Looking for Prior Art

(Non-Patent-Literature)

In the Internet Haystack

Today's Moderators

1. Steve Pederson – *CEO & Chairman*
2. Will Bushee – *V.P. of Development*
3. Troy Mentzer – C*hief Navigator...trolling the deepest and darkest the Internet has to offer...finding relevant content for the great unwashed.*

stevep@harvestIP.com / +1 (612) 840-6888
willb@harvestIP.com / +1 (605) 212-6396
troym@harvestIP.com / +1 (605) 261-0450

Agenda

1. About HarvestIP
2. What is the problem?
3. What is the Deep Web?
4. What Validity Search & Google Misses
5. Ten Things Missing from Search
6. How HarvestIP Harvests the Deep Web
7. Deep Web Report Examples
8. Q&A

*Harvesting **Deep Data** from the Public Web, whether it is "known", "unknown" or "hidden".*

*Providing normalized, **Relevant Content** for analysis in the Legal Community.*

The Company

HarvestIP, LLC is the legal industry's leader in harvesting high quality content from and the conventional Surface Web (Google) and the more authoritative, but largely inaccessible ***Deep Web*** *sources ("unknown and hidden" data.)*

HarvestIP was formed to serve the legal community building on the Deep Web Federated Harvesting technologies developed over the past 10 years by BrightPlanet Corporation.

Focusing on the Legal Community

The IP research market is growing rapidly.

Finding unstructured, Non-Patent-Literature (NPL) on the web is fast becoming THE key essential "missing link" for IP researchers.

HarvestIP is in a position to quickly become the leading "real-time" research resource for litigators, analysts, and researchers from the Public Web.

A Proven History

With over 10 years of Deep Web extraction expertise, and four years of direct experience working with U.S. Intelligence Agencies, HarvestIP's parent technology company, BrightPlanet Corporation, has achieved a strong reputation and is acknowledged as the resource for Deep Web Harvesting in the Intelligence Community (IC) on behalf of the U.S. Government's War on Terror.

HarvestIP has been charged with providing these same Deep Data Harvesting resources to the Legal Community - successfully integrating with ""best of breed" analytic and visualization tools to provide highly relevant content for the legal community from the Deep Web .

harvestIP

Three Problems

1. *Data & Documents, both (structured and unstructured) on the Open Source Public Web and on private/proprietary Databases are growing exponentially…at the "scale of the Internet".*
2. *Current search technologies (like Google) simply do not have the scope or capacity to find all "**unknown and hidden**" **Deep Data** on the web.*
3. *Current search technologies return popular links – they do not provide content, nor can they federate and normalize it into a form that is **qualified, relevant, and useable** to the projects at hand.*

harvestIP

The Solution

To find data that is
"known", "unknown", "protected" or "hidden",
we must first rely on a technology that does not simply crawl data, but can Harvest, Federate, Normalize and Qualify data.

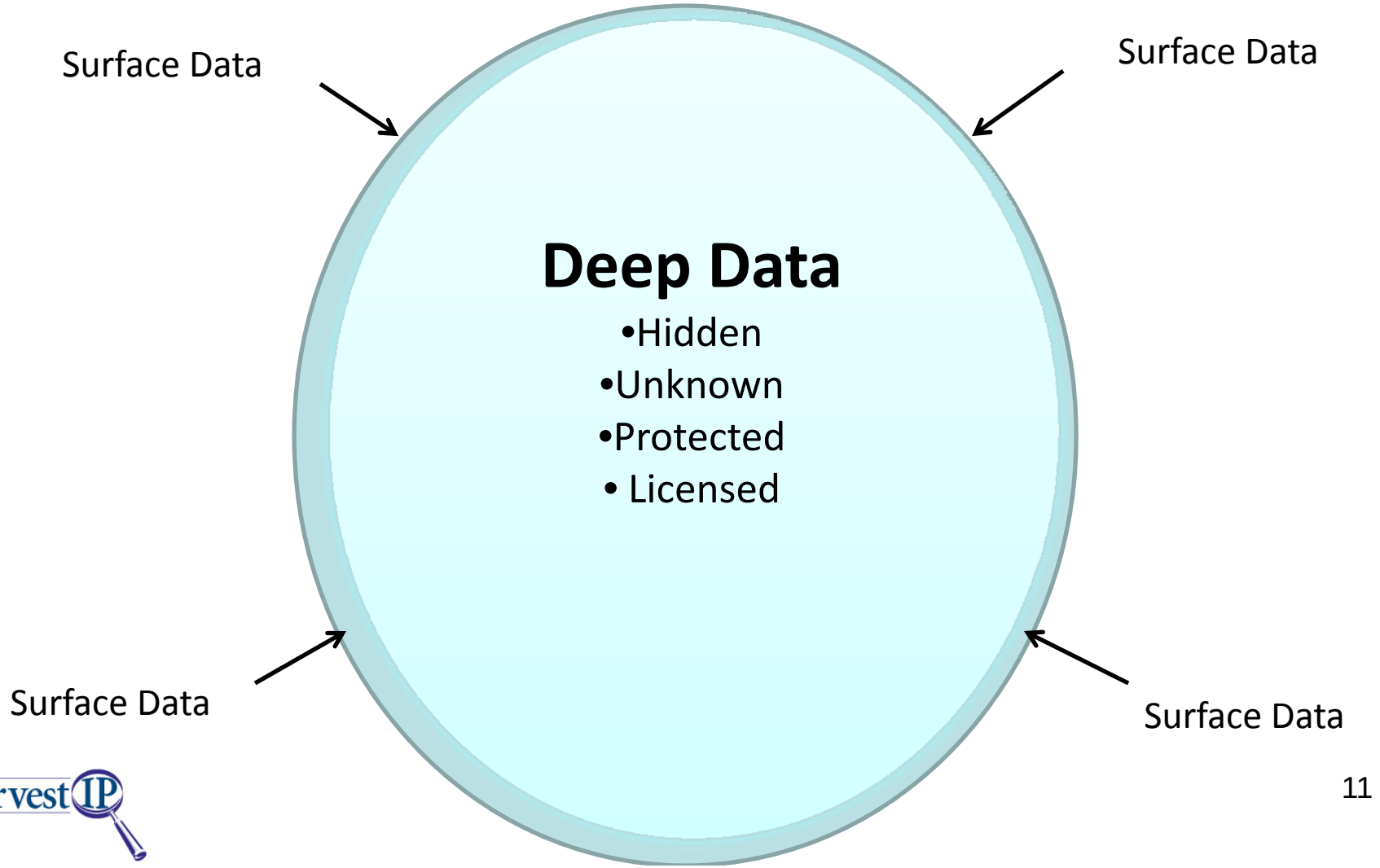
HarvestIP harvests and delivers fully qualified, relevant Deep Data to analysts and analytic tools from the
Open Source Public Web, Proprietary and Private sources.

