



European Data Forum 2015

November 16-17, 2015, Luxembourg

New Frontiers for Official Statistics

Mariana KOTZEVA – EUROSTAT, Deputy Director General

Key issues

1. A dynamically changing data ecosystem
2. Official Statistics in the age of big data
3. Statistical Office of the future: *Smart Statistics?*



**Statistics are the mirror through
which we view society**

David Hand @ Royal Statistical Society 2010

Key issues

1. A dynamically changing data ecosystem
2. Official Statistics in the age of big data
3. Statistical Office of the future: *Smart Statistics?*



African proverb

when the music changes so does the dance

If we fail to listen we will be out of step
[Denise Lievesley]



European
Commission

Datafication

Digital footprint

Sensors



European
Commission

What Happens in an Internet Minute?

1,572,877 GB of global IP data transferred¹

10 Million
ads displayed²

347,222
Tweets³

3.3 Million
pieces of
content shared⁴

6.9 Million
messages sent⁴



Netflix + Youtube =
more than 1/2 of
all traffic⁵

438,801
Wiki page views⁷

\$400 Million
during Alibaba
peak day sales⁶

10 Million
WeChat messages at its peak⁹

34.7 Million
instant messages
(MIM) sent⁸

194,064
app downloads¹⁰

\$133,436
in sales¹¹

31,773
hours of
music played¹²

38,194
photos
uploaded¹³

57,870
page views¹⁴

100
hours of video
uploaded¹⁶

4.1 Million
searches¹⁵

138,889
hours of video
watched¹⁶

23,148
hours of video
watched¹⁷

And Future Growth is Staggering



By 2017, mobile
traffic will have grown
13X in just
5 years¹



In 2017, there will be
3X more connected devices
than people on Earth¹

All digital data created reached
4 zettabytes in 2013¹⁸

Proclamation of pope Benedict 2005



Proclamation of pope Francis 2013





European
Commission

What Frontiers for Official Statistics?



What frontiers for official statistics (in the age of big data)? -1

- Data sources:** many, different, new, promising...
- Analysis:** limitless possibilities, nearly everything
- Methodology:** new horizons, revisiting, blending
- Privacy and ethics:** time for action
- Quality assurance:** a new function - certification?
- Statistical products and services:** data brokers?
- Standards:** a new and complex world
- Borders of correlation and causality:** revisit
- Research:** limitless possibilities

What frontiers for official statistics (in the age of big data)? -2

- Skills:** multidisciplinary in practice
- Evidence-based policy vs policy-based evidence:** risky territory

-.....

In the age of big data, official statistics and their frontiers are in transition. Eventually, the point of equilibrium will be set with the deployment of *Smart Statistics*, that is Statistics in a full-fledged IoT environment.

Key issues

1. A dynamically changing data ecosystem
2. Official Statistics in the age of big data
3. Statistical Office of the future: *Smart Statistics?*

The response of the European Statistical System:

- [Scheveningen Memorandum September 2013](#)
 - [Big Data Action Plan and Roadmap 1.0 Sept. 2014](#)
 - ↳ ESS Pilots 2016 - 2019
- Implementation of [ESS Vision 2020](#)

Areas in Big data roadmap

Policy

**Quality
framework**

Skills

**Experience
sharing**

Legislation

**IT
Infrastructures**

Methods

**Ethics /
Communication**

Pilots

Areas in Big data roadmap



Challenges

- cooperation, sharing of know-how
- development of a sound methodology ("from design-based to model-based approach")
- exploration & tentative implementation

Action (example)

- **Pilot projects, carried out by the Member States**
[2015-2019] Large scale partnerships

