TDM STORIES #1

ADRIANA HOMOLOVA, DATA JOURNALIST

As a data journalist, the bulk of my work consists in collecting, analysing and visualising data for journalistic stories. Text and Data Mining comes in handy in my work, because by writing a script I can read all the information I need using a computer doing it at a much faster pace than if I used the naked eye, and have it analysed for patterns and correlations.

I developed a project called "Elvis". It is a tool I created to visualise public spending through tenders in EU countries. The ultimate objective is to help journalists reveal potential corruption cases. It allows journalists to examine preferences of public institutions when choosing suppliers, discover which companies sell services to governments, as well as dig into public contract details and money flows. Most of the datasets we use for this tool are freely available on most government websites.

In this age of information, we, journalists, rely on Text and Data Mining to break news. It is important that we areable to pull together and

make sense of all the data available out there. If we had to comb through thousands of websites or documents by hand, we would never have been able to uncover stories like the Panama Papers, the Lux Leaks, or stories of tennis match-fixing.

The issue today is that journalists are currently not covered by the TDM exception, as we are not research organisations, neither non-for-profit organisations. Limiting our ability to use TDM in our activities would damage investigative journalism in Europe.

In democratic societies, investigative journalism holds government and institutions private accountable the public. Technological to developments have been making our job as investigative journalists more efficient, and more valuable. That is why it is of the utmost importance that journalists can continue to use TDM technologies with legal certainty, and be able to provide citizens with the knowledge they need to make informed decisions.