OpenData500

GOVLAB

Report of Findings from an
Open Data Roundtable
with the
U.S. Patent & Trademark Office

GOVLAB

Report of Findings from an Open Data Roundtable with the U.S. Patent & Trademark Office

A dialogue for feedback on ways to improve the patent data system.

April 2015

Joel Gurin
Audrey Ariss
Katherine Garcia
Laura Manley

Table of Contents

Foreword

Today, data is widely recognized as not just an important tool, but a platform for innovation, and this Administration has taken many important steps towards making government data more open and accessible to the public. As Secretary Pritzker has said, the Department of Commerce is "America's Data Agency," and has a unique and central role in unlocking the potential of government data.

For the USPTO, disclosing and disseminating data go to the very heart of our mission. After all, the patent system rests on the bargain an inventor makes when he or she discloses an invention and teaches the public how it works, and in turn receives the right to exclude others from practicing it. From that perspective, the USPTO has been in the business of open data for a very long time. Our data lets inventors and businesses know when they need to invent around a technology and when they need to consider a license, which serves as fuel for our nation's innovation engine. If we are going to live up to our mission to disseminate information about patents and trademarks, it requires an agency-wide commitment to the principles of open data—a commitment that the USPTO has made.

As we redouble our efforts to make our data accessible in a useful format, we want to be responsive to our users' needs. We've taken this feedback seriously and moved ahead with many of the ideas including the creation of the first public GitHub account within the Department of Commerce to share design elements and the creation of the Club for Open Data Enthusiasts (C.O.D.E.) affinity group within the USPTO to bring together USPTO employees that are open data advocates and enthusiasts to showcase the power of open data in the community. We look forward to using the findings and suggestions from this report today, and in the months and years to come.

I'd like to thank Joel Gurin and The GovLab as well as the USPTO's Open Data team for putting on this event. I especially want to thank the attendees for sharing their thoughts for the first of what I hope will be many productive discussions between this agency and the diverse people and institutions that use our data.

Michelle K. Lee
Under Secretary of Commerce for Intellectual Property and Director of the USPTO

Introduction

In June 2014, The Governance Lab (GovLab) at NYU launched a series of Open Data Roundtables to bring together government data providers with data users in the private and non-profit sectors. The goal of the Roundtables has been to help agencies prioritize their work on open data more effectively, meet the needs of businesses and other data users more efficiently, and create social and economic value from the government's vast data resources.

The first [Open Data Roundtable](#) was held at the U.S. Department of Commerce on June 18th, 2014. Attendees at that event expressed an interest in providing continued feedback on the USPTO open data efforts. USPTO's data holds the history of American ingenuity, and the USPTO's mission includes data dissemination. However, the USPTO is faced with many challenges, ranging from legislative hurdles to legacy systems. The technology leadership of the USPTO is focused on the Office's long-term commitment to providing patent and trademark data as public resources.

On December 8th, 2014, The GovLab held a follow-up Open Data Roundtable event at the U.S. Patent and Trademark Office (USPTO) in Alexandria, VA focused on the USPTO's patent data. At the USPTO Roundtable, attendees heard both Michelle K. Lee, who was then Deputy Under Secretary of Commerce for Intellectual Property and Deputy Director of the USPTO, and Lynn Overmann, Deputy Chief Data Officer of the Department of Commerce (DOC), emphasize the agency's desire to more actively engage data users and technologists by involving the developer community and creating ongoing feedback loops and collaborations.

This Roundtable was attended by representatives from 14 companies and three academic institutions, the chairman of the Patent Information Users Group (PIUG), two Presidential Innovation Fellows from the U.S. Department of Energy, and twenty participants from USPTO and DOC. In addition to Michelle K. Lee and Lynn Overmann, the participants heard from Alan Marco, Chief Economist of the USPTO, and Thomas Beach, Senior Advisor at the Office of the Under Secretary and Director.

At the Roundtable, the USPTO presented a roadmap for open data initiatives designed to provide better access to USPTO's data assets and better engagement of the USPTO's community of data users. The main elements of this roadmap included:

Deliver: The USPTO will provide data and application programming interfaces (APIs) to the public.

▸ The publication of technical architecture, to include shared services, data formats, documentation, security/access measures, standards for APIs (application program interfaces), and collaboration tools;

▸ The release of an API for Public PAIR. The Patent Application Information Retrieval (PAIR) system provides its users a safe, simple, and secure way to retrieve and download documents and status information on publically available patent applications;

▸ The release of an API for Patent Assignments data; and

▸ Publication of a public-facing data inventory.

