THE WHAT, WHY AND HOW OF OPEN DATA

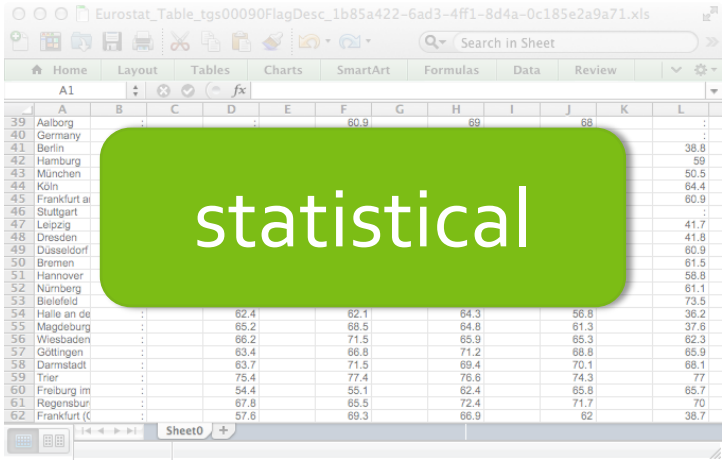
JENI TENNISON
TECHNICAL DIRECTOR

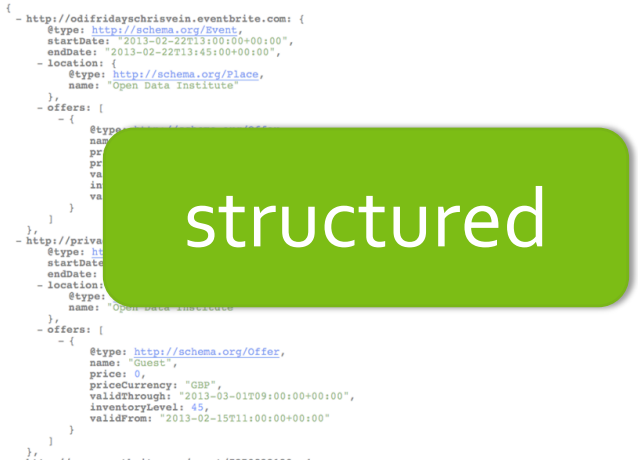
@JENIT
JENI@THEODI.ORG

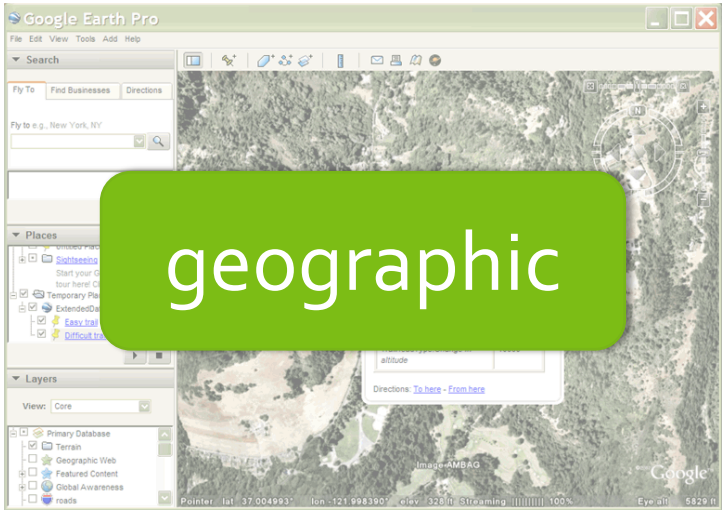
CC BY SA

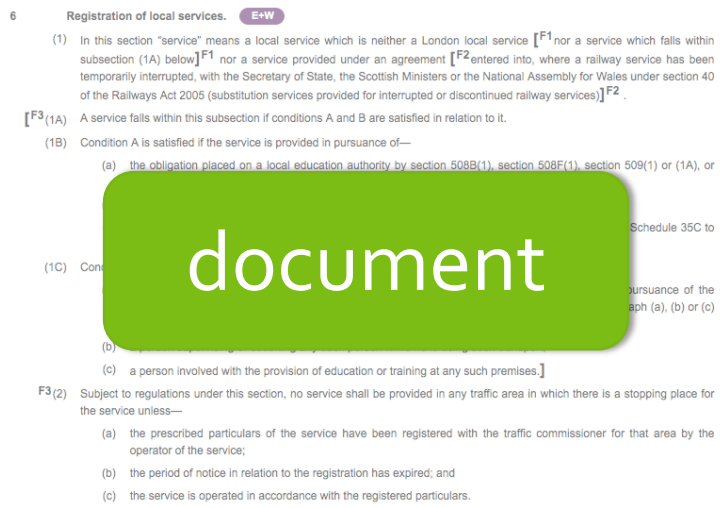
WHAT?

WHERE DO YOU GET "DATA"?