Surface Web vs. Deep Web
Surface Data vs. Deep Data
Search-Crawl-Mine vs. Harvest

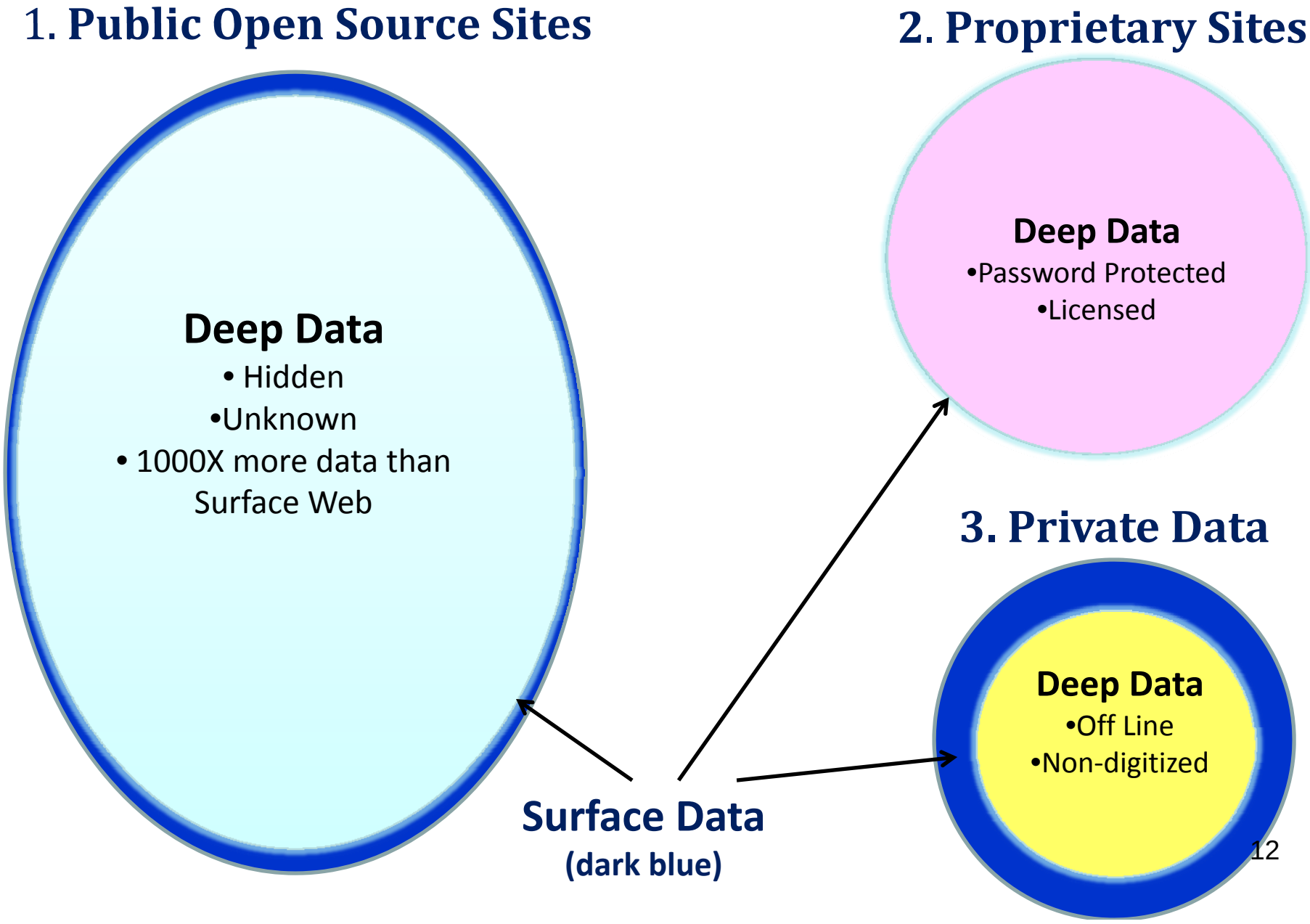
Exploring a Deep Web that Google Can't Grasp
New York Times, February 2009
(http://www.nytimes.com/2009/02/23/technology/internet/23search.html)

harvestIP

Surface Data: *Accessed by link traversal via Google, Yahoo!, etc.*