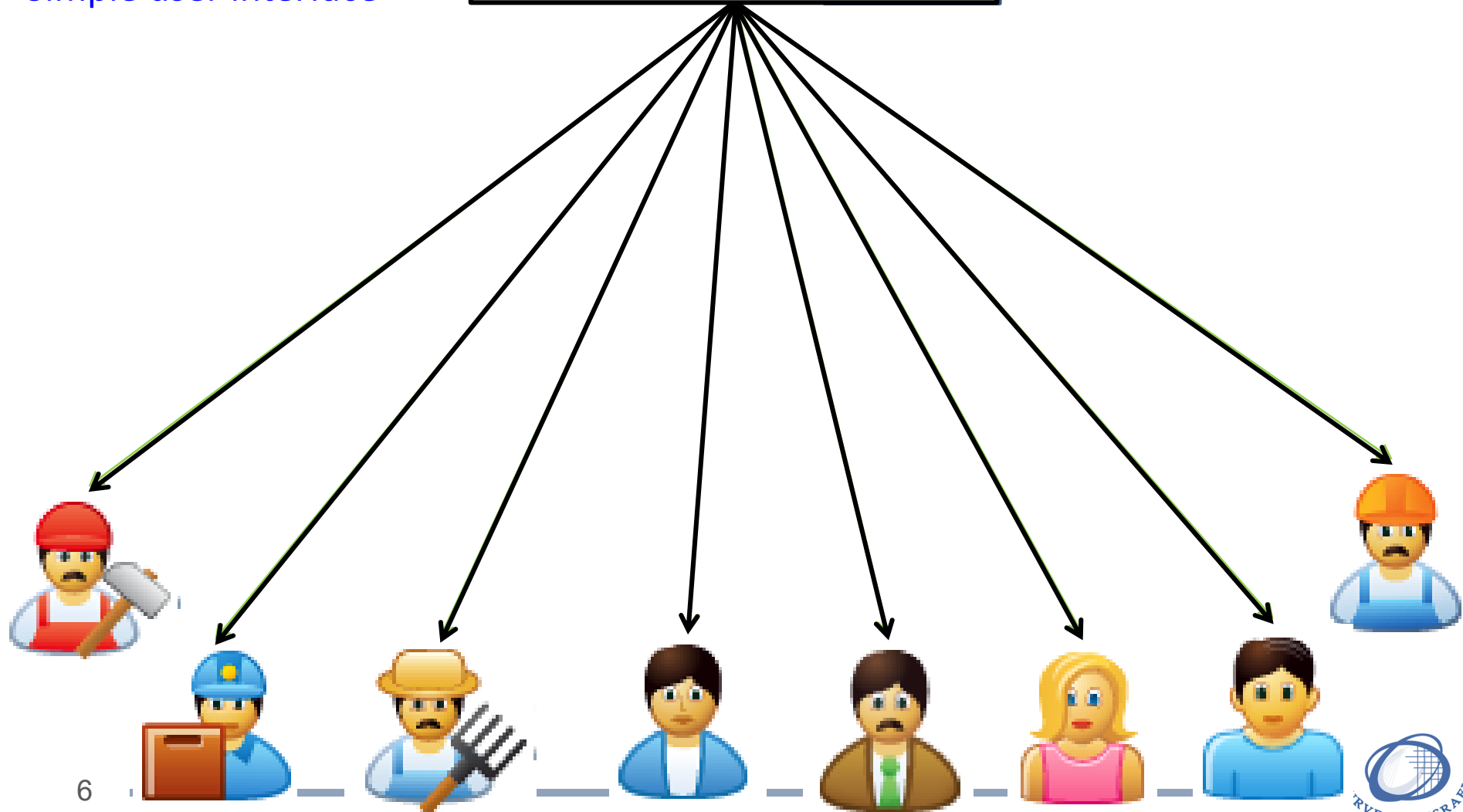
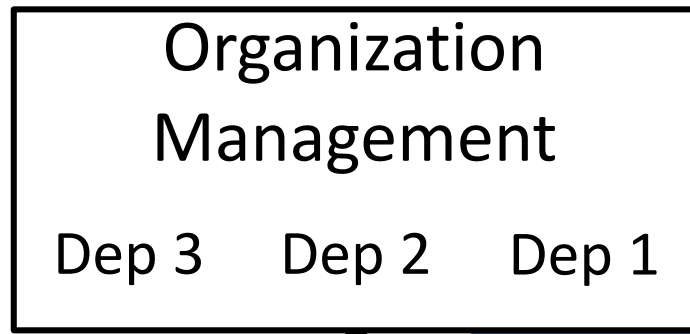


Organization Management



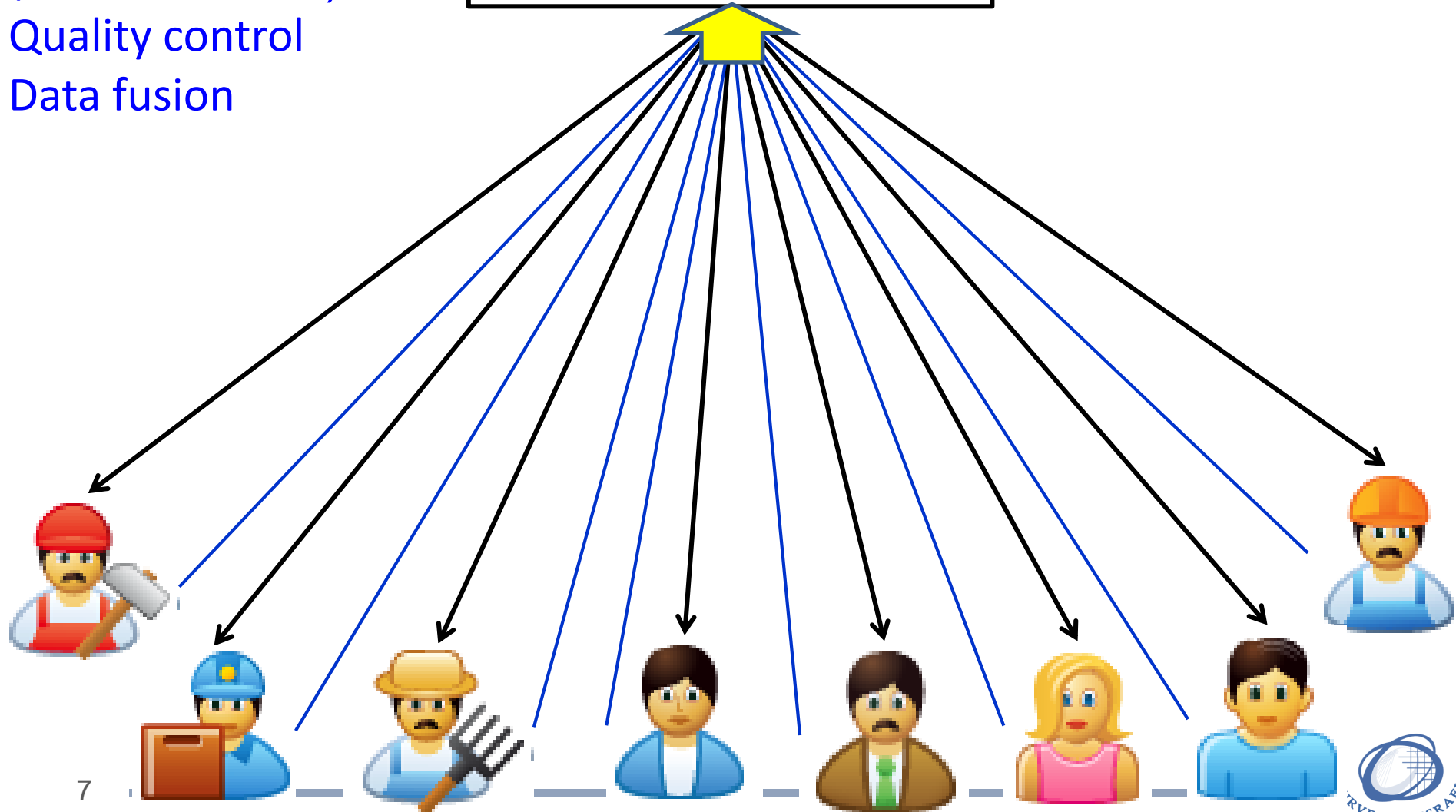
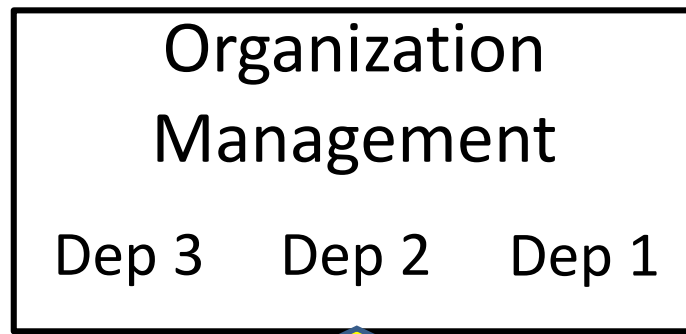
Web 1.0