Areas in Big data roadmap



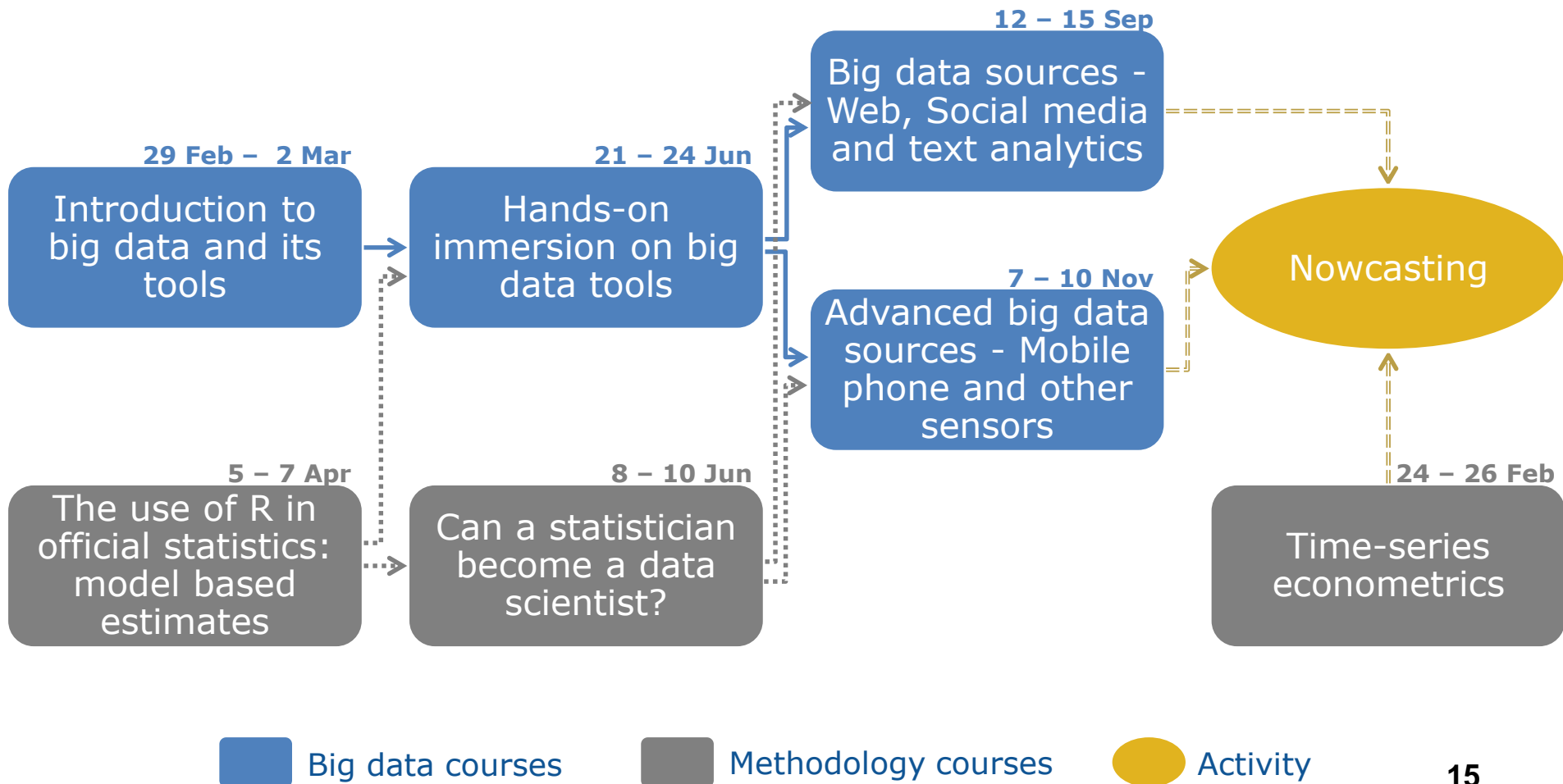
Challenges

- new skills for NSI staff: statisticians vs. data scientists ?
- computing capacity, hardware ?
- analytical tools, software?
- storage ?