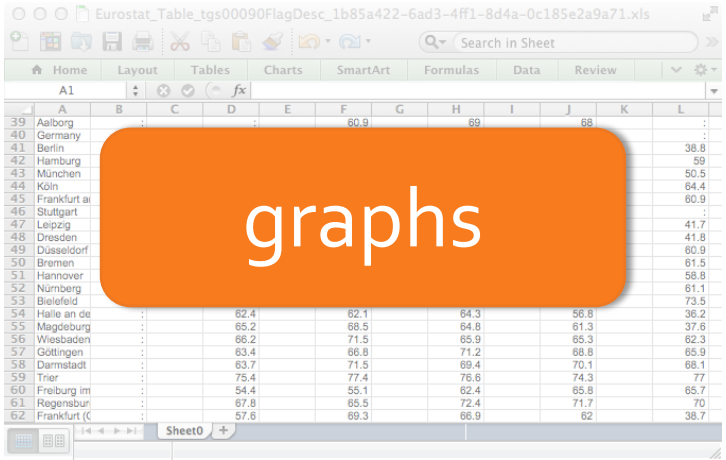


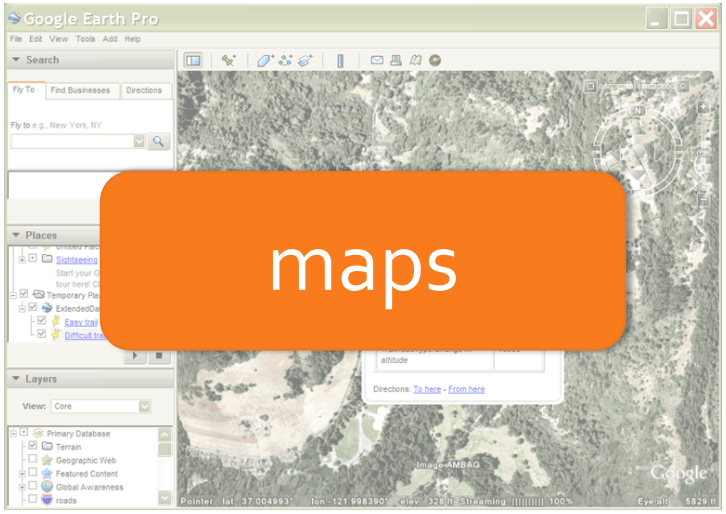




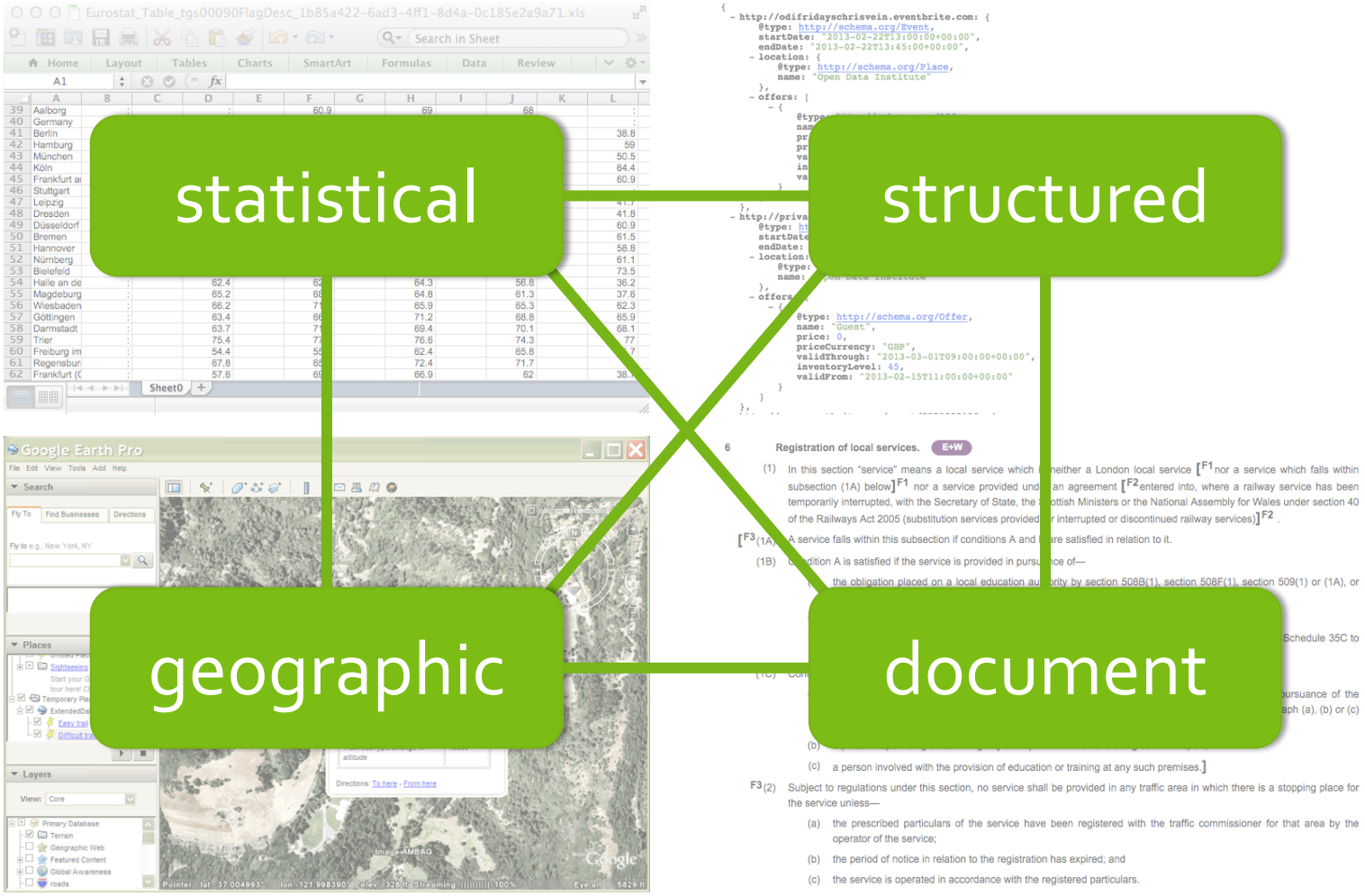


DATA CAN BE REUSED





DATA CAN BE COMBINED



WHAT DOES "OPEN" MEAN?

- **for everyone**
  - not limited by funds
  - not limited by who they are
- **to do anything**
  - analyse and present in different ways
  - combine with other data
  - republish for other people
  - make money

MOST DATA IS NOT OPEN

- You must not use or attempt to use any automated program to access our system or this web site, or to search, display or obtain links to any part of this site. All access to and use of the services via any automated software agent is prohibited. This includes without limitation, any mechanical program, screen scraper, spider or other web crawler.

Landmark's Energy Performance Certificate website

3.1 Restrictions on Commercial Use of TomTom Licensed Content.

(a) You are not permitted to print more than 5,000 copies of sales collateral materials containing a screenshot of TomTom Licensed Content for commercial sales lead generation ("Direct Marketing"). If you desire to do so, you must (i) enter into a Google Enterprise license agreement or (ii) contact TomTom to obtain a direct license to do so.