- One data source
- Connectivity standards
- Uniformity
- Simple user interface

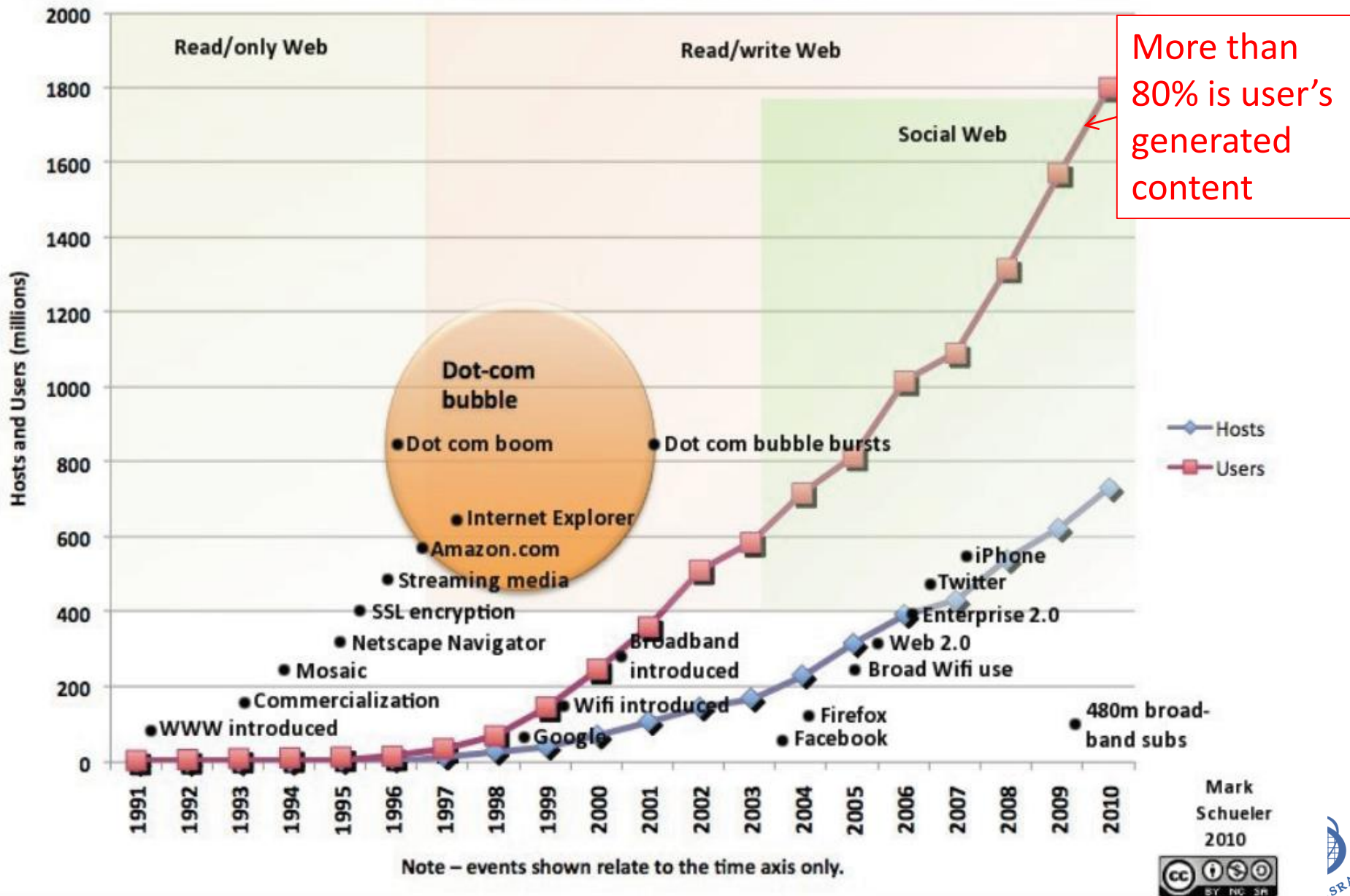


Web 2.0

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



The world is ready for that...

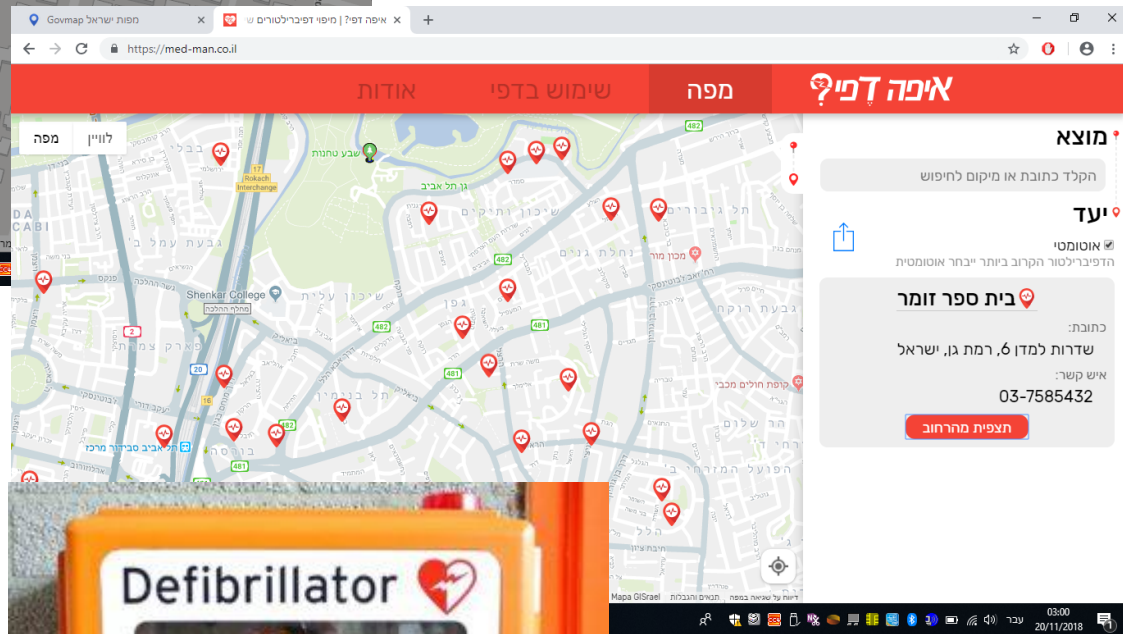
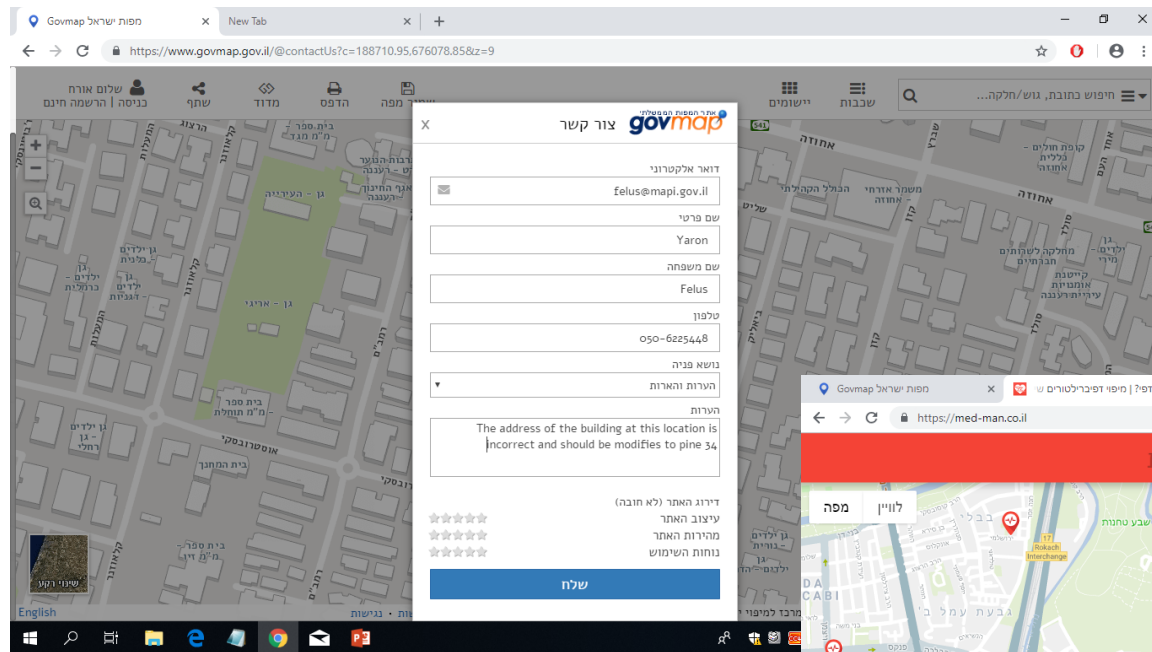


