

# The Story Behind the Big Clean

Submitted on 09 Oct 2012 by

On November 3rd, at the National Technical Library in Prague, the 2012 Big Clean will take place. Big Clean? Yes, it is a spring cleaning for open government data! So that more data becomes better usable throughout the PSI re-use lifecycle. How did the Big Clean come about, and what is in store for the 2012 edition in Prague?

The idea behind the [Big Clean](#) was first put forward by [Antti Poikola](#) during the [Open Government Data Camp 2010](#). At the core, the idea of the Big Clean was a plan for a screen-scraping hackathon focused on local public sector data. The event was framed as a big Spring cleaning sweeping the dusty corners in the public sector. However, the focus of the event was ultimately broader, involving the whole life-cycle of reuse of public sector data, including data acquisition and cleaning, data analysis, information visualization and data-driven journalism.



The [last year's Big Clean](#) was a successful event, considering its DIY nature and no funding. It attracted 80 participants overall, coming from a range of different backgrounds spanning from developers to journalists. And what is more, it provided a venue in which members of different organizations working on similar topics can meet. I believe that the event contributed to the integration of the overlapping efforts in the domain of public sector data, minimizing the time spent on duplicating the previously unknown work done by others instead of reusing it directly.

Fortunately, the Big Clean attracted attention of funders as well. For the upcoming Big Clean we have received funding from Google Community Grants of [Tides Foundation](#). The grant allows us to set up a better and more international Big Clean for the year 2012, by supporting invited speakers and opening the conference up for a broader audience by using English as the conference language. While this is a significant change when compared with the last year's Big Clean, the events still forms a clear succession from the previous one, holding on to the core topics of screen-scraping, data refining and data-driven journalism.

The Big Clean 2012 will take place at the [National Technical Library](#) in Prague, Czech Republic, on Saturday November 3rd. Thanks to the funding received we are able to make the event's admission free. More details about the conference can be found on its [web site](#).