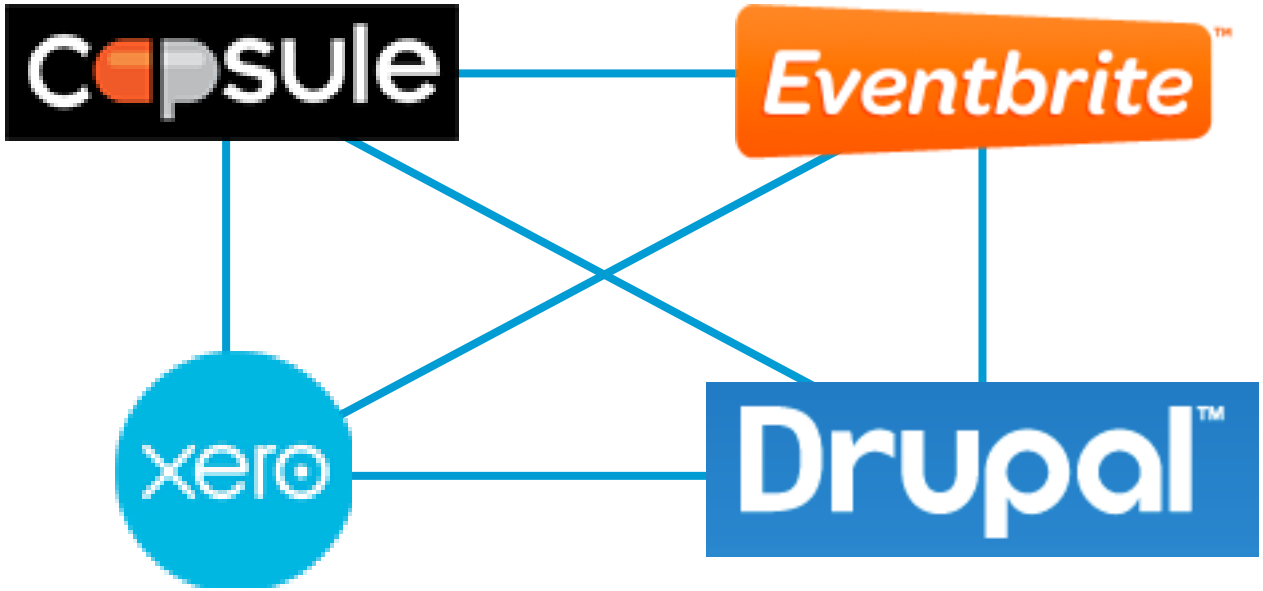
(b) You are not permitted to incorporate TomTom Licensed Content as a core part of printed matter (such as printed maps or guide books) that you redistribute for a fee. If you desire to do so, you must contact TomTom to obtain a license.

(c) You are not permitted to offer a batch geocoding service that uses the TomTom Licensed Content contained in any Google products or services.

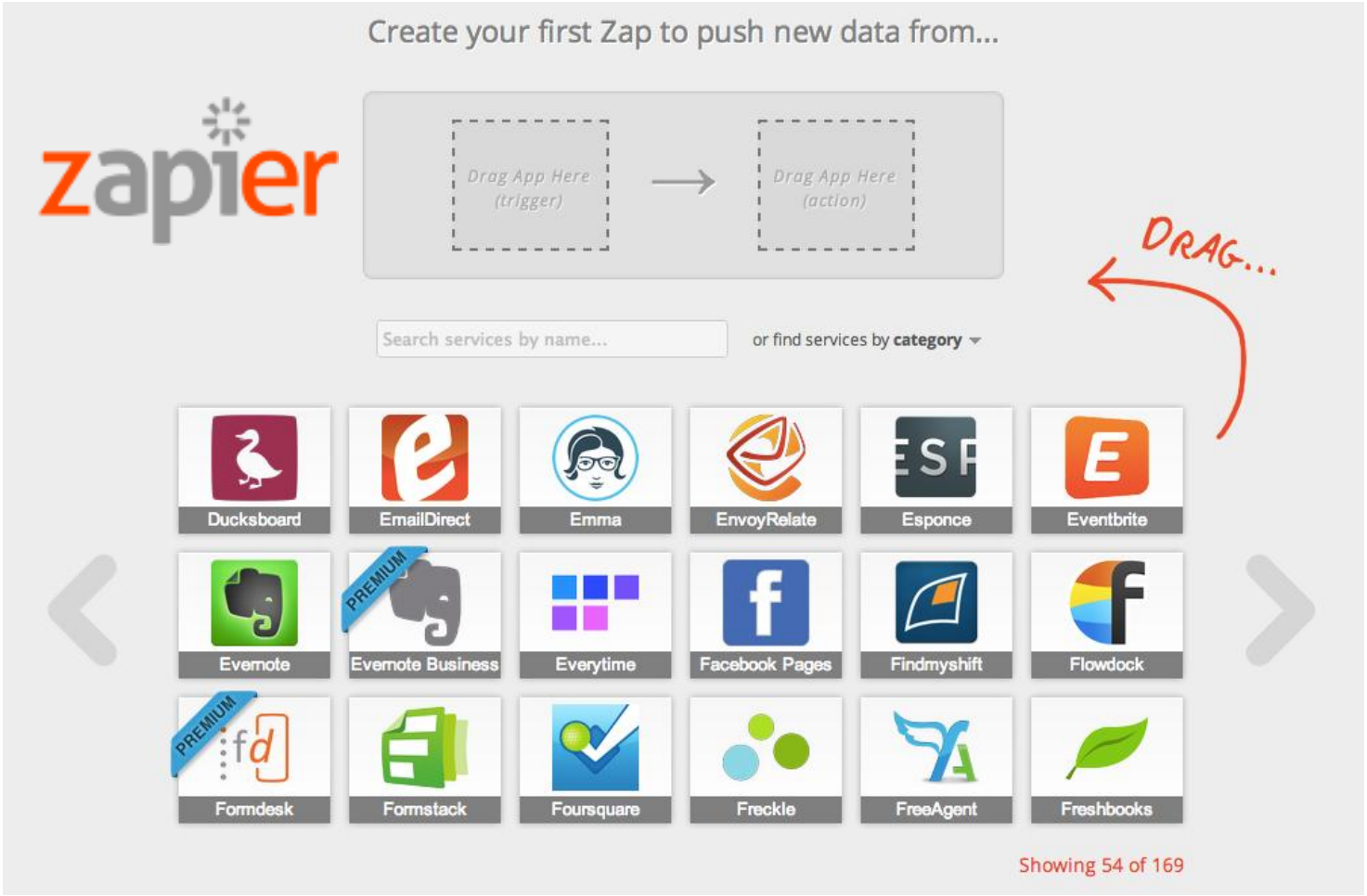
http://www.google.com/intl/en-us/help/legalnotices_maps.html

