17 July 2013

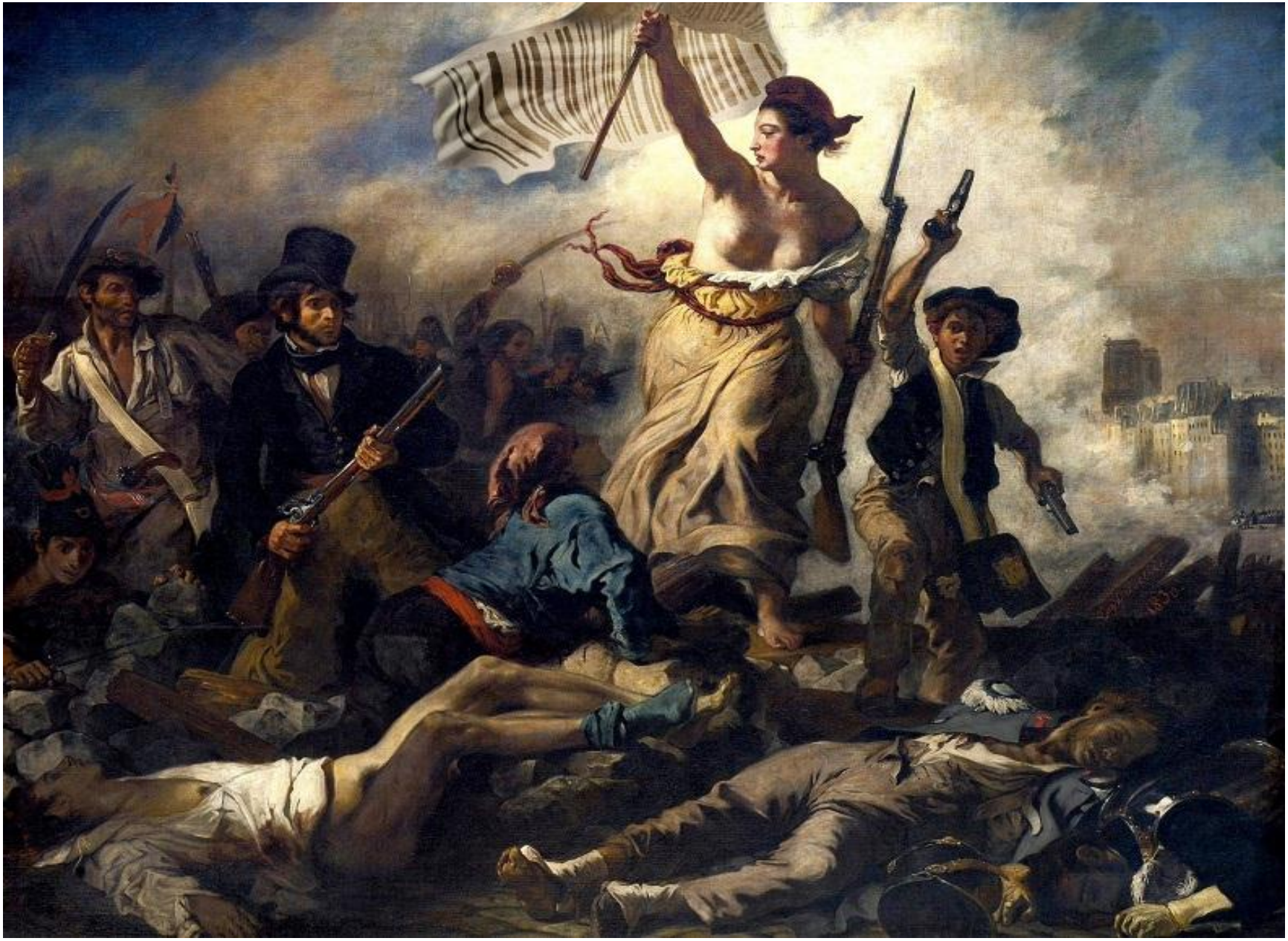
Product Open Data

Status – 17/07/2013

Philippe Plagnol

Free access to product data

product-open-data.com



La liberté guidant le peuple, Eugène Delacroix, 1830 redesigned by Jessica Dere

Current Status 13-07-2013

- **POD Official supporters** : INC, 60 millions de consommateurs, Etalab (data.gouv.fr)
- **POD database users** :
  - OpenDataSoft (pod.opendatasoft.com)
  - LSA (coming soon)
- **POD facts :**
  - 2 millions of GTIN codes collected (including 1million of high tech products provided by icecat.fr)
  - 600 000 products with a brand and owner assigned
  - 350 000 GCP (prefix codes)
  - 5 000 brands (managed manually)

Overview

Data printed on the products is not published in numeric format

Product public data is stored in their database

Stores / e-commerce

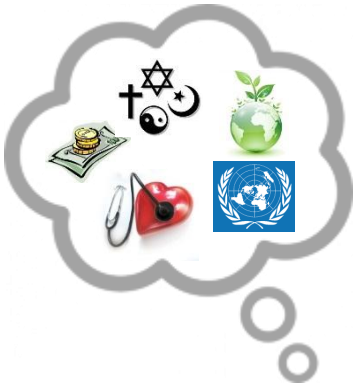
Create their own list of products and classifications

Hide barcode to be protected from competitors

Technology

Big Data enables the management and analysis of huge amounts of data

Smartphones enable real time product data by scanning the barcode



Public Product Repository based on barcode

FACTUAL DATA

OPEN DATA

Manufacturers

Numeric catalog (RSS for products)

Use barcode as a communication channel

Consumers are waiting for tools helping them to choose

They are ready to use barcode as a new communication channel

Consumers

Owners of barcodes + GPC classification

GS1 barcode system owner & supply chain focused

GS1 is currently investigating how open data about products is already being used - and the potential benefits, opportunities and risks.