Crowdsourcing Vs wisdom of the crowd Vs Expert Sourcing

- **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



Crowdsourcing applications



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



Crowdsourcing of defibrillator locations

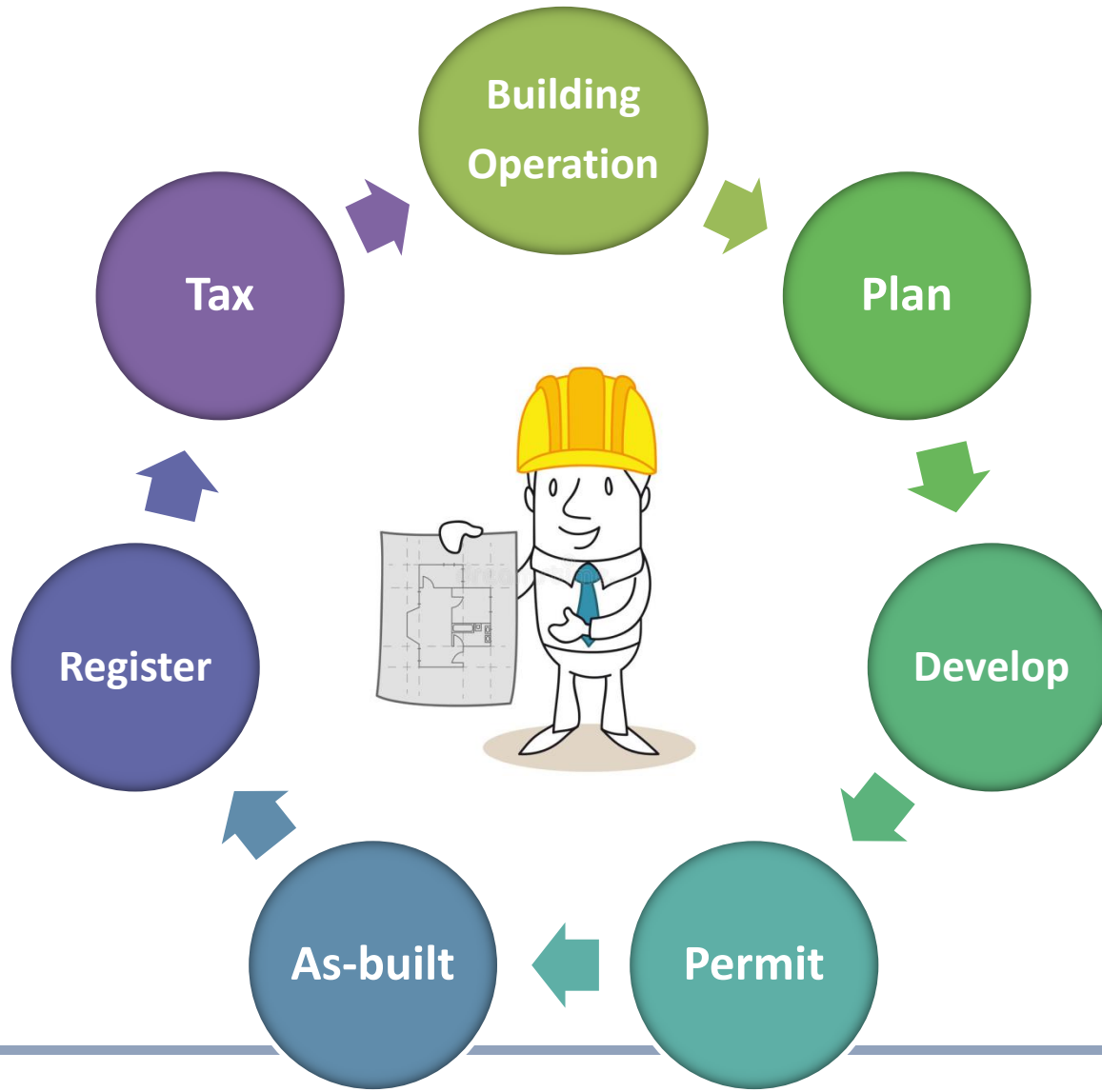
Contents

- **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.

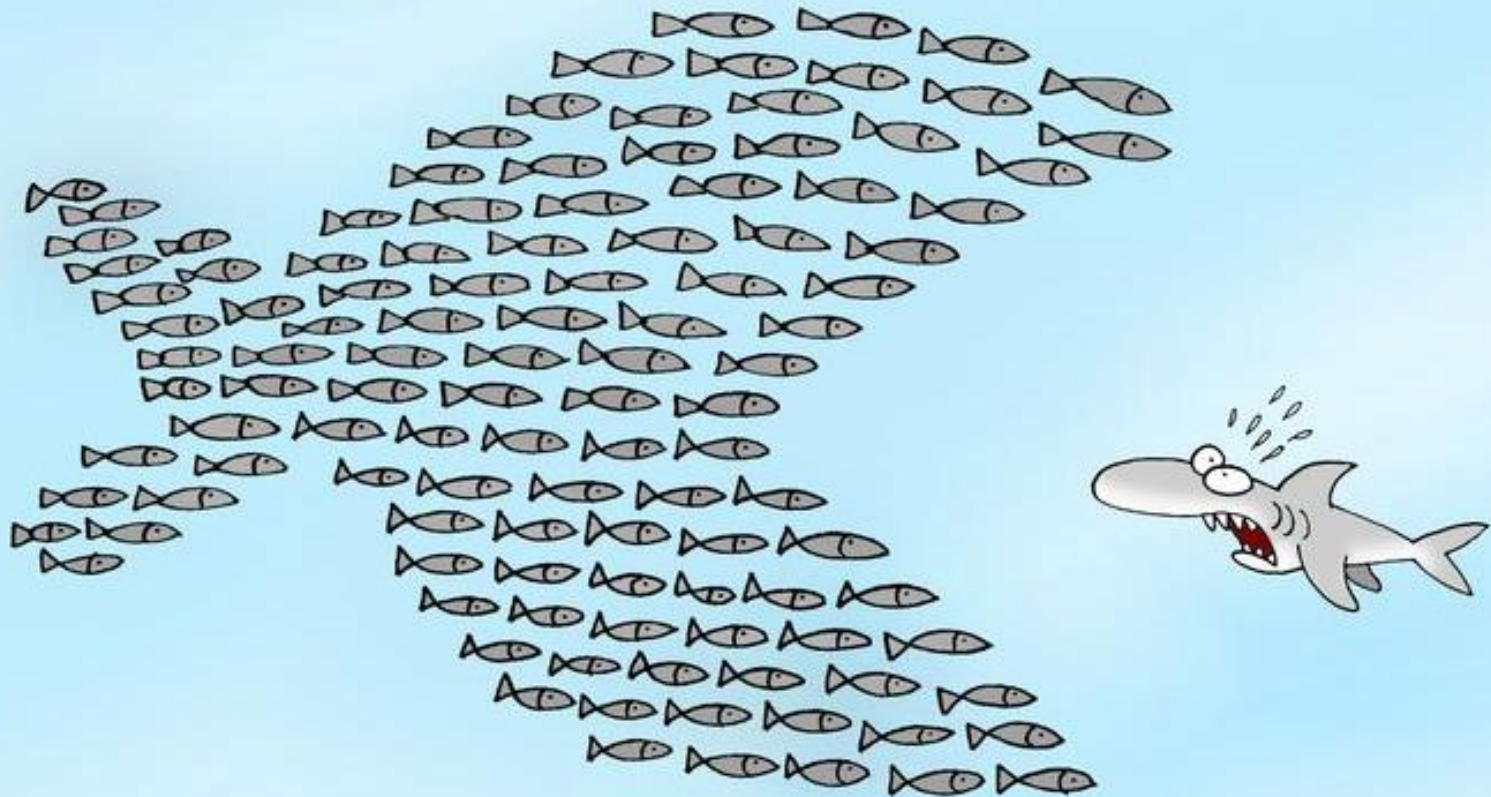


How to create a data lossless culture?

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?

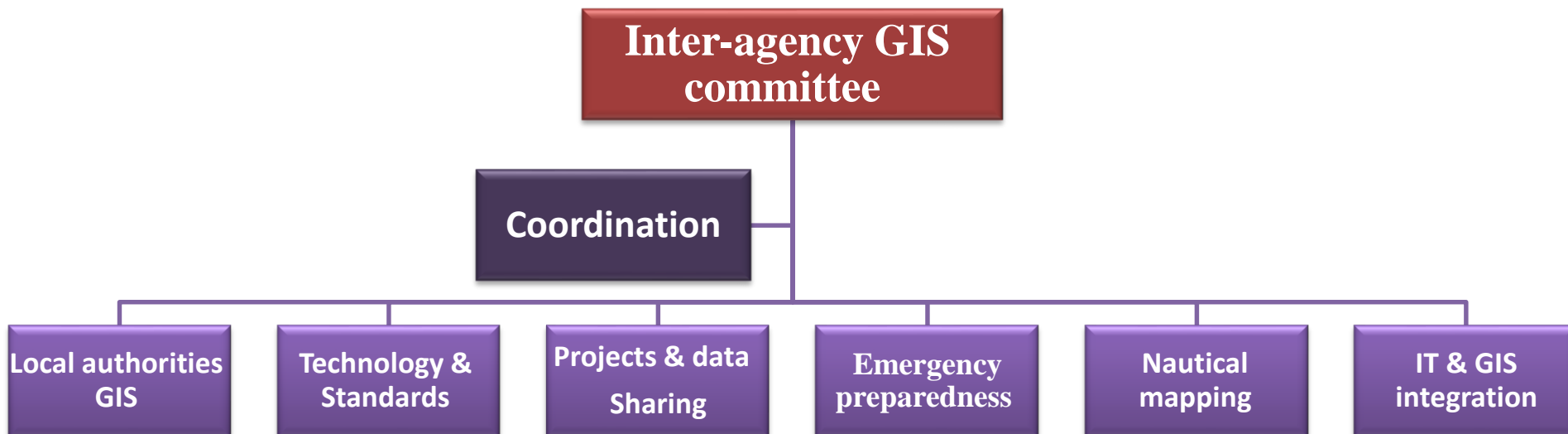


The power of collaboration



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.