WHY? (1)

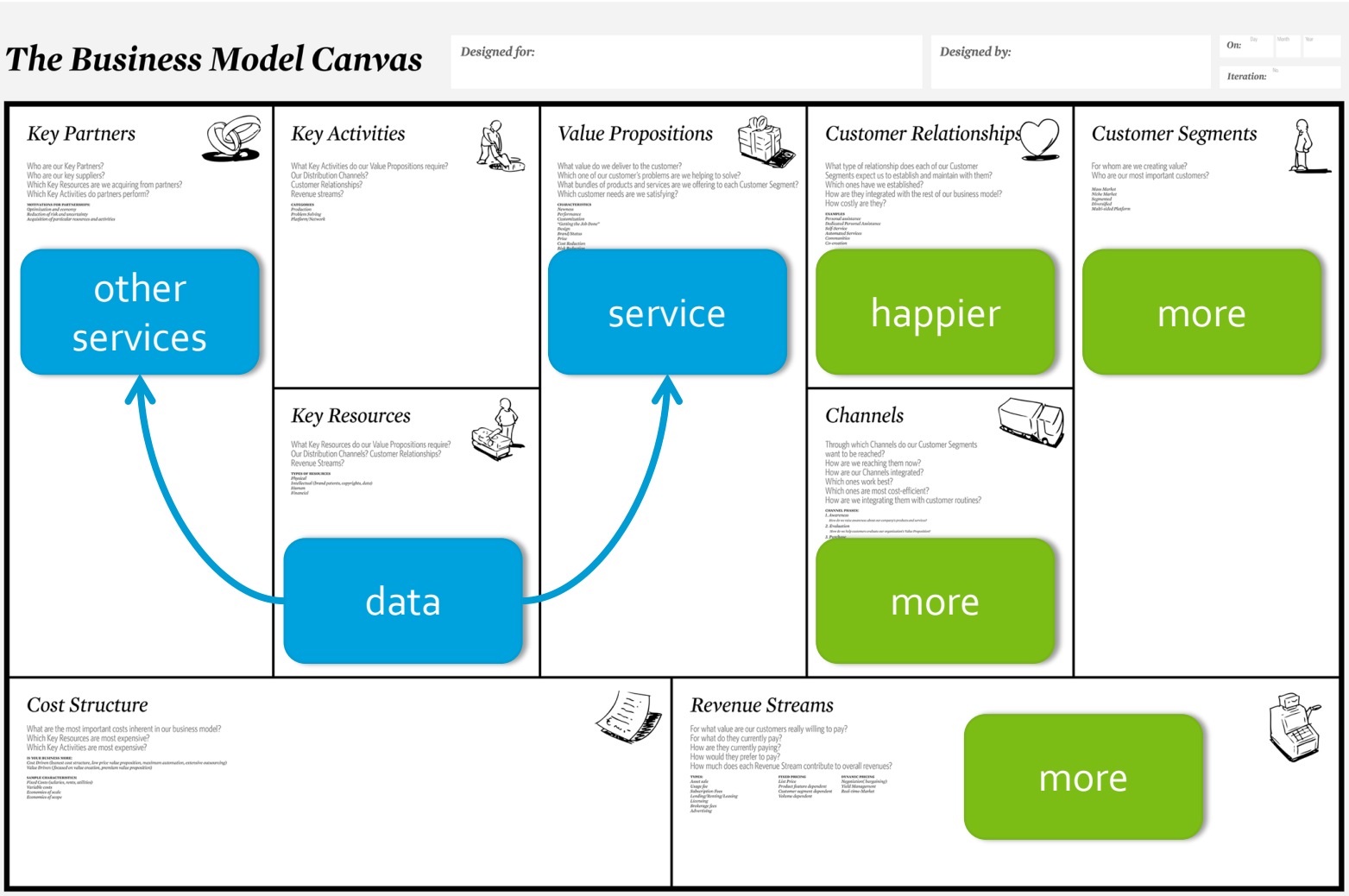
GREAT ONLINE SERVICES



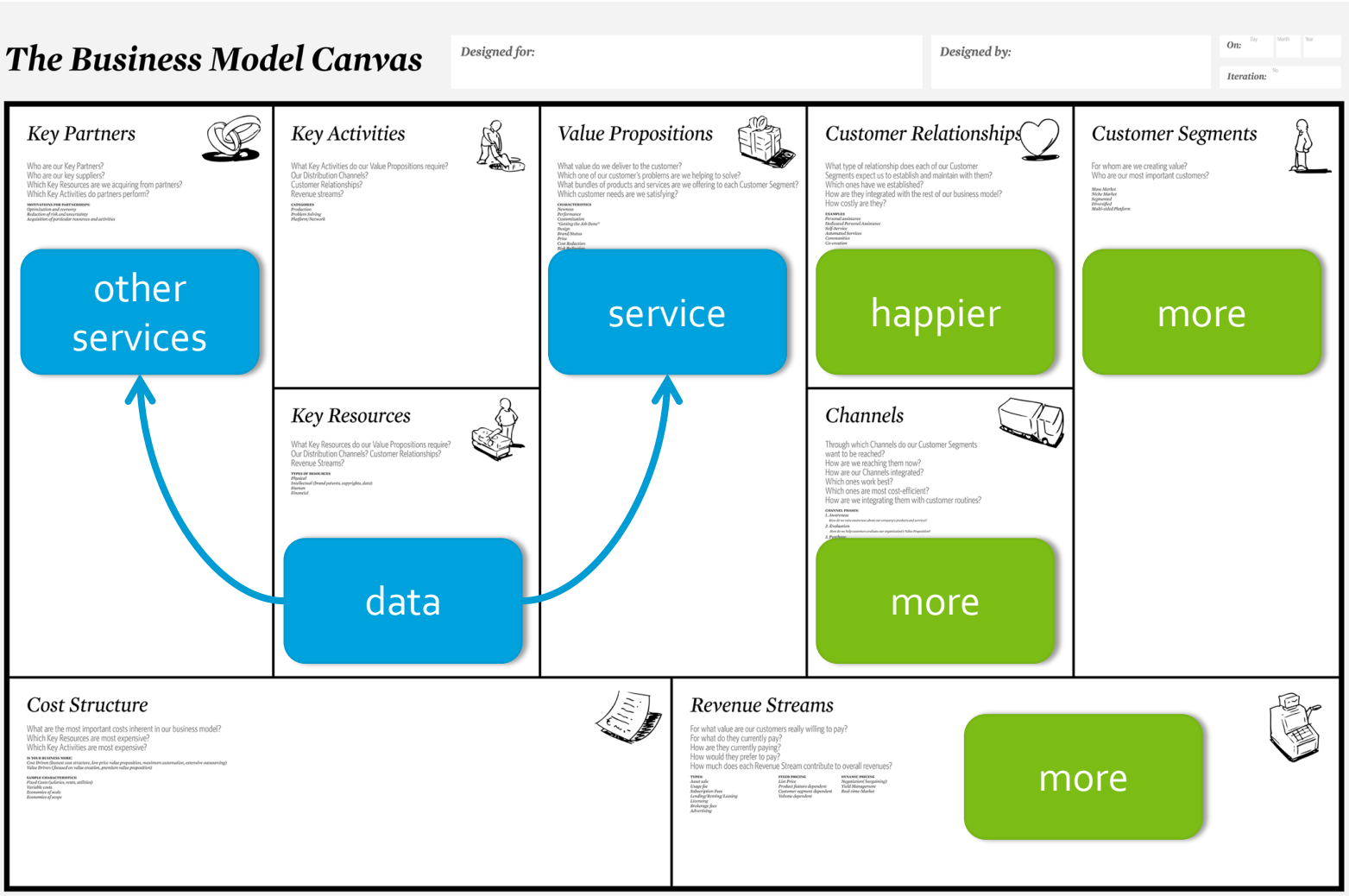
APIS EVERYWHERE



WHY DO THEY HAVE APIS?



WHY OPEN DATA?



WHY? (2)