This might result in a rather different approach from their B2C/MobileCom project, which did not provide open data about products.

Laws

Governments

French government is handling the subject

EU law * (end 2014) obligation to provides nutrition information

Use barcode to communicate on products

Consumers protection/information organizations will wake up soon…

Consumers Protection/Information organizations (+ ecology, environment, human rights organizations,…)

* http://ec.europa.eu/food/food/labellingnutrition/foodlabelling/proposed_legislation_en.htm

Global approach

- POD must belong to an organization assuring the neutrality (Manufacturers vs Consumers) and free access to product data
- **No CrowdSourcing** ! Product data must be provided by manufacturers: these data are already stored in their database and used to print packaging.
- GS1 should be the leader of Product Open Data movement because of its product expertise, international presence (150 countries) and critical role in the barcode registration process.
- POD will continue to use e-commerce websites to prevent manufacturers from refusing to provide product data. The core idea is "Product data should be available to public, rather than being restrained by manufacturers".
- POD is lobbying government to pass a law requesting manufacturers to publish a free numeric catalogue of their products.

Sourcing

- POD project follows a very simple guideline: product data has to be provided by manufacturers because the information is already in their database and printed on packages.
- A simple flat file (Excel, XML,…) is enough to publish a numeric catalogue (list of their barcode + data printed on packages).
- Potential resistance from manufacturers:
  - Arguments about Financial issues(Cost of Website upgrade, Microdata implementation, web service,…)
  - Arguments about technical issues(Flash website, use of a CMS,…)

Identification of the source	Accessibility	Exploitability	Matching
No CrowdSourcing • Manufacturer • E-commerce website • Search Engine (Google,…) • Retailer Listing of products • …	• Scraping (HTML page, Google,…) • Flat File (Excel, XML,…) • Web Service (Soap, Rest,…) • Microdata (RDF,…) • RSS for products (PSS) • …	• Data printed on packages is public • Visuals/Pictures are not public • Data has different level of sensitivity (ex: nutritional info vs. brand name)	• GTIN Code will join the product data with other kind of info i.e. price, comments • Life cycle of GTIN code and potential reuse of the code

Data Quality

Quality issues can be detected by using the following steps:

- **Step 1**: Check code validity by verifying the last digit following the GS1 rule www.gs1.org/barcodes/support/check_digit_calculator
- **Step 2**: Check if the first 3 digits are approved by GS1 by using the following prefix list (ex: there's no barcode starting with 2) www.gs1.org/barcodes/support/prefix_list
- **Step 3**: Verify first 3 digits: code for countries vs. books/publications
- **Step 4**: GEPIR is a repository of the manufacturers' info provided by GS1. First 6 to 10 digits are assigned by GS1 to the manufacturers (ex: products starting with 0004900 belongs to Coca-Cola company) gepir.gs1.org/v32/xx/gtin.aspx
- **Step 5**: GTIN code can be reused for the following reasons:
  - New Product www.gs1.org/docs/idkeys/GS1_GTIN_Allocation_Rules.pdf
  - Seasonal offers (Christmas, Summer, Valentine,…)
  - Special offers:
    - Apply on offers of additional value (ex: 1 liter + 20% free needs a new GTIN code)
    - Apply on offers of giveaways (ex: : 6 bottles of water + 2 free)
    - Others

Use case #1

Purchase Intelligence

- Barcode is an unique identifier for products and can be scanned by smartphone apps
- Consumers will have supporting tools to help them make responsible choices upon purchase, as they are equipped with real time information about Heath, Nutrition, Environments, Human Rights…etc.
- Consumers can easily access other consumer' comments and opinions by using the barcode (ex: "what do you think of 321212456545 ?")

Smartphone apps with barcode scanner engine will be the "Product GPS" of tomorrow - will help consumers to make informed decisions upon purchase.

Health & Nutrition

Environment

Religion & Moral values

Human rights

Pricing & benchmark

Use case #2

Media

- Media publishes news about products, brands or industries can use barcode as a communication channel
- Searched by GTIN is a new way to access all the concerning published articles
- Different elements can be displayed when products are scanned, in addition to product data:
  - Published articles
  - Information of brands and manufacturers
  - Advice on product usage, related subjects,…
  - Ads (financial purpose)

Barcode (EAN / UPC) can be used as a new communication channel – 'product speaks for itself'

Use case #3

Supply Chain & Logistics

- Products are collected by volunteers of humanitarian organizations
- In face of challenges from Supply chain & Logistics (storage, transportation, distribution,…)
- French Government requires reporting of nutrition details
- Volunteers enter data into ERP system (using barcode hand scanner):
  - "Collect" data (what, where, when, how much,…)
  - Product data (name, brand, weight,…)
- Insufficient manpower to input all the details of product data
- Need to use international system for Product Classification

Volunteers should only have to deal with scanning GTIN codes.

Use case #4

Selective sorting of waste

- Each city has different specific rules for sorting of waste based on the capacity of the local recycling factories
- Sorting rules are based on the contents or different types of packaging materials (Carton, Plastic, PET,…)
- If the selective waste sorting is not done properly, all the containers could be rejected
- People need to have real time information to apply the local sorting rules.

Smartphone will propose apps soon to support proper sorting.