Action (example)

- **Training program for European statisticians (ESTP)**
- Nearly 30 days of expert training in 2016 alone

ESTP courses supporting big data (2016)



Areas in Big data roadmap



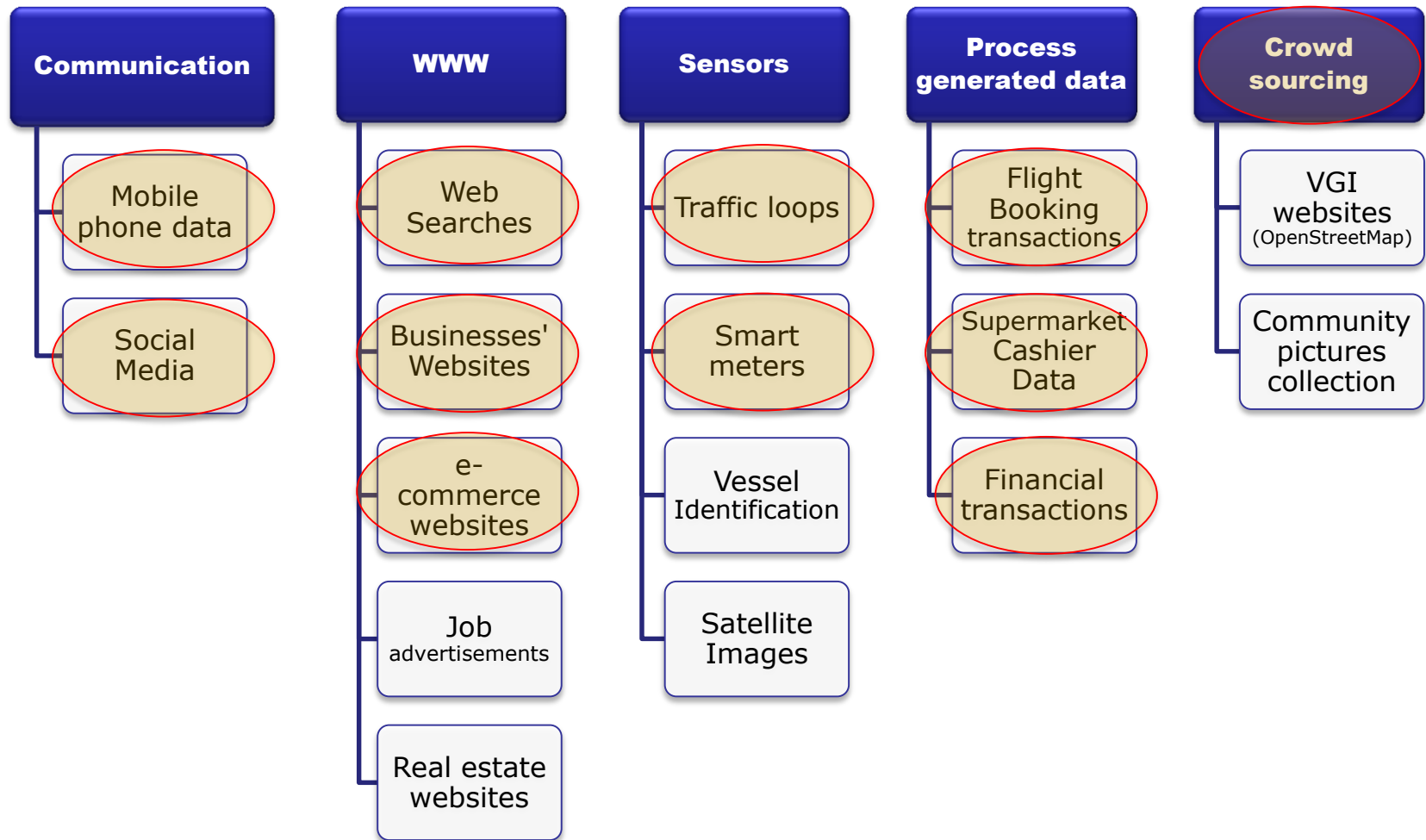
Challenges

- transversal challenges to all big data activities: **quality** and **ethics & communication**
- big data vs. statistics : "goodness of fit" (concepts, representativeness,...)
- impact on the public opinion of privacy and security concerns ?