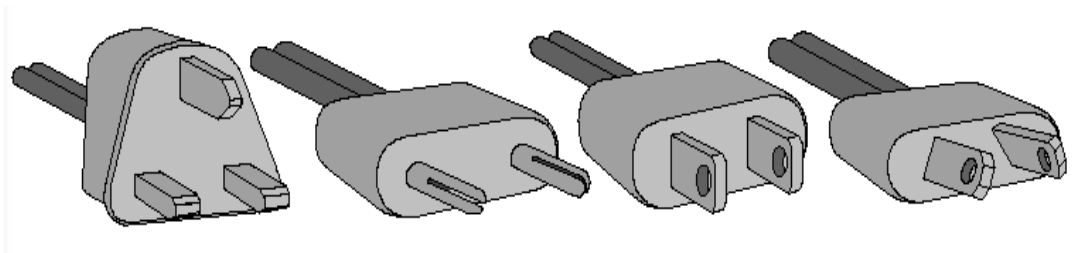
How to create a data lossless culture?

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



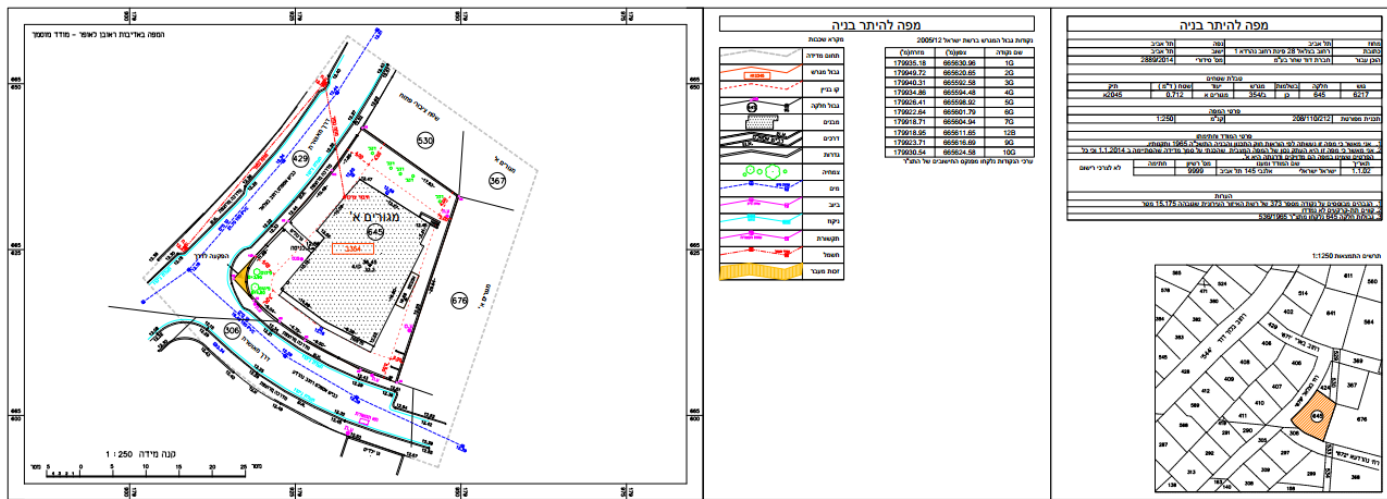
National mapping standard (published 2016)

Define data format and layers

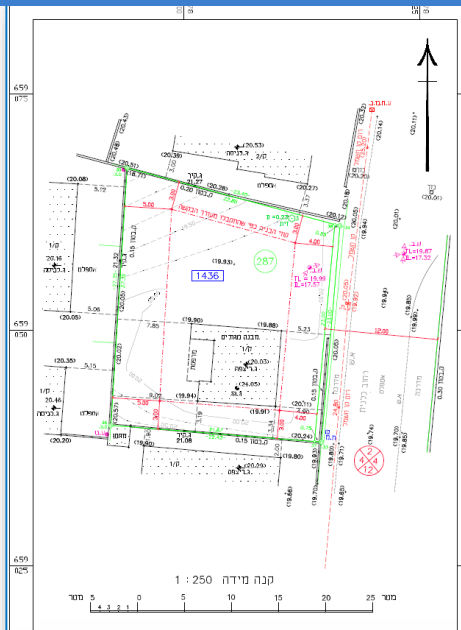
	פוליגון גוש	C1600
	נתוני גוש	C1601
	פוליגון חלקה	C1602
	נתוני חלקה	C1603
	מידת חוית	C1609
	נקודת גבול קיימת	C1610
	נקודת גבול חדשה	C1611
	מספר אליפסה	C1642

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

Define the content and the cartography



Uniform mapping



תוכנית

החוק: 1277.2016.4
 החוק: 1277.2016.4
 תחום: 1436
 חלק: 545

הערות:

1. תוכנית זו נכנסה לתוקף ב-27.12.2016.
2. תוכנית זו נכנסה לתוקף ב-27.12.2016.
3. תוכנית זו נכנסה לתוקף ב-27.12.2016.

נתונים כלכליים:

מס	שטח (מ"ר)	ערך (₪)
118	17814.59	659051.76
126	17816.54	659061.04
127	17825.97	659067.20
128	17825.97	659067.20
129	17825.97	659067.20
128	17825.97	659067.20
270	17825.97	659067.20
271	17825.97	659067.20
272	17825.97	659067.20
273	17825.97	659067.20
274	17825.97	659067.20
275	17825.97	659067.20
276	17825.97	659067.20
277	17825.97	659067.20
278	17825.97	659067.20
279	17825.97	659067.20
280	17825.97	659067.20
281	17825.97	659067.20
282	17825.97	659067.20
283	17825.97	659067.20
284	17825.97	659067.20
285	17825.97	659067.20
286	17825.97	659067.20
287	17825.97	659067.20
288	17825.97	659067.20
289	17825.97	659067.20
290	17825.97	659067.20
291	17825.97	659067.20
292	17825.97	659067.20
293	17825.97	659067.20
294	17825.97	659067.20
295	17825.97	659067.20
296	17825.97	659067.20
297	17825.97	659067.20
298	17825.97	659067.20
299	17825.97	659067.20
300	17825.97	659067.20

הערות המידות:

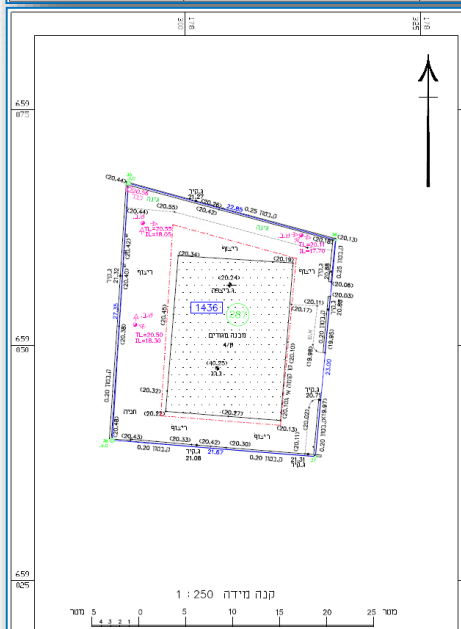
על מנת להימנע מבלבול, נקבעה גודל המידות...
 1. המידות נכונות לנתוני היישוב 2005-12.
 2. המידות נכונות לנתוני היישוב 2005-12.
 3. המידות נכונות לנתוני היישוב 2005-12.

שירות דב דור
 6781 הלש

Permit



Register



תוכנית

החוק: 1277.2016.4
 החוק: 1277.2016.4
 תחום: 1436
 חלק: 545

הערות:

1. תוכנית זו נכנסה לתוקף ב-27.12.2016.
2. תוכנית זו נכנסה לתוקף ב-27.12.2016.
3. תוכנית זו נכנסה לתוקף ב-27.12.2016.

נתונים כלכליים:

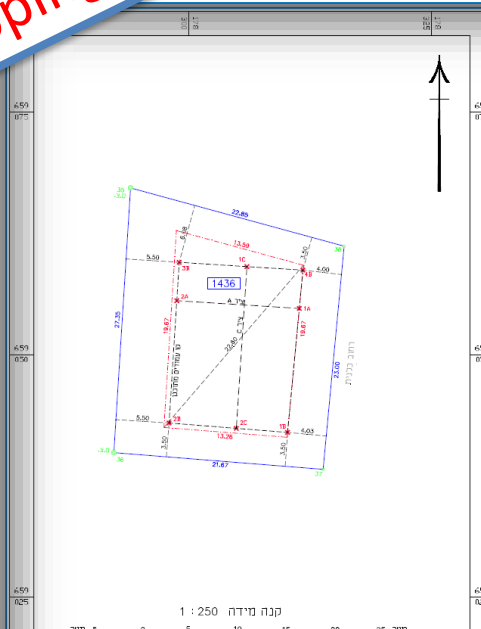
מס	שטח (מ"ר)	ערך (₪)
118	17814.59	659051.76
126	17816.54	659061.04
127	17825.97	659067.20
128	17825.97	659067.20
129	17825.97	659067.20
128	17825.97	659067.20
270	17825.97	659067.20
271	17825.97	659067.20
272	17825.97	659067.20
273	17825.97	659067.20
274	17825.97	659067.20
275	17825.97	659067.20
276	17825.97	659067.20
277	17825.97	659067.20
278	17825.97	659067.20
279	17825.97	659067.20
280	17825.97	659067.20
281	17825.97	659067.20
282	17825.97	659067.20
283	17825.97	659067.20
284	17825.97	659067.20
285	17825.97	659067.20
286	17825.97	659067.20
287	17825.97	659067.20
288	17825.97	659067.20
289	17825.97	659067.20
290	17825.97	659067.20
291	17825.97	659067.20
292	17825.97	659067.20
293	17825.97	659067.20
294	17825.97	659067.20
295	17825.97	659067.20
296	17825.97	659067.20
297	17825.97	659067.20
298	17825.97	659067.20
299	17825.97	659067.20
300	17825.97	659067.20

הערות המידות:

על מנת להימנע מבלבול, נקבעה גודל המידות...
 1. המידות נכונות לנתוני היישוב 2005-12.
 2. המידות נכונות לנתוני היישוב 2005-12.
 3. המידות נכונות לנתוני היישוב 2005-12.