Engage: The USPTO will work with the community to do innovative things with our data and APIs.

▸ Organizing an Open Data Affinity Group at the USPTO;

▸ Hosting a series of roundtables, data jams and hackathons throughout 2015 and 2016; and

▸ Inaugurating a USPTO Datapalooza in 2016.

Collaborate: The USPTO will partner with the private sector to help shape the direction of open data at the USPTO.

▸ The development of a community governance plan and USPTO community partnerships;

In addition to its Open Data Roadmap, the USPTO presented demonstrations of three data projects at the Roundtable:

i. PatentsView. This prototype patent visualization tool is intended to allow users to explore inventors and their assignees, patent classes, and locations for over 5 million U.S. patents. Its three aspects are aimed at engaging all types of users: an API for developers, a query tool for economic and legal research, and visualization for general users. Starting with the U.S. Department of Agriculture (USDA), PatentsView is working with federal agencies to track federal funds going to inventors.

ii. Electronic Data Hosting. This prototype repository of public bulk patent and trademark data utilizes Swagger to represent a RESTful architecture solution to supplement the current file-based approach to accessing bulk data. It allows users to use APIs to search data and download customized datasets on both mobile and desktop devices.

iii. Assignment Search. This ongoing project provides a searchable database containing all recorded Patent Assignment information dating back to August 1980. Its data is updated weekly. It allows users to also 'add synonyms' to search companies using different names to register their patents.

This report is being released as a public document with the hope that it will encourage further input and dialogue. It encapsulates the exchange of information among the USPTO's individual data users, both internal and external, who participated in the Roundtable, and describes the open data priorities they have identified. The report hopes to be of value to the USPTO open data efforts; officials in other government agencies working to improve their agencies' open data capabilities; individuals and organizations who use the USPTO's patent data; and interested members of the media and the public.

Roundtable Feedback and Suggestions

The Roundtable discussions covered three key areas where the USPTO can further develop its open data strategy and increase the potential to leverage patent data: user engagement, data access and availability, and data quality and interoperability. Beyond the value to be gained from feedback and collaboration among open data stakeholders, The GovLab recognizes the need for continued research and mapping of the open data ecosystem.

These areas for improvement mirror the findings of The GovLab's earlier [Roundtable with the U.S. Department of Commerce](#), and may apply to other federal agencies as well. That earlier Roundtable identified seven areas of improvement: (i) treating data users as customers by engaging with them and getting their input and feedback on a regular basis; improving data (ii) discoverability and findability, (iii) access and availability, (iv) quality, and (v) interoperability; and new strategies to (vi) collect and (vii) disseminate data. The USPTO's 'in flight' initiatives (though in beta) will go a long way towards making it easier to find and discover patent data, and achieve wider data dissemination. In the case of the USPTO's data, the data collection processes are inseparable from data collection issues, and these concerns are treated as a single area for improvement.

The subsequent sections are a reporting of the individual feedback and suggestions expressed by attendees on the three key areas where the USPTO can further develop its open data efforts.

The USPTO's Current Data Assets

Patent Information Products (daily/weekly /monthly/cumulatively)

▸ Patent Grants – Images/Text/Bibliography with selected information from 1790 to present

▸ Patent Applications – Images/Text/Bibliography from 2001 to present

▸ Patent Assignment/Classification Products

▸ Patent Official Gazette (every Tuesday)

▸ Public PAIR Data – highly detailed patent information from 1980 to present

Trademark Information Products (daily/weekly /monthly/cumulatively)

▸ Trademark Registration Images from 1870 to present

▸ Trademark Application Images from 2010 to present and text from 1884 to present

Patent Technology Monitoring Reports

▸ Patent Summary Reports by demographics, geocoding and domain categories

---

Data User as Customer

The USPTO, like other federal agencies, is beginning to think of its data stakeholders – both current and potential – as "customers" for their data. This concept is the basis for developing a more user-focused approach to data collection and dissemination. In the context of patent data, stakeholders include not only users of patent data but also users of the patent system: inventors, patent experts and attorneys, patent examiners, and the USPTO itself.

Goal

To better identify patent data stakeholders, establish communities of interest and feedback loops, and increase the number and depth of USPTO's interactions with its data users.

