UPU postal big data and trade

UN GWG on Big Data

Use cases for data, services and applications: trade data lake

10 November 2017
Outline

Digital trade context and international postal exchanges
- International e-commerce growth through postal networks
- UPU postal tracking data capture and the adoption of digital customs declaration systems

4V postal data, international postal flows and trade
- Data on international postal shipments volumes and shifting digital trade patterns
- Data on value and categories of goods traded online and trade in low-value shipments

UPU big data platform technology
- UPU big data zoo
- Connecting to trade data lake

UPU big data governance and collaborations
- Collaborations and cooperation agreement frameworks
- Experience with UN Global Pulse
**Quantitative aspects**

**Volume**

*Scale of data*

Over 30 billion historical international postal tracking records by 2020, exponentially growing since 2010 – 3 billion in 2016 alone

**Velocity**

*Analysis of streaming data*

Hourly updates of millions of potential predicted delivery times and events along the international postal supply chain

**Variety**

*Different forms of data sources*

Geo-located data (GIS standards, geocoded information), EDI postal tracking data (postal data standards), customs data, and aviation/transportation data

**Veracity**

*Uncertainty of data*

Data reconciliation between all tracking data and sample-based QoS measurement system data (e.g. GMS), geographic coverage of different international postal data exchange systems

**Qualitative aspects**
### Tonnage

<table>
<thead>
<tr>
<th>Year</th>
<th>Total (kg)</th>
<th>Letter-Post (kg)</th>
<th>Parcel-Post (kg)</th>
<th>EMS (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>457'579'770</td>
<td>163'106'100</td>
<td>168'253'513</td>
<td>126'220'157</td>
</tr>
<tr>
<td>2014</td>
<td>566'648'650</td>
<td>213'610'159</td>
<td>188'587'382</td>
<td>164'451'109</td>
</tr>
<tr>
<td>2015</td>
<td>657'681'986</td>
<td>219'095'693</td>
<td>220'296'032</td>
<td>218'338'735</td>
</tr>
<tr>
<td>2016</td>
<td>742'928'846</td>
<td>332'658'171</td>
<td>203'304'623</td>
<td>209'583'027</td>
</tr>
</tbody>
</table>

### Volume

<table>
<thead>
<tr>
<th>Year</th>
<th>Total (kg)</th>
<th>Letter-Post (kg)</th>
<th>Parcel-Post (kg)</th>
<th>EMS (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>396'747'127</td>
<td>318'548'512</td>
<td>38'290'757</td>
<td>39'907'858</td>
</tr>
<tr>
<td>2014</td>
<td>460'525'763</td>
<td>373'012'947</td>
<td>35'961'086</td>
<td>51'551'730</td>
</tr>
<tr>
<td>2015</td>
<td>569'431'727</td>
<td>461'646'446</td>
<td>41'040'849</td>
<td>67'380'432</td>
</tr>
<tr>
<td>2016</td>
<td>689'771'865</td>
<td>580'826'534</td>
<td>40'920'615</td>
<td>68'435'558</td>
</tr>
</tbody>
</table>

### Change (%)

<table>
<thead>
<tr>
<th>Year</th>
<th>Tonnage Change</th>
<th>Volume Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016/15</td>
<td>13.0%</td>
<td>21.1%</td>
</tr>
<tr>
<td>2015/14</td>
<td>16.1%</td>
<td>23.6%</td>
</tr>
<tr>
<td>2014/13</td>
<td>23.8%</td>
<td>23.8%</td>
</tr>
<tr>
<td>2011 (from/to)</td>
<td>Developed countries</td>
<td>Africa</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Developed countries</td>
<td>46.3</td>
<td>2.4</td>
</tr>
<tr>
<td>Africa</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Asia and Oceania</td>
<td>21.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>1.7</td>
<td>0</td>
</tr>
<tr>
<td>Transition economies</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>World</td>
<td>70.8</td>
<td>2.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2016 (from/to)</th>
<th>Developed countries</th>
<th>Africa</th>
<th>Asia and Oceania</th>
<th>Latin America and Caribbean</th>
<th>Transition economies</th>
<th>World</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed countries</td>
<td>26.3</td>
<td>0.9</td>
<td>20.8</td>
<td>2.7</td>
<td>2.4</td>
<td>53.1</td>
</tr>
<tr>
<td>Africa</td>
<td>0.7</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>1.1</td>
</tr>
<tr>
<td>Asia and Oceania</td>
<td>33.2</td>
<td>0.4</td>
<td>4.2</td>
<td>1.4</td>
<td>4.0</td>
<td>43.2</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>1.0</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0</td>
<td>1.3</td>
</tr>
<tr>
<td>Transition economies</td>
<td>0.7</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.5</td>
<td>1.3</td>
</tr>
<tr>
<td>World</td>
<td>61.9</td>
<td>1.5</td>
<td>25.4</td>
<td>4.3</td>
<td>6.9</td>
<td>100</td>
</tr>
</tbody>
</table>

International deliveries (tonnage) of small packets, parcels and packages, 2011 and 2016, distribution of regional flows as a share of global flows, percent

Source: Universal Postal Union.
From postal volumes to international trade

Currently: 379,301,707 electronic customs declarations recorded in our UPU big data platform

285,952,139 with value less 50 (in any currency), i.e. 75% of recorded declared postal shipments
UPU Big Data Platform and Systems

- Tracking messages: personal data
- Tracking messages: transactional data
- Tracking messages: operational data
- UPU postal statistics questionnaires
- Key UPU surveys

UPU surveys

UPU Big Data Platform and Systems

Key UPU surveys

UPU postal statistics questionnaires

Tracking messages: operational data

Tracking messages: transactional data

Tracking messages: personal data
Products & services
Part that is visible to the end-user

Processes, systems and technologies
Part that is visible to the producers
import org.apache.spark.sql.functions.explode
val nm = testTriniti.select(testTriniti.col("id"), explode(testTriniti.col("value")))
nm.createOrReplaceTempView("nm")

import org.apache.spark.sql.functions.explode
me: org.apache.spark.sql.DataFrame = [id: string, key: string ... 1 more field]

spark2

ns.printSchema

root
|-- id: string (nullable = true)
|-- key: string (nullable = false)
|-- value: string (nullable = true)

Filtering the exploded DataFrame for a specific key

df2 = sqlContext.sql("select cast(value as double) from nm where key = 'cp/d.val/amt'")
df2.count

df2: org.apache.spark.sql.DataFrame = [values: double]
res11: long = 379301727

spark2

val df2 = sqlContext.sql("select cast(value as double) from nm where key = 'cp/d.val/amt'")
df2.count

df2: org.apache.spark.sql.DataFrame = [values: double]
res11: long = 379301727

spark2

val df2 = sqlContext.sql("select cast(value as double) from nm where key = 'cp/d.val/amt'")
df2.count

df2: org.apache.spark.sql.DataFrame = [values: double]
res11: long = 379301727

spark2

val df2 = sqlContext.sql("select cast(value as double) from nm where key = 'cp/d.val/amt'")
df2.count

df2: org.apache.spark.sql.DataFrame = [values: double]
res11: long = 379301727

spark2

val df2 = sqlContext.sql("select cast(value as double) from nm where key = 'cp/d.val/amt'")
df2.count

df2: org.apache.spark.sql.DataFrame = [values: double]
res11: long = 379301727
Thank you