Action (example)

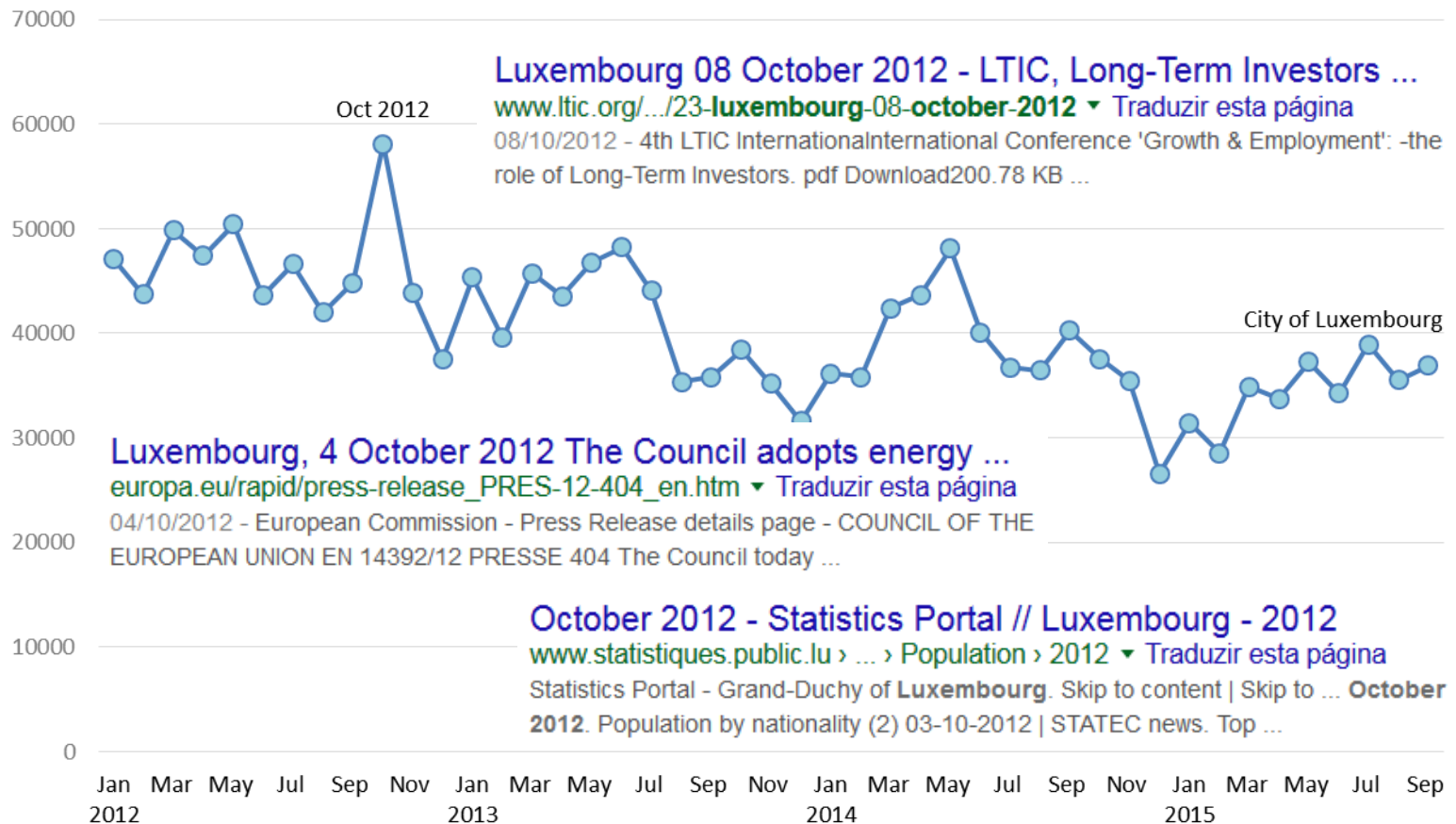
- Develop a quality framework for big data
- Propose ethical guidelines and communication (but also legislation and a data strategy)

Big variety in big data sources



Insights about World Heritage from Wikipedia page views

Page views for Luxembourg City in 30 languages (Jan12-Sep15)