Desired impact

To help the community of patent users become better informed, connected, and engaged.

Feedback and Suggestions

▸ Develop an outreach plan for all patent stakeholders to include: inventors, the public, data vendors, known USPTO data users and intermediaries, law firms and patent attorneys and patent examiners, and internal USPTO groups. Individual suggestions for an outreach plan included:

- Collecting user information held internally. This can include identifying patent users from requests and emails received by individual members of the USPTO, and creating a listserv of external patent data users.

- Forming 'affinity groups' of USPTO staff and external stakeholders; including a group for interested examiners and USPTO personnel that also provides a bridge with volunteer technical experts.

- Engaging patent experts within industry and academia.

- Expanding USPTO staff attendance at conferences.

- Identifying USPTO datasets for different kinds of communities to engage with on a geographical and industry basis. Patent data on selected sectors – such as energy, automotive and others – would be of interest to different innovation communities around the country.

- Building a developer community, perhaps a virtual community, that is specific to USPTO's data. Elsewhere, the DOC is starting to build a developer relations team, and the Census Bureau has created a developer community of its own.

- Providing internal training for USPTO's data users.

- Creating a GitHub account for the USPTO.

- Exploring cooperative research and development agreements (CRADAs) – agreements between a government agency and a private company to work together on specific research and development projects.[1]

- Obtaining public feedback on proposals, with data vendors participating earlier in data quality and standardization efforts, allowing these stakeholders to test and give input on initiatives.

▸ Articulate the value of the information held in USPTO's data for non-technical users, and add more context to make it usable by non-experts. For instance, lawyers are rarely technological experts, while inventors may be intimidated by legal jargon.

▸ Provide a public-facing point of contact for all data queries: a data librarian, data concierge, or data "contentist".

▸ Provide a point of contact or communication channels for civic tech volunteer groups and civic tech entrepreneurs.

▸ Align events such as hackathons, data jams, and datapaloozas[2] with open data releases and/or tools. For instance, hackathons can be timed to follow beta launches of APIs as an opportunity for coders, developers and designers to build apps or visualizations using the API. Similarly, data jams could be used as a part of agile development within open data efforts, for example as a means to pilot data cleaning efforts. It might be useful to engage not only data users but also other users of the patent system – lawyers, inventors, large and small firms providing data – to see how they would react to different applications and uses of patent data.

▸ Better promote events through the USPTO's network of patent and trademark institutions, the Patent Information Users Group (PIUG), 18F resources at General Services Administration (GSA), the USPTO's website, an internal listserv, the Inventor's Eye newsletter, and other means.

▸ Explore further opportunities for patent examiners to interact with the public. While the public may challenge patent applications by submitting examples of prior art voluntarily, Section 122(e) of Title 35 of the U.S. Code, Section 122(c) prohibits the USPTO from encouraging such challenges to applications.[3] In practice, this affects examiners who can get technical information on prior art from Ask Patents,[4] but cannot look at arguments from the people submitting the prior art.

▸ Look for ways to provide additional feedback mechanisms across all UPSTO open data efforts.

▸ Provide measurements of success for open data initiatives.

---

1   USGS: CRADA. *http://www.usgs.gov/tech-transfer/what-crada.html*
2   Hackathons are events where programmers, technologists, software developers and designers collaboratively build solutions over a short period of time, usually one or two days. Data jams assemble innovators and entrepreneurs from the government, nonprofits, educational institutions and the private sector in small groups to brainstorm products, services and technological tools that could solve national challenges using open data, and be created within 90 days. Datapaloozas are events that focus on demonstrating and highlighting the uses of open data to address social problems.
3   The USPTO Manual of Patent Examining Procedure, Chapter 1100, Section 1134, Third Party Inquiries and Correspondence in a Published Application. *http://www.uspto.gov/web/offices/pac/mpep/s1134.html*.
4   Ask Patents, StackExchange. *http://patents.stackexchange.com*.

Data Access and Availability

Data repositories use multiple formats and languages to help users access data, which may be stored in a variety of different and sometimes incompatible formats. Roundtable participants highlighted overwhelming priorities for information dissemination: access to bulk Public PAIR and CPC data, data on the status and expiration of patents, and increased notification of changes affecting data classification, access, and availability. The first of these – access to public PAIR data – was addressed by the USPTO at the opening of the Roundtable, and validated by participants as addressing one of biggest access challenges faced by users of patent data.

