

Unleashing Patent Data to Drive the IP Policy Change

Some Big Data sources are publicly available but not yet “unleashed”. An example is patent data: all patents are freely downloadable from the Web, and – still – no one seems to have full visibility on the patent world. As a result, companies are issuing patents on topics protected by patents of other companies, technologies are being developed without a prior assessment of ‘freedom to operate’, and patent wars burn billions of dollars and years of litigation.

We have built a global map of over 5 million patents issued in the USA over the last 39 years. The map reveals semantic relationships between patents: similar patents are close to each other on our map. It allows detecting communities, monitoring trends, and drilling down into any region of the patent world to understand local effects. We can visualize any region, and discover opportunities and vulnerabilities in any patent portfolio.

Most importantly, our patent map can answer questions regarding the structure of the American patent system itself. It is not a secret that the patent system is flawed but it has not been possible to quantify the flaws. We can do it now. For example, the US patent office requires each patent to cite related patents. We discovered that, for an average patent, only 14% of its citations are among 100 patents that are most semantically related to it. In fact, for 54% of patents, no single citation is among their 100 most related patents. If so many relevant patents are missed during the patent evaluation process, it is now clear that pretty much every patent is vulnerable and the protection it provides is questionable.

In this talk, we will introduce the audience to a series of revealing, and not particularly flattering statistics on the US patent system. We will make data-driven suggestions for the IP policy changes.

Speaker’s bio: Dr. Ron Bekkerman has recently joined the Faculty of Management (Business School) of the University of Haifa, Israel, as Assistant Professor of Data Science. In addition to his academic activities, Ron is co-founder of a stealth-stage startup, and a Data Science consultant. In 2013 Ron served as Chief Data Officer of Carmel Ventures – the leading Israeli Venture Capital fund. Previously, Ron was a founding member of the Data Science team at LinkedIn in Mountain View, CA. Prior to that, he was a Data Mining researcher at HP Labs, Palo Alto. Ron was among the organizers of KDD-2012 and KDD-2013. He served on program committees of KDD-2009 and KDD-2010. At KDD-2011, together with Misha Bilenko and John Langford, Ron gave an invited tutorial on Scaling up Machine Learning. At KDD-2013, Ron participated in a popular panel titled “A Data Scientist’s Guide to Making Money from Start-ups”. Ron holds his PhD in Computer Science from the University of Massachusetts, Amherst, and has authored a long list of publications in top-tier Data Mining venues.