Insights about World Heritage from Wikipedia page views

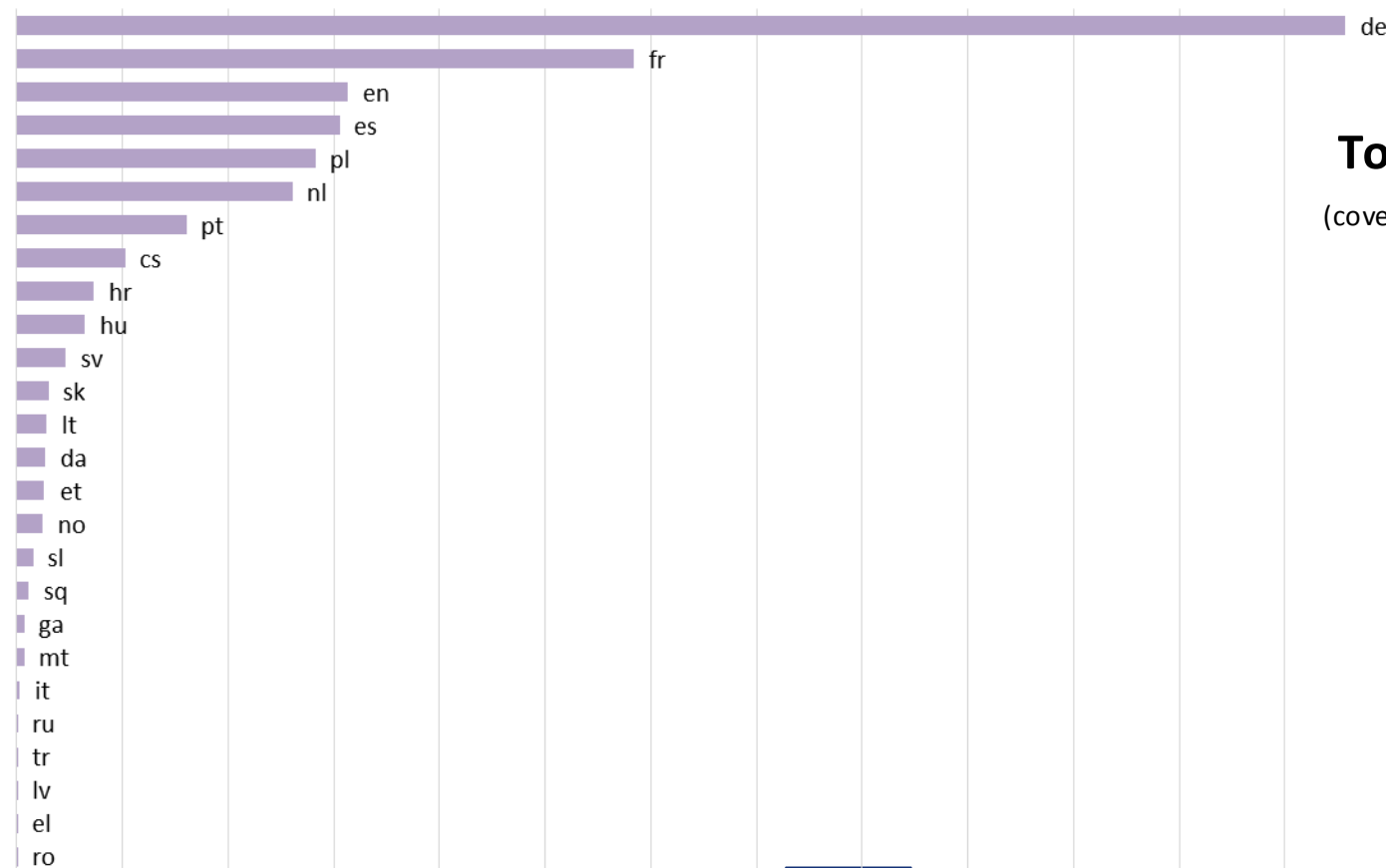
Page views for Luxembourg City in 30 languages (Jan12-Sep15)



Insights about World Heritage from Wikipedia page views

Page views for Luxembourg City in 30 languages (Jan12 - Sep15, by language)

0 50000 100000 150000 200000 250000 300000 350000 400000 450000 500000 550000 600000