Goal

To provide better access to more USPTO data, and better information on policies and procedures affecting quality, access and availability of USPTO data.

Desired Impact

To increase both the number of users of and ease of access to USPTO data in an egalitarian manner.

Feedback and Suggestions

▸ Provide daily updates and bulk access to Public PAIR data[5] to better facilitate statistical analysis of patent data. To facilitate faster access, participants suggested providing an option to access Public PAIR without the 'file wrappers'. These are the written record of negotiations between an applicant and the office preceding the issuance or rejection of a patent, and are frequently in PDF/image formats. Concurrently, for those who desire easier image search and comparisons of patent design, USPTO can provide image file wrappers in a machine-readable format, integrated as part of the API.

▸ Map Cooperative Patent Classifications (CPC) codes to Master Classification Files (MCF) from 1976 to the

---

5   Public PAIR data includes:

Application data – all of the bibliographic data such as application number, status, publication number, examiner name, class/subclass, inventor, and title among other fields;

Transaction history – a list of transactions during the prosecution as well as the date on which they occurred;

Image file wrapper – the file wrapper contains PDF images of the various stages of prosecution (ie. applicant submitted claims, fee worksheets, non-final rejections, list of references, issue information etc.);

Patent term adjustments – this gives the history of patent term adjustments petitioned and received, although no individual documents are available for public download;

Continuity data – if the record has any parent or child continuity data, the document numbers is listed (and linked, if possible);

Fees – This tab generates a pop up window with the data needed to inquire about whether fees are up to date (from another part of the USPTO website);

Published documents – If there are related published documents they are listed here with links to view the full-text and image of said documents as well as access their respective PAIR entries;

Address & attorney/agent – Correspondence address for the attorney prosecuting the case as well as the other attorney/agents involved

(Summary from The Intellogist: *https://intellogist.wordpress.com/2010/08/03/using-uspto-public-pair-part-1*.)

present. The U.S. Patent Grant MCF contains classification information on all patent grants issued by the USPTO, and is available bimonthly.[6]

▶ Provide USPTO's users with more information about the Office's open data initiatives, including any potential ramifications of changes in policy for data users. Suggestions included:

- Providing information and notice on any changes in USPTO policies and practices or related legislation. Users would like advance notice of policy initiatives coming up and how they will affect the delivery of USPTO's data. For instance, when changing options for accessing data, the USPTO should inform users of the availability of old links to the data; any retagging (and inclusion) of older data in new data products; the formats in which bulk data is available via APIs (JSON or XML); and the availability of permalinks to the patents themselves within the new products and APIs.

- Providing more public notice of data releases.

- Clarifying how USPTO open data initiatives relate to and are affected by legislative reform, international initiatives, other USPTO initiatives such as crowdsourcing and Patent Trial and Appeal Boards (PTAB) trials.[7]

- Providing information on the process of patent examination in the interest of transparency.

- Addressing concerns about whether USPTO data will be permanently available and accessible. For example, perhaps USPTO could provide a service-level agreement (SLA) that data will be available over time.

- Providing as much information as possible for developers, such as a schedule for when files are added or deleted.

- Providing more granularity and updated information on processes, implementation, and implications for users – including information on the patent process and the examining process.

- Providing information on serial numbers and their periodic recycling, including records of the data structures. This data is a lengthy time series, including many changes, but there is limited available documentation of the evolution of serial numbers.

- Providing the North American Industry Classification System (NAICS) code on Patents.

▶ Explore the possibility of making more information on patents granted via government contracts through iEdison (Interagency Edison) – the interface for grantees and contractors to interact with any participating agency. iEdison helps government grantees and contractors comply with the Bayh-Dole Act requiring that government-funded inventions be reported to the federal agency that made the award.[8]

▶ Provide data on the legal status and enforceability of applications: whether an application is in progress/provisional, litigated, approved, reinstated or abandoned/expired. If litigated, users would like to know the outcome

---

6   The CPC effort is a joint effort between the USPTO and the European Patent Office (EPO) where the Offices have agreed to harmonize their existing classification systems (European Classification (ECLA) and United States Patent Classification (USPC) respectively) and migrate towards an internationally compatible classification system. This common scheme applies to technical documents, in particular patent publications, used by both offices in the patent granting process.

7   PTAB Trials, U.S. Patent and Trademark Office, *https://ptabtrials.uspto.gov*.

8   iEdison. *https://public.era.nih.gov/iedison*.

of the process.