שירות דב דור
 6781 הלש

As-built



תוכנית

החוק: 1277.2016.4
 החוק: 1277.2016.4
 תחום: 1436
 חלק: 545

הערות:

1. תוכנית זו נכנסה לתוקף ב-27.12.2016.
2. תוכנית זו נכנסה לתוקף ב-27.12.2016.
3. תוכנית זו נכנסה לתוקף ב-27.12.2016.

נתונים כלכליים:

מס	שטח (מ"ר)	ערך (₪)
118	17814.59	659051.76
126	17816.54	659061.04
127	17825.97	659067.20
128	17825.97	659067.20
129	17825.97	659067.20
128	17825.97	659067.20
270	17825.97	659067.20
271	17825.97	659067.20
272	17825.97	659067.20
273	17825.97	659067.20
274	17825.97	659067.20
275	17825.97	659067.20
276	17825.97	659067.20
277	17825.97	659067.20
278	17825.97	659067.20
279	17825.97	659067.20
280	17825.97	659067.20
281	17825.97	659067.20
282	17825.97	659067.20
283	17825.97	659067.20
284	17825.97	659067.20
285	17825.97	659067.20
286	17825.97	659067.20
287	17825.97	659067.20
288	17825.97	659067.20
289	17825.97	659067.20
290	17825.97	659067.20
291	17825.97	659067.20
292	17825.97	659067.20
293	17825.97	659067.20
294	17825.97	659067.20
295	17825.97	659067.20
296	17825.97	659067.20
297	17825.97	659067.20
298	17825.97	659067.20
299	17825.97	659067.20
300	17825.97	659067.20

הערות המידות:

על מנת להימנע מבלבול, נקבעה גודל המידות...
 1. המידות נכונות לנתוני היישוב 2005-12.
 2. המידות נכונות לנתוני היישוב 2005-12.
 3. המידות נכונות לנתוני היישוב 2005-12.

שירות דב דור
 6781 הלש

Construct

National mapping standard

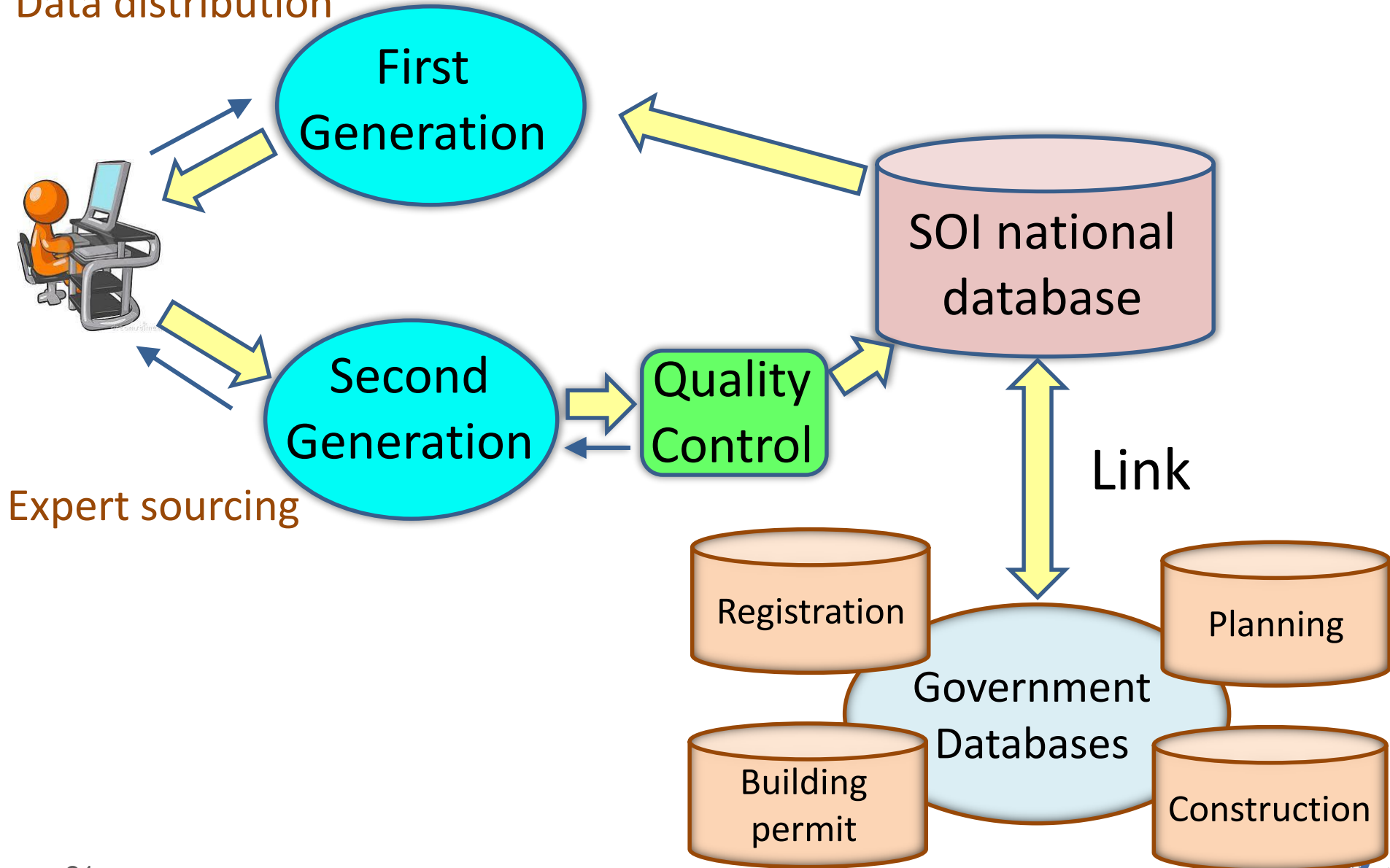
How to create a data lossless culture?