Top 8 languages

(covering 90% of page views)

de	35%
fr	16%
en	9%
es	8%
pl	8%
nl	7%
pt	4%
cs	3%

Statistical domains



- Tourism
- Employment
- Population
- Migration
- Balance of payments
- Regional and GIS
- Transport
- ICT usage
- Prices and inflation
- Land use
- Agriculture

BDCOMP

(Big Data for Official Statistics Competition)

- ✓ **Nowcasting economic indicators – EU and country level**
 - ✓ **Unemployment**
 - ✓ **HICP**
 - ✓ **Tourism**
 - ✓ **Retail Trade**
- ✓ **Registration deadline 10.1.2016**
- ✓ **First submissions due 20.1.2016**
- ✓ **Track winners present at NTTS 2017**
- ✓ **Open to everyone – you are welcome to participate!**

<http://www.cros-portal.eu/content/bdcomp>

Key issues

1. A dynamically changing data ecosystem
2. Official Statistics in the age of big data
3. Statistical Office of the future: *Smart Statistics?*

The statistical office of the future

- ✓ **Data flows** in addition to **surveys and censuses**
- ✓ Embedded in data flow – **statistics 'everywhere'**
- ✓ **Product designers** in addition to **data collection designers**
 - Statistical modelling will be a major activity
 - From descriptive indicators to nowcasting (and forecasting)
- ✓ **Trust** and **quality** will be key
- ✓ New role in teaching **digital literacy**
- ✓ **Accreditation** and **certification** in addition to pure production
- ✓ Address issues linked to quality & transparency, privacy & confidentiality, access to third party data sources & data sharing, scientific standards & methodology, professional ethics, skills, ...

The NSI of the future: Official Statistics in a full-fledged IoT world

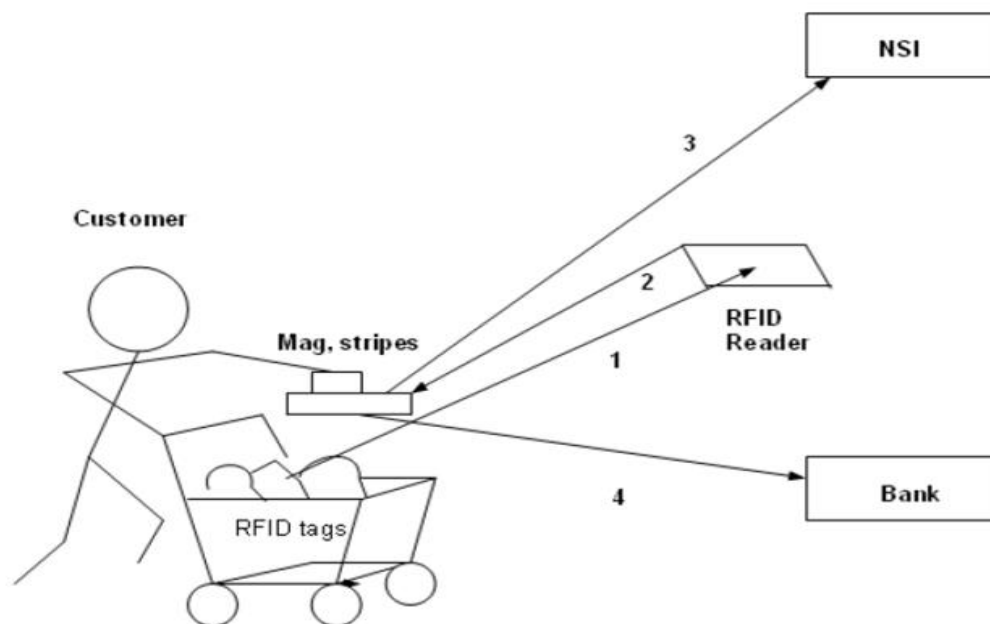


Figure 1: Shopping transactions

Svein Nordbotten

Concluding remark

- Frontiers in official statistics are in a dynamic transition phase.
- In designing the smart systems of tomorrow it is imperative to think of 'smart statistics'.
- Joint research in official statistics, big data and IoT will prepare the transition to smart statistics.

Thank you for your attention !

[Learn more about Eurostat's work on Big Data](#)