- Provide bulk access also to rejected patent applications. Participants noted a desire for more transparency in patent examination to enable statistical analysis, for example, on the impact of case law or legislation on the admissibility of patents.

- Provide certifications of corrections of patents when these are made.

- Explore opportunities for moving to new systems at USPTO, with easy, discrete changes. For instance, integrating File Forms AO120 from local district courts would provide information on whether a patent has been filed in a suit.[9] Such changes can be implemented on a 'moving forward' basis, while later going back to integrate historical data.

---

9   When a civil action is filed regarding a patent or trademark, the U.S. Patent and Trademark Office (USPTO) must be notified via the submission of a form entitled Report on the Filing or Determination of an Action Regarding a Patent or Trademark (AO 120). Plaintiff's counsel is required to fill out and submit this form at the time the complaint is filed.

Data Quality and Interoperability

At the Roundtable, data users emphasized the need for cleaner and more complete, valid, and accurate data. The USPTO recognizes that its data sometimes contains omissions, ambiguities, and inaccuracies.

The USPTO is faced with three significant challenges in improving the quality of its data. First, statutory requirements for the submission of patents make fixes difficult when errors are detected after these have become legal documents. Second, patent applicants are often reticent to provide unambiguous information: there are market incentives to obfuscate data about patent ownership. Third, the process of data cleaning is, in itself, a somewhat subjective one.

An additional challenge is the multiplicity of data models within the USPTO, and internationally. This prevents data interoperability, the ability of users to compare U.S. patent data across datasets and with patent data from around the world.

Goal

To make the USPTO's data more complete, representative, accurate, and interoperable.

Desired Impact

To improve the usability of the USPTO's data and reduce the amount of work necessary to clean data for use, including allowing users to better and more easily combine different datasets and variables. This makes it possible for users to analyze and visualize information and develop actionable insights in new, more efficient ways.

Feedback and Suggestions

▸ Address errors *early* in the patent application process. The USPTO could give an applicant real-time feedback on his or her application, with an incentive for providing correct information such as a faster processing time (or a refusal to process the application unless the information provided is accurate.) An API could provide instant feedback regarding errors at the time of filing, ranging from data placed in the wrong fields to errors in federal contract numbers, location, etc.

▸ Explore creating a marketplace or crowdsourcing approach for *finding* errors. Spelling errors, information entered in the wrong fields, missing information and other inconsistencies, prevalent throughout USPTO data, making it difficult to harmonize.

▸ Explore a marketplace also for *fixing* errors. A stop-gap way to improve data quality could be to mirror the data and store a copy of the data in the cloud, updating it periodically (weekly or daily), and correcting it for errors. The USPTO could publish corrections (or a certificate of corrections) in machine-readable format.

▸ Support a legislative change from paper-based or PDF systems to an e-filing system for patent applications: a requirement that they be filed in an all-digital text-searchable format. If the patent bar opposes this change, e-filing could be promoted as a voluntary option for patent applicants. Identify ways to get public/industry support for this.

▸ Address accuracy issues arising from reclassification of USPTO patents to CPC classifications. While Roundtable participants did not object to the CPC system, they were concerned about the importance of mapping the new classification to old information accurately. The USPTO believes the public can be of help, for instance by using crowdsourcing to verify the accuracy of CPC classifications, or by contributing to a glossary project to help translate old classifications to new one.

▸ Provide a more current registry of active patent attorneys and agents and identify incentives for attorneys to update these records.

▸ Improve the quality of pre-1970 image data and full-text backfiles.

▸ Look at ways to improve bibliographic data.

▸ Investigate the legislative barriers or other concerns that prevent using potential third-party solutions to data problems. For example, a private company could theoretically clean patent data and donate the improved dataset to the USPTO. But if a government agency were to receive such a 'gift' of data, it could give the appearance of the USPTO endorsing a company or favoring it over its competitors.

▸ Normalize formats for data from different countries/states (E.g. Canada vs. California).