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

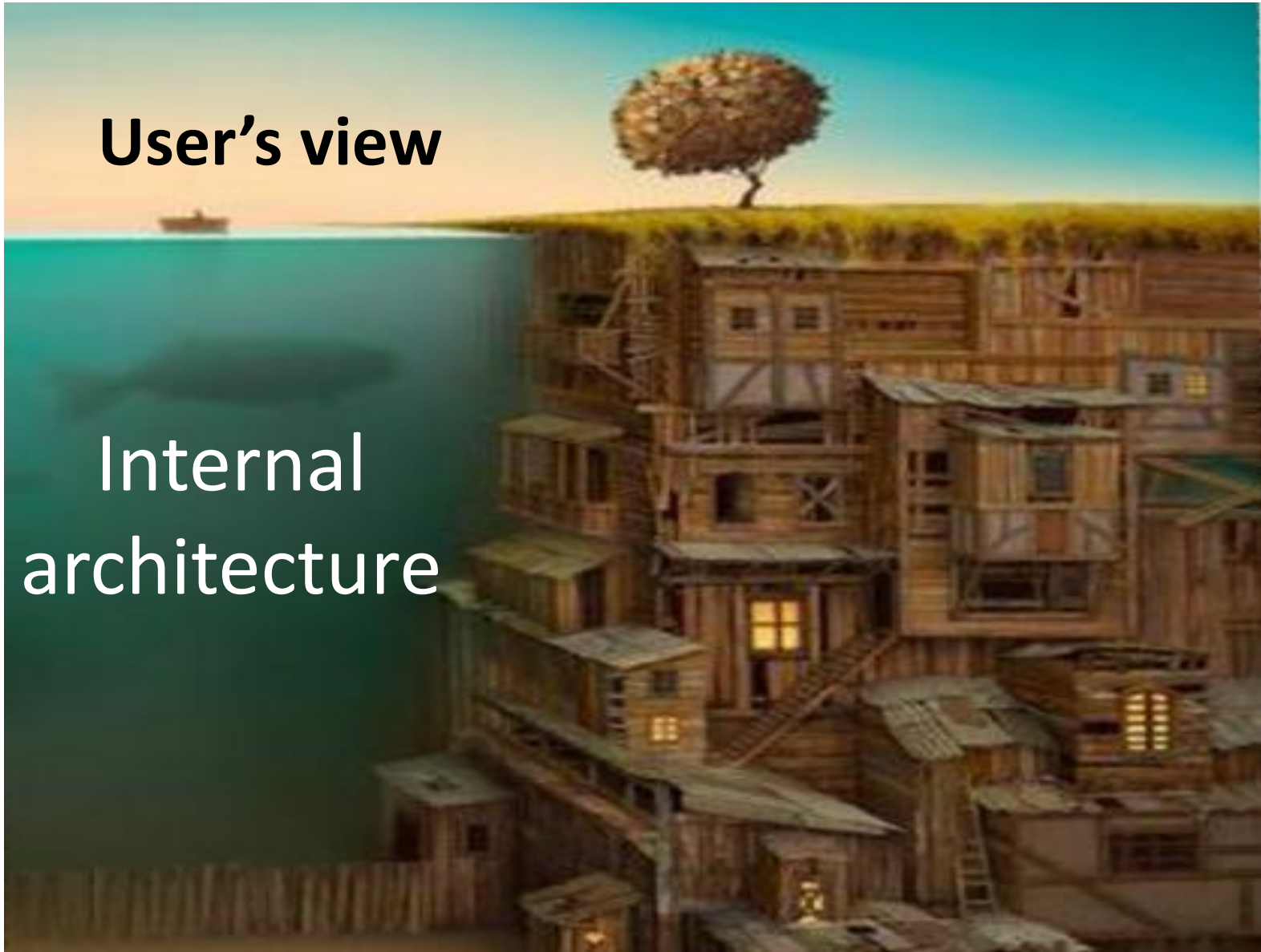
Data distribution



Simplicity – friendly user experience

User's view

Internal
architecture

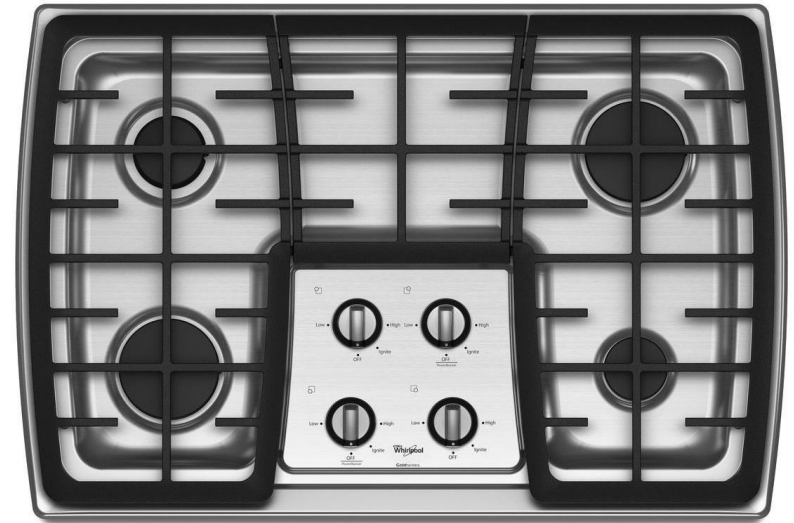


Simplicity – friendly user experience



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



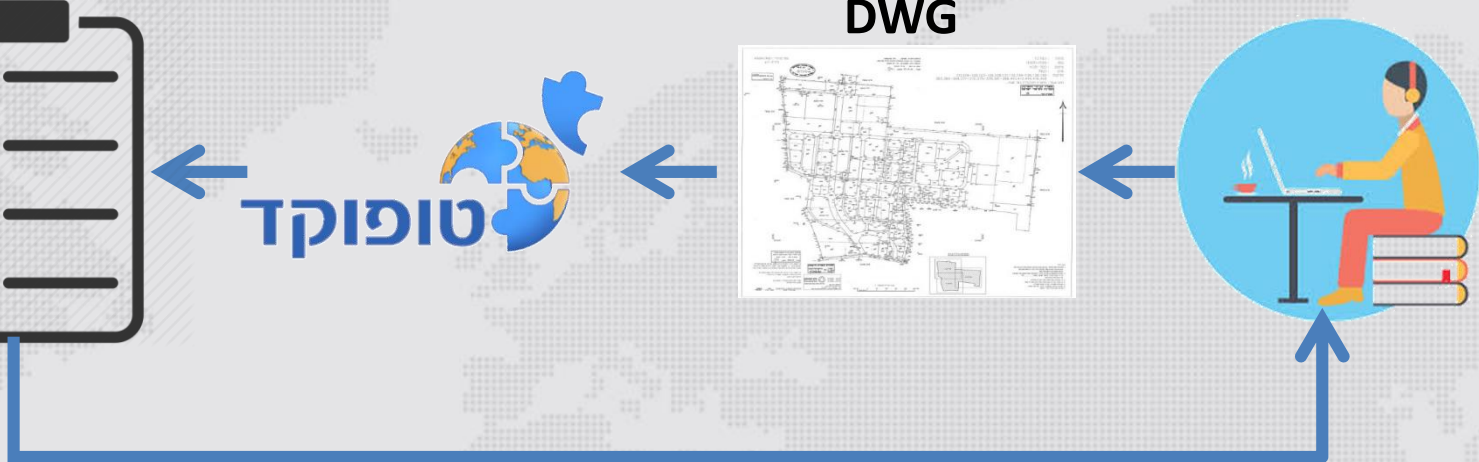
QA report



DWG



טופוקד



כל הזכויות שמורות

התקנת תעודות gov.il

כתוב אלינו

תנאי שימוש

אודות המערכת



Online QC tool

The screenshot shows the Topocad web application interface. The browser address bar displays "Secure | https://www.topocad.gov.il/#". The page header includes the user name "שלום ירון פלוס" and the "התנתק" (Logout) button. The main navigation bar contains "מלל חופשי", "גוש/חלקה", "נק' בקרה", "פסק דין", and "תצ"ר". Below this, there are buttons for "חיפוש" (Search), "חיפוש מתקדם" (Advanced Search), "תיחום במפה" (Map Delineation), and "האזור האישי שלי" (My Area). The main content area is titled "פרויקטים טופוגרפיים" (Topographic Projects) and features a "פרטי פרויקט טופוגרפי" (Topographic Project Details) form. The form includes fields for "שם פרויקט" (Project Name) set to "PROJECT GOLD", "סוג פרויקט" (Project Type) set to "טופוגרפית" (Topographic), and a "בחר עם מי לשתף את הפרויקט" (Share project with) dropdown. A file selection dialog is open, showing a list of DWG files. The file "AOI2.dwg" is selected and highlighted in yellow. The dialog also shows a preview of a map with various markers and a table of files.

Name	Date modified	Type	Size
AOI1 - Copy.dwg	26/09/2017 12:26	DWG Drawing	18 KB
AOI1.dwg	26/09/2017 12:26	DWG Drawing	18 KB
AOI2.dwg	26/09/2017 12:26	DWG Drawing	48 KB
AOI3.dwg	26/09/2017 12:26	DWG Drawing	
AOI4.dwg	26/09/2017 12:26	DWG Drawing	
AOI5.dwg	26/09/2017 12:26	DWG Drawing	
AOI6.dwg	26/09/2017 12:26	DWG Drawing	
declaration_OK.DWG	09/11/2017 12:14	DWG Drawing	
SURVEY_INON.dwg	02/10/2017 15:29	DWG Drawing	
Survey_mountain.dwg	02/10/2017 15:29	DWG Drawing	
TestTopo.dwg	07/11/2017 15:39	DWG Drawing	
tstFile2010.DWG	09/11/2017 10:02	DWG Drawing	

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

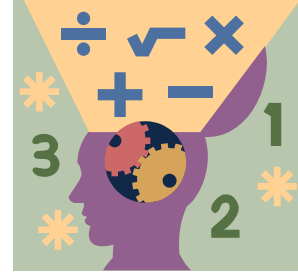
Temporal quality

Incompatible



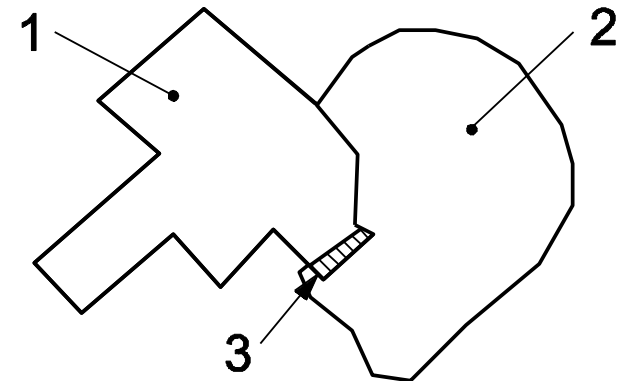
Compliance (metadata report)

Logical consistency



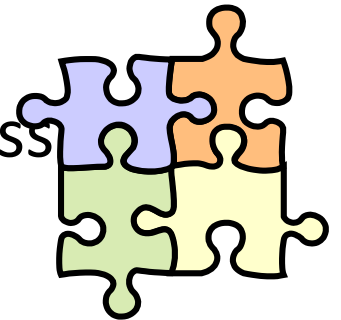
The program finds elements that do not conform to (logical) rules:

- 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

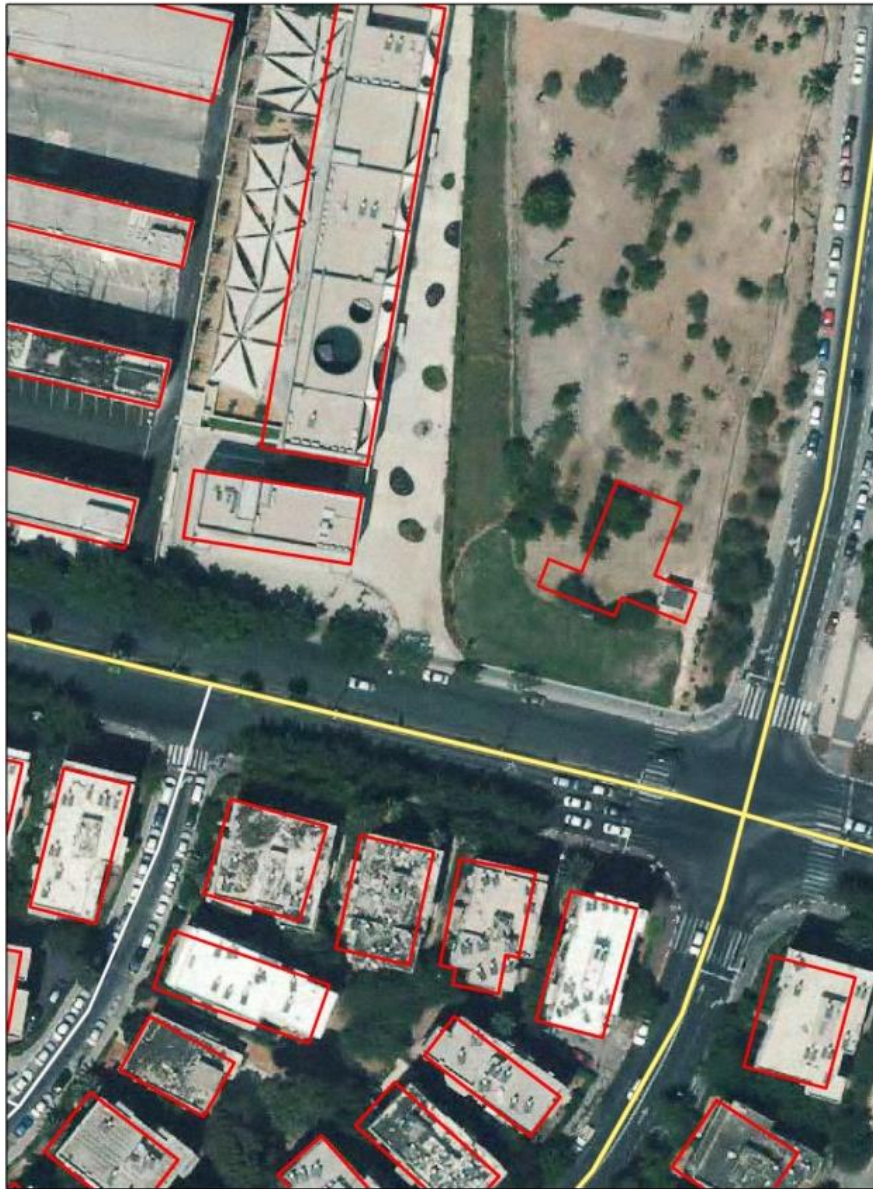


Completeness

- 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



© זכויות יוצרים: למוציא לאור כי כל הזכויות שמורות למרכז לסיימון ישראל, משרד"ש, יוני און ראשון, לרעם, להנצחה, רח, לשרעם, להפיק, לאחסן במאגר מידע, לשחזר או לקלוט בכל דרך או בכל אמצעי אלקטרוני, אופטי, מכני, או אחר, כל חלק שהוא מהחומר הכלול בנפיש דו אסיר בהחלט אלא באישור מפורט ובכתב מנכ"ל המרכז לסיימון ישראל

Omission



© זכויות יוצרים: למוציא לאור כי כל הזכויות שמורות למרכז לסיימון ישראל, משרד"ש, יוני און ראשון, לרעם, להנצחה, רח, לשרעם, להפיק, לאחסן במאגר מידע, לשחזר או לקלוט בכל דרך או בכל אמצעי אלקטרוני, אופטי, מכני, או אחר, כל חלק שהוא מהחומר הכלול בנפיש דו אסיר בהחלט אלא באישור מפורט ובכתב מנכ"ל המרכז לסיימון ישראל

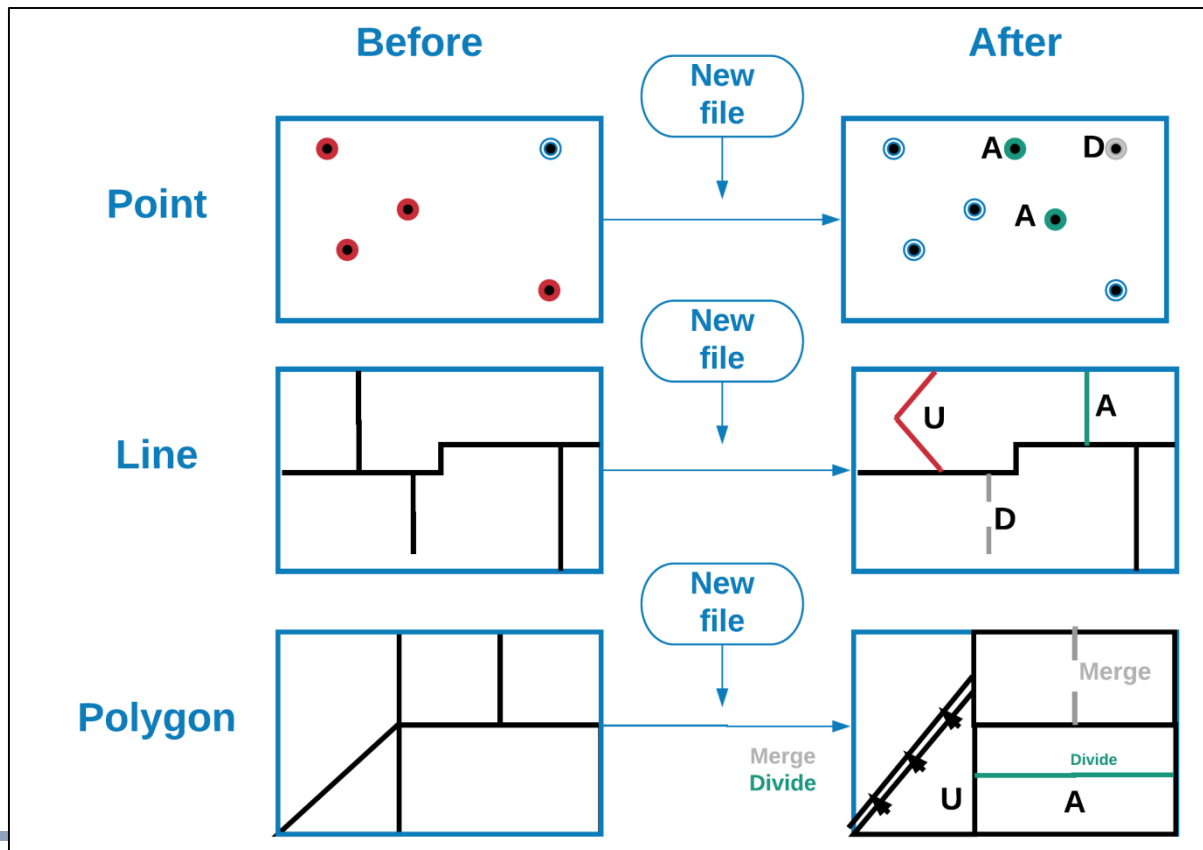
Thematic and positional accuracy

- 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.

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

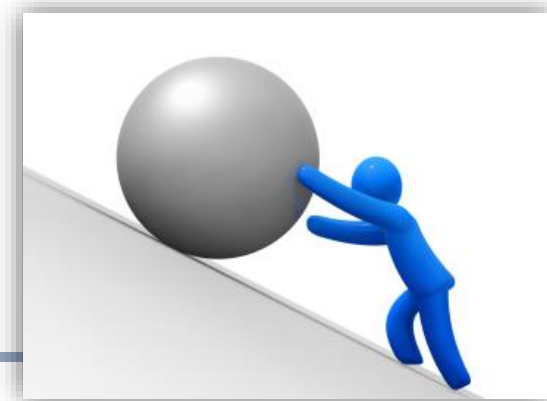
The data integration process (conflation)

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

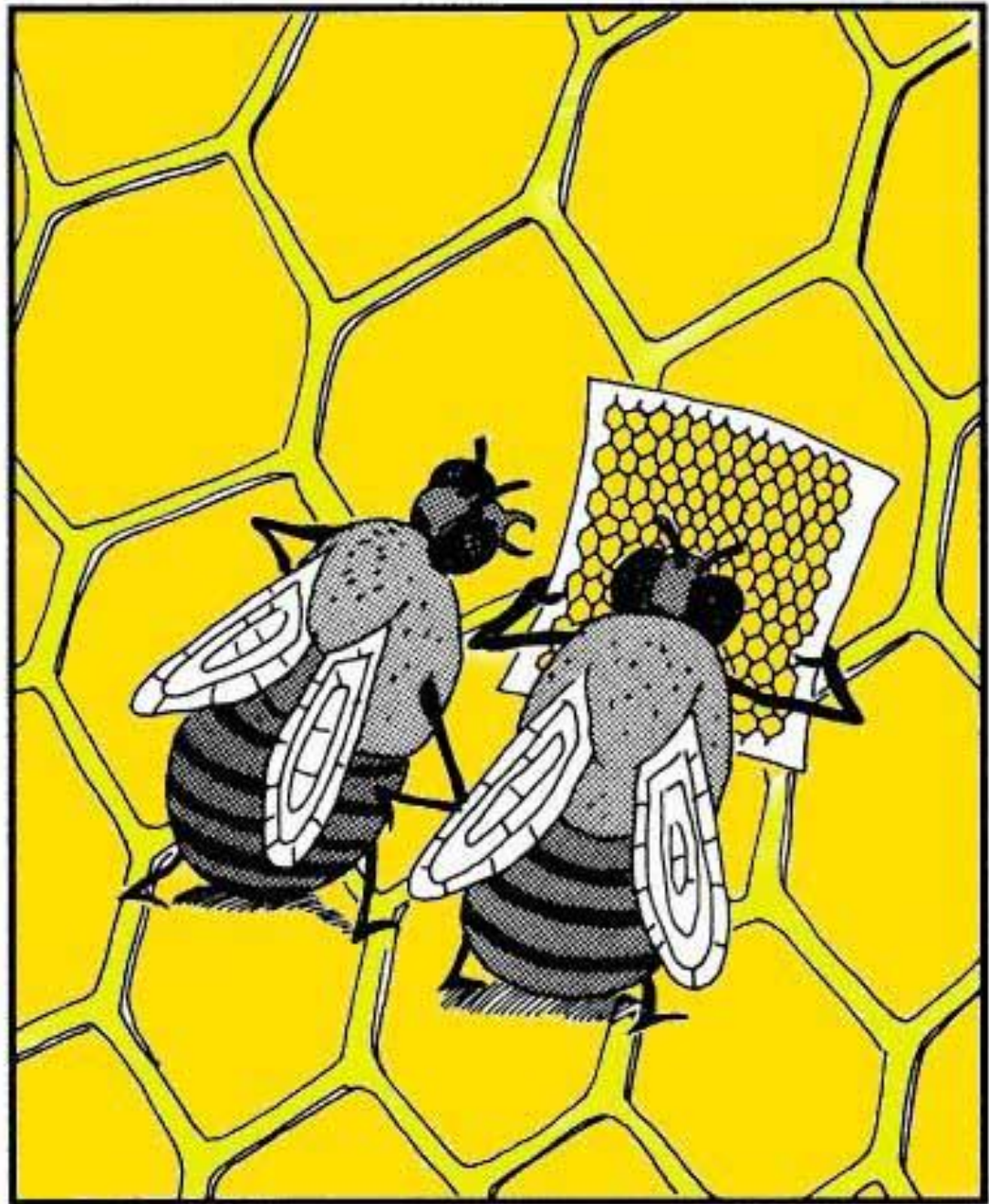


Summary

- 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

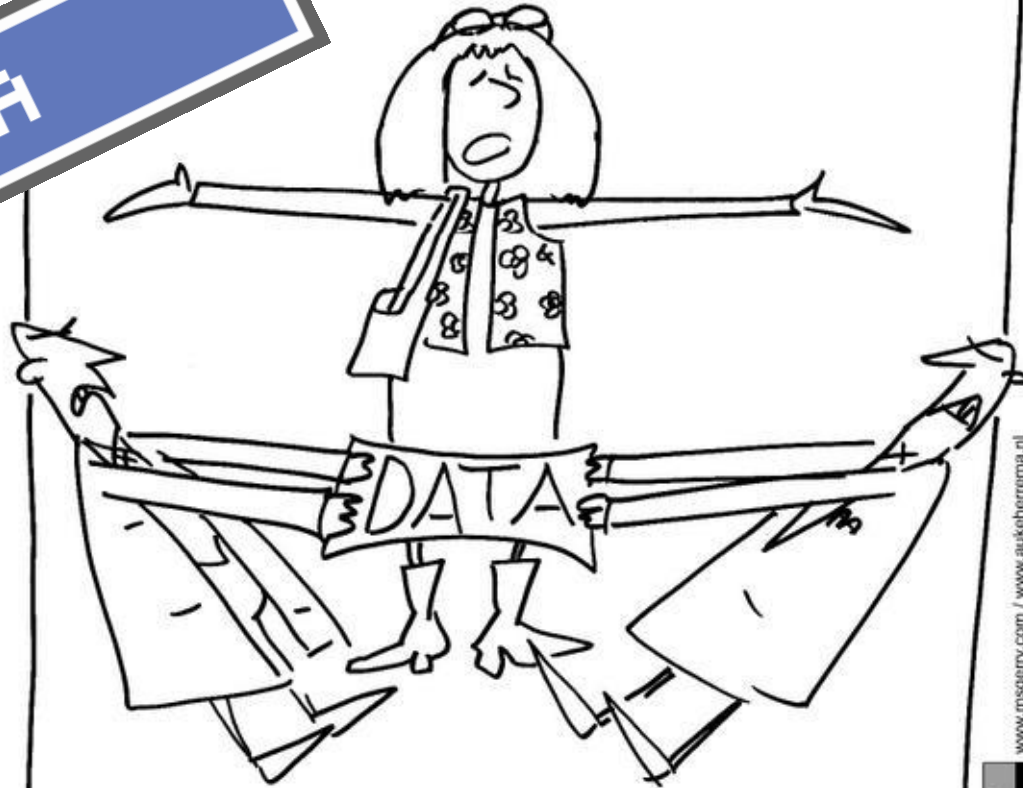


So, Where are we exactly?

DO NOT BE SO 2014

SHARE!

OPEN DATA

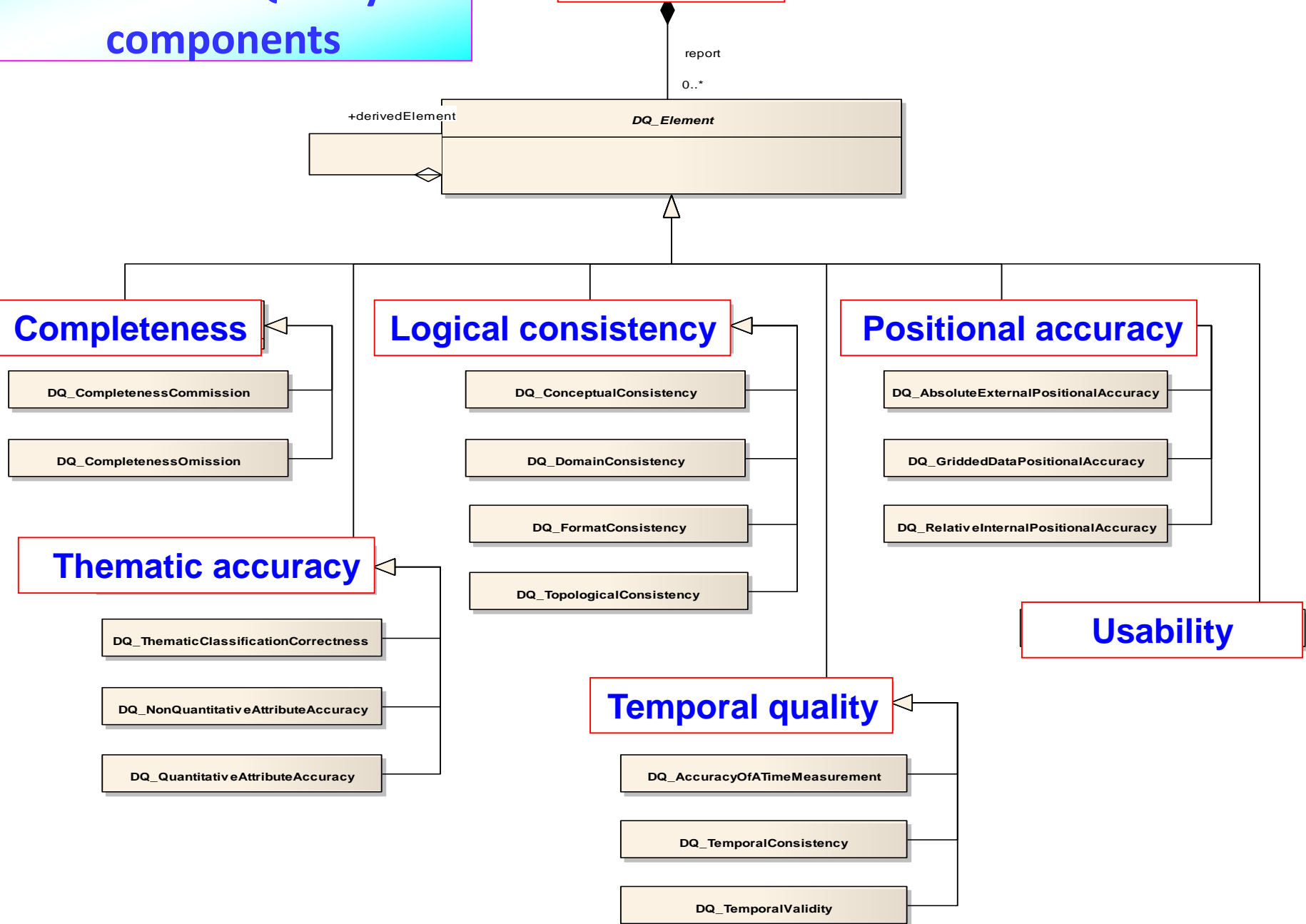


AMAF

www.insgerry.com / www.aukehorrema.nl

ISO/CD 19157 Standard Quality components

Data Quality



Why Geo-spatial technologies?

Geo-spatial technology underpins smart nation operations