ENCYCLOPEDIAS

WIKIPEDIA
The Free Encyclopedia

ENCYCLOPÆDIA
Britannica®

CODE

MAPS

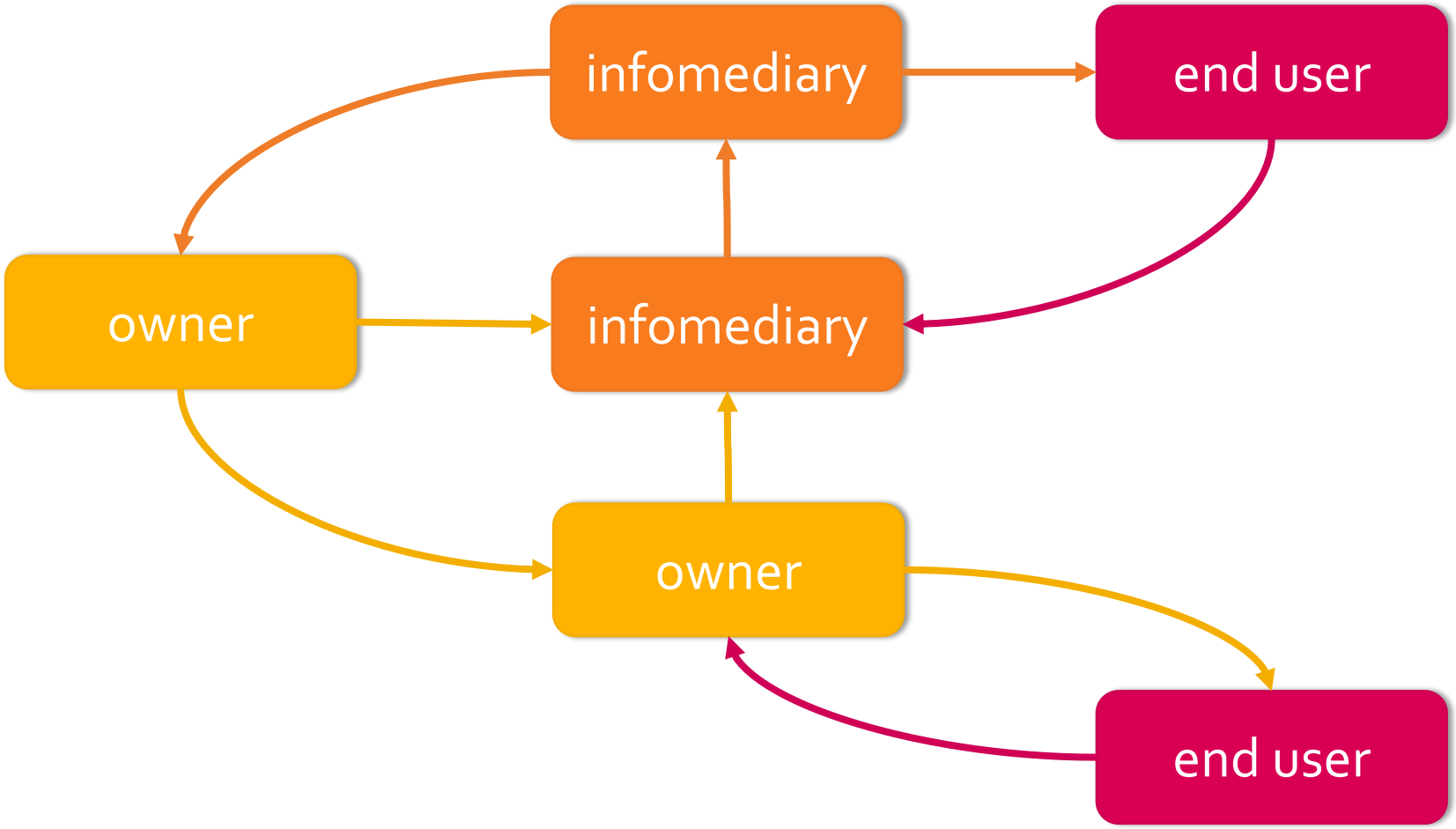
OpenStreetMap

Ordnance Survey®

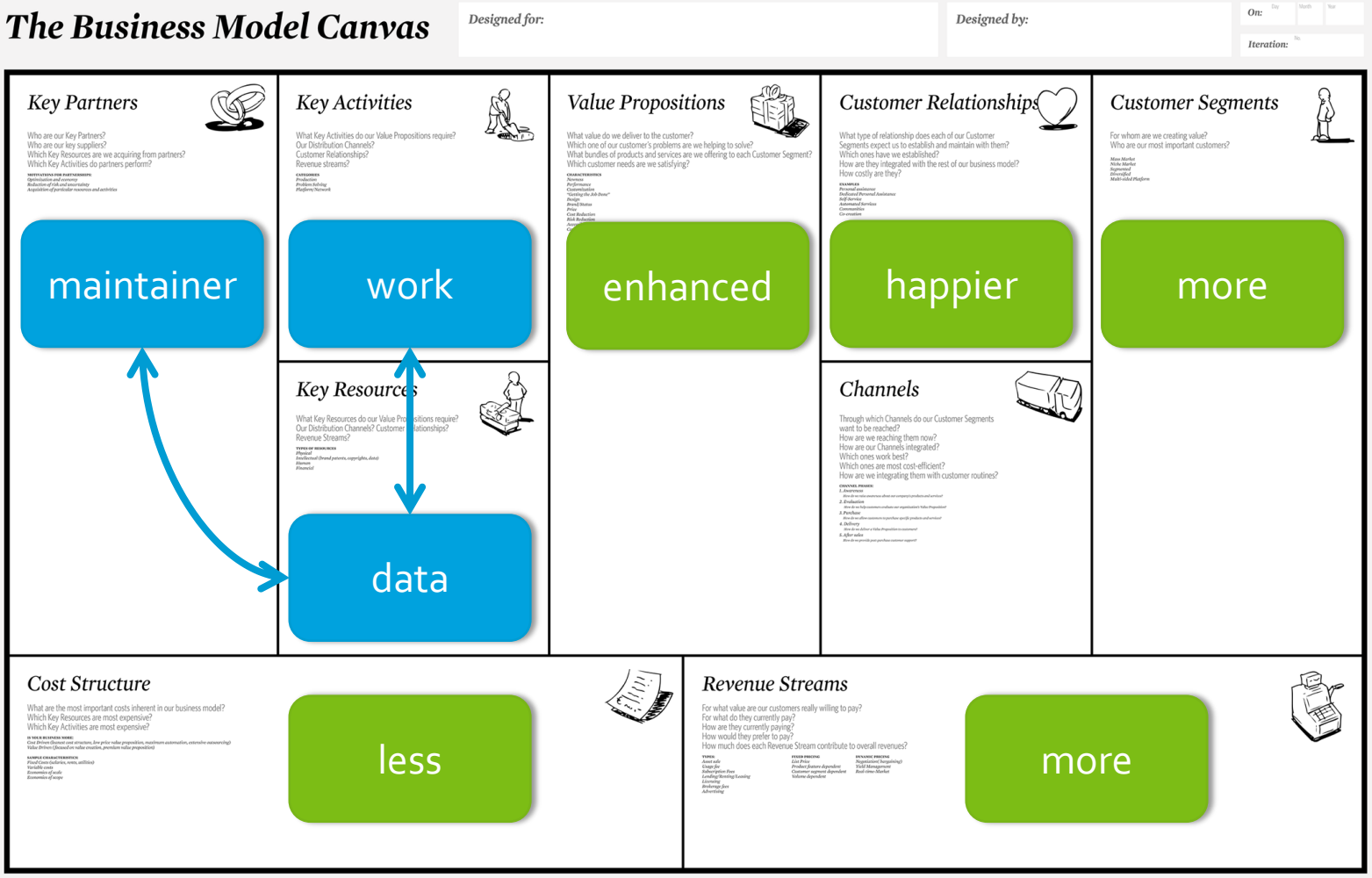
COMMONS-BASED PEER PRODUCTION

"a new model of socio-economic production in which the creative energy of large numbers of people is coordinated (usually with the aid of the Internet) into large, meaningful projects mostly without traditional hierarchical organization"

NETWORKS OF DATA



WHY OPEN DATA?

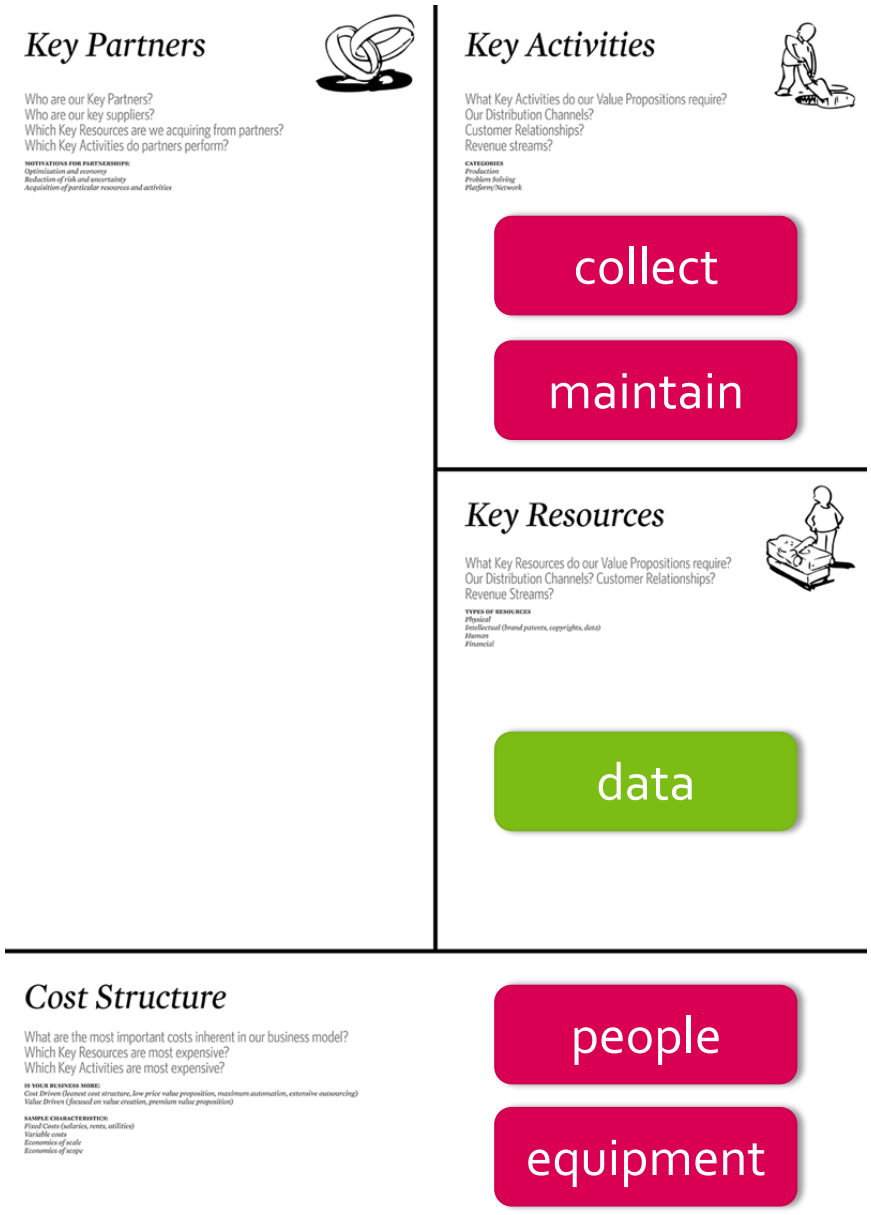


HOW?

1. IDENTIFY THE DATA YOU HAVE

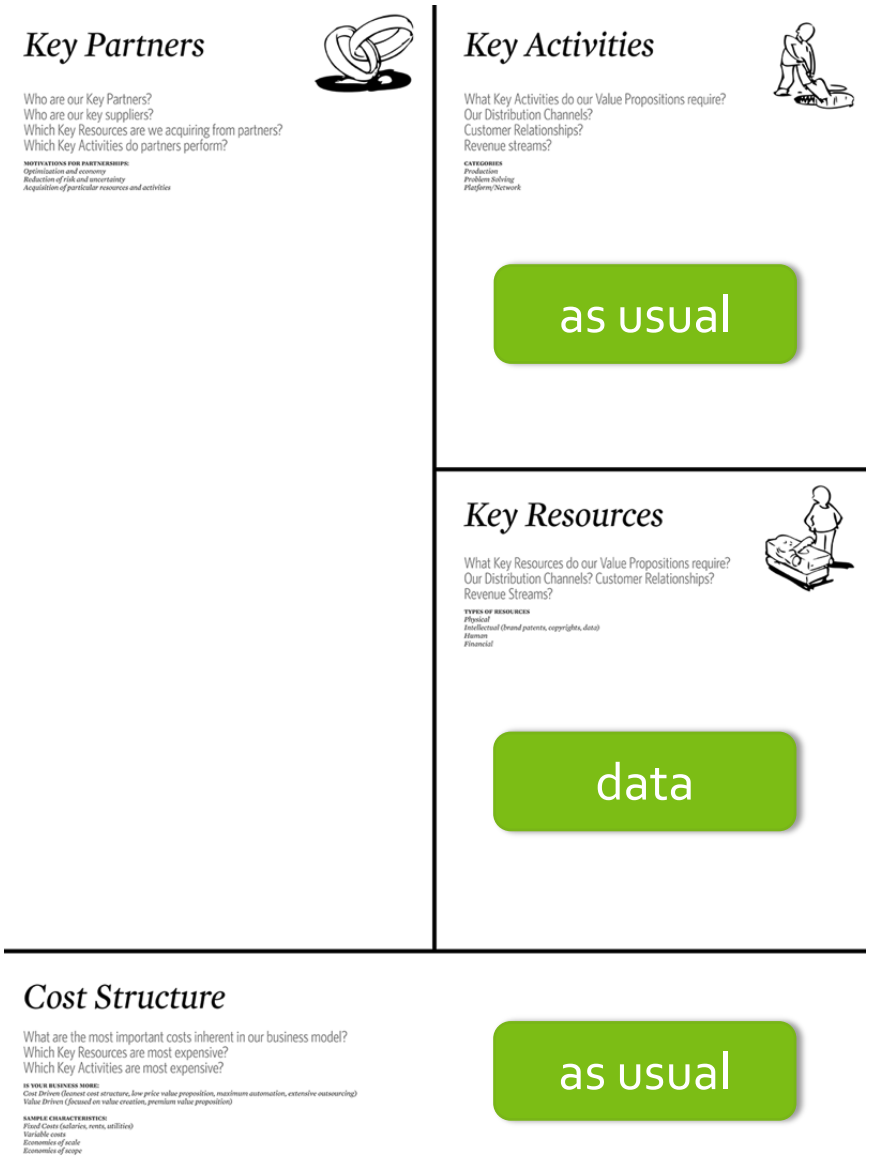
PRIMARY DATA

- **takes effort**
  - collect
  - maintain
- **and investment**
  - people
  - equipment
- **examples**
  - Met Office
  - book publishers
  - census



EXHAUST DATA

- **support activity**
- **no extra effort**
- **no extra cost**
- **examples**
  - till receipts
  - phone usage
  - customer data
  - accounts



DO YOU OWN IT?

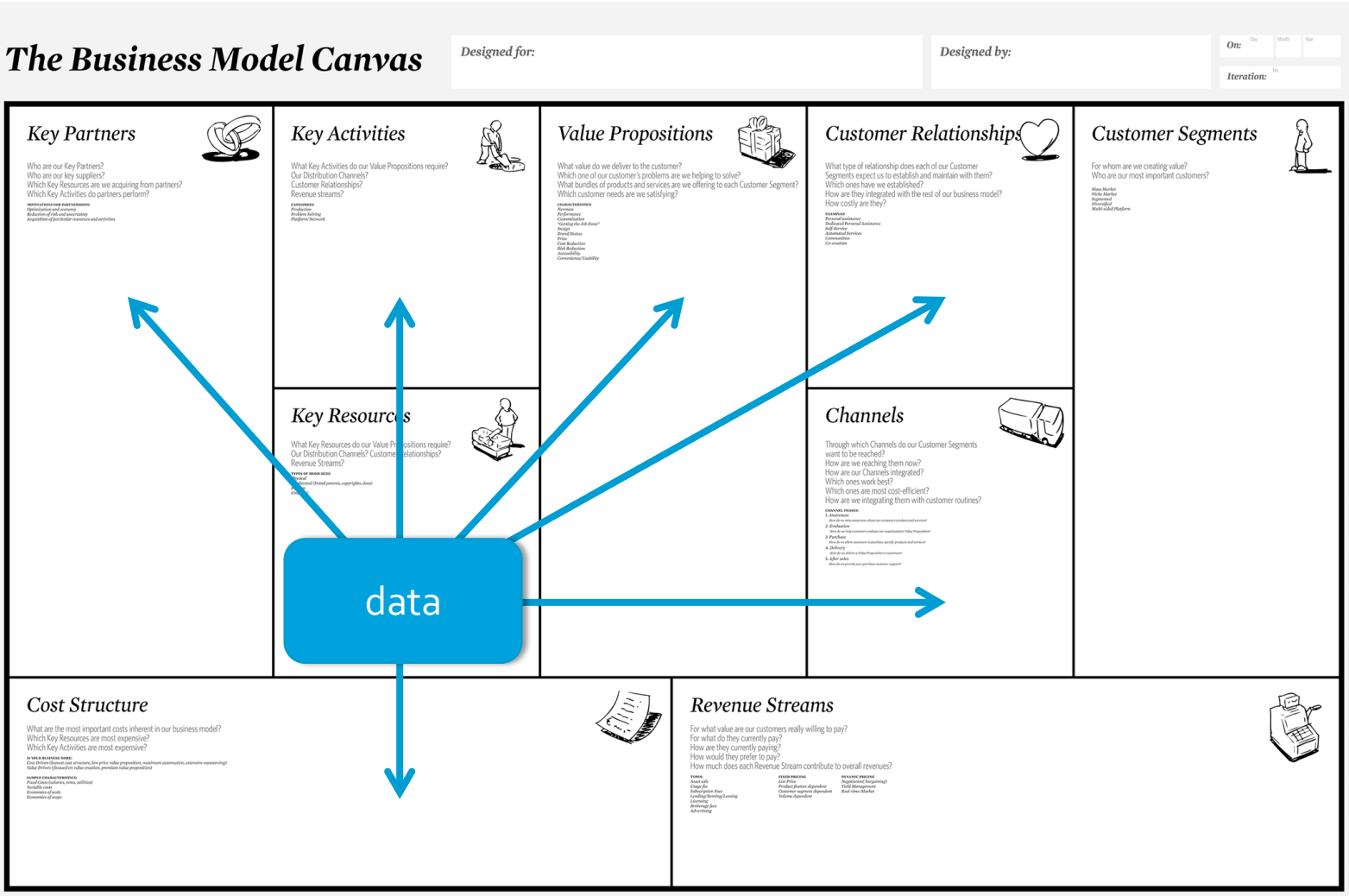
created data	you own the data you gather
derived data	you might own data you calculate from others' data
bought data	you almost certainly don't own data you've got from someone else

IS IT PERSONAL?

non-personal	data that has nothing to do with people or groups of people
aggregate	summaries of data about people or groups & their activity
personal	data about people or groups & their activity

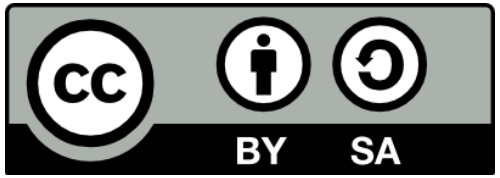
2. WORK OUT WHAT YOU WANT TO ACHIEVE

OPEN DATA IS A TOOL



3. MAKE IT OPEN

CHOOSE A LICENCE

	attribution	attribution & share-alike
content		
data	ODC-by	ODC-ODbL

ADD THE LICENCE

in HTML

```
<a resource="{{database-url}}"
   rel="license"
   href="{{licence-url}}">
   {{licence-name}}
</a>
```

in HTTP

```
Link: <{{licence-url}}>; rel=license
```

4. ITERATE & IMPROVE

TECHNICALLY

- **standard formats**
- **APIs help some reusers**
- **dumps help other reusers**
  - whole dataset analyses
  - also provide feeds of changes
- **links & link relations**
  - pull together disparate sources

HIGH QUALITY OPEN DATA

- **make it trustworthy**
  - accurate, timely
  - guaranteed availability
- **create a community**
  - feedback, including APIs
  - discussion lists
  - code libraries

CHALLENGES

WE DON'T KNOW YET

- which business models work
- which licences drive good behaviour
- how to measure open data use
- how to make data findable
- whether data formats really matter
- whether links in data really matter
- whether the world is better in the end

Open Data Institute

@JENIT JENI@THEODI.ORG

QUESTIONS?