Appendix A: About the Open Data Roundtable

AGENDA

Monday, December 8, 2014

USPTO Madison Building – GIPA, Paris/Berne Room

9:00 AM **Registration**

9:30 AM **Welcome & Goals of the Day**
Michelle K. Lee
Deputy Under Secretary of Commerce for Intellectual Property and Deputy Director-U.S. Patent and Trademark Office

Lynn Overmann
Deputy Chief Data Officer - U.S. Department of Commerce

9:45 AM **Structure of the Day**
Joel Gurin
Project Director - Open Data 500, Senior Advisor - The GovLab

10:00 AM **Presentation of USPTO's Open Data Roadmap and Project Demonstrations**
Alan Marco
Chief Economist - U.S. Patent and Trademark Office

10:40 AM **Breakout Session 1: Understanding USPTO's Data Customers**

12:10 PM **Lunch**

1:00 PM **Breakout Session 2: Responding to the Open Data Roadmap**

2:20 PM **Break**

2:40 PM **Breakout Session 3: Brainstorming Solutions**

4:00 PM   **Commitments and Next Steps**
                 Thomas Beach
                 Senior Advisor - U.S. Patent and Trademark Office

4:20 PM   **Closing**
                 Joel Gurin
                 Project Director - Open Data 500, Senior Advisor - The GovLab

4:30 PM   **Adjourn**

Participants

GOVERNMENT AGENCIES AND OFFICES

U.S. Department of Commerce (DOC)	The U.S. Department of Commerce promotes job creation, economic growth, sustainable development and improved standards of living for all Americans by working in partnership with businesses, universities, communities and the nation's workers. Representatives: **David Langdon** – *Economist and Senior Policy Advisor* **Lynn Overmann** – *Deputy Chief Data Officer* **Rand Ruggieri** – *Deputy Digital Strategist*
U.S. Patent and Trademark Office (USPTO)	The United States Patent and Trademark Office is the federal agency for granting U.S. patents and registering trademarks. Representatives: **Grayling Achiu** – *Project Management Lead* **Thomas Beach** – *Senior Advisor* **Scott Beliveau** – *Supervisory Patent Examiner, Technology Center 2400* **David Chiles** – *Chief Technology Officer* **Tony Chiles** – *Deputy Chief Information Officer* **Rajeev Dolas** – *Portfolio Manager, Trademark Next Generation* **Greg Gabel** – *Director, Application Architecture Division* **Janet Gongola** – *Senior Advisor* **Jack Harvey** – *Director, Technology Center 2800* **Jeremy Idol** – *Director, Corporate Systems Division* **Ajay Kundaria** – *Senior Advisor* **Michelle K. Lee** – *Deputy Under Secretary of Commerce for Intellectual Property and Deputy Director* **Chris Leithiser** – *Computer Scientist, Information Products Division* **Alan Marco** – *Chief Economist* **Amanda Myers** – *Economist, Office of the Chief Economist* **Bob Vanni** – *Program Manager, Dissemination and PTAB* **Christopher Wong** – *Presidential Innovation Fellow*

Economics and Statistics Administration (ESA)	The ESA plays three key roles within the Department of Commerce (DOC). ESA provides timely economic analysis, disseminates national economic indicators and oversees the U.S. Census Bureau (Census) and the Bureau of Economic Analysis (BEA). In this latter role, ESA works closely with the leadership at BEA and Census on high priority management, budget, employment, and risk management issues, integrating the work of these agencies with the priorities and requirements of the Department of Commerce and other government entities. Representative: **Austin Durrer** – *Chief of Staff*
U.S. Department of Energy	The mission of the U.S. Department of Energy is to ensure America's security and prosperity by addressing its energy, environmental and nuclear challenges through transformative science and technology solutions. Representatives: **Timothy Jones** – *Presidential Innovation Fellow* **Clarence Wardell** – *Presidential Innovation Fellow*

COMPANIES, NONPROFIT, ACADEMIC AND RESEARCH ORGANIZATIONS	
Amazon Web Services	Amazon provides cloud computing services through Amazon Web Services to a range of clients. Representatives: **Cindy Brent** – *Federal Sales Manager* **Jed Sundwall** – *Open Data Technical Business Development Manager*
American Institutes for Research (AIR)	AIR's mission is to conduct and apply the best behavioral and social science research and evaluation towards improving peoples' lives, with a special emphasis on the disadvantaged. Representative: **Rebecca Rosen** – *Principal Researcher*
Code for Northern VA	Northern Virginia locals helping Arlington County, Alexandria City and other Northern Virginia municipalities open their data and create meaningful web and mobile apps. Representative: **Michelle Koeth** – *Brigade Captain*

Collective IP	Collective IP is a comprehensive business intelligence platform for discovering and analyzing technologies emerging from universities, companies and inventors. Representative: **Will Butler** – *VP, Software Engineering*
Computer Packages Inc (CPI)	Computer Packages provides Patent and Trademark Intellectual Property Management Systems and Patent Annuity Payment Services. CPI helps reduce the cost of patent, trademark and annuity management. Representative: **Linda Wright** – *Programming Director*
Enigma	Enigma is a platform that centralizes, mines and relates big public data about companies, people and locations. It draws together sources as diverse as SEC filings, government spending contracts, liens, patents, asset ownership, bills of lading, and more. Representative: **Marc DaCosta** – *Co-founder*
Georgia Institute of Technology	Georgia Tech is a science and technology-focused learning institute renowned for a deeply-held commitment to improving the human condition. Representative: **Stuart Graham** – *Professor*
Google	Google's mission is to organize the world's information and to make it universally accessible and useful. Representative: **Frank Italiano** – *Patent Counsel*
IBM	IBM is an international technology and consulting company. Representative: **Manny Schecter** – *Chief Patent Counsel*
IFI CLAIMS Patent Services	IFI's databases include chemical and biological indexing, US patent assignee name standardization, and 'probable' assignee name additions on US published application records, calculated expiration dates, and class code updates on all patent records. IFI also offers comprehensive US patent legal status information, including ownership changes. Representative: **Mike Baycroft** – *CEO*
IP.com	IP.com is an intellectual property solutions provider that helps organizations make complex business decisions around their innovations Representative: **Brian McGlynn** – *VP Strategy*

The Patent Board	The Patent Board helps clients unlock the value of patent assets to increase performance, improve profits, gain knowledge, identify opportunities, research competitors, and pursue the right investments Representative: **Karl Wilhelm** – *CEO*
Patent Information Users Group (PIUG)	The objective of PIUG is to create a forum to discuss issues and concerns relevant to the patent searcher community. The group promotes the importance of intellectual property, especially patent publications, and most specifically, the computerized patent databases. Representative: **Elliott Linder** – *Chair*
Reed Technology	Reed Technology, a LexisNexis company, is a leader in electronic content management services, including data capture and conversion, e-submission and publication. Representative: **Dave Abbott** – *VP Strategy*
Thomson Reuters	Thomson Reuters is a leading source of information for businesses and professionals. Representative: **Rosalind Cheslock** – *Senior Product Development Manager*
Trea	Trea is a service where people can analyze relationships between innovation stakeholders and map ideas to the unified knowledge graph. Representative: **Max Yuan** – *Founder*
University of Missouri School of Law	A national leader in the field of dispute resolution, University of Missouri School of Law seeks to complement a strong traditional curriculum with an orientation toward lawyering as a problem-solving endeavor. Representative: **Dennis Crouch** – *Professor*
Way Better Patents	Way Better Patents chronicles the latest developments in intellectual property, innovation, inventions and patents (I3P) and the emerging innovation economy. Representative: **Arleen Zank** – *President*

OPEN DATA ROUNDTABLE SERIES SPONSORS

Amazon Web Services offers a broad set of global compute, storage, database, analytics, application, and deployment services that help both large enterprises and start-ups move faster, lower IT costs, and scale applications. Amazon Web Services Worldwide Public Sector is helping government and education customers employ cloud services to reduce costs, drive efficiencies, and increase innovation across the globe. Public Sector organizations of all sizes use AWS to build applications, host websites, harness big data, store information, conduct research, improve online access for citizens, and more. *For more information, see aws.amazon.com/gov.*

PricewaterhouseCoopers is a network of firms in 158 countries that delivers quality in assurance, tax and advisory services. PWC helps federal agencies meet the challenge of integrating financial and performance systems to help make decisions and improve accountability. Clients look to the PricewaterhouseCoopers Public Sector Practice to bring direct hands-on knowledge of federal standards for systems, internal controls, and financial reporting. The Practice assists clients through creating interactive data, developing an agile reporting and analytic framework, and identifying and implementing improvements to the data and information supply chain. *To find out more, visit www.pwc.com/publicsector.*

Appendix B: Glossary of Acronyms

API   Application Program Interface

CPC   Cooperative Patent Classification

CRADA   Cooperative Research and Development Agreement

ECLA   European Classification

EPO   European Patent Office

MCF   Master Classification File

NAICS   North American Industry Classification Standards

PAIR   Patent Application Information Retrieval

USPC   United States Patent Classification

USPTO   U.S. Patent and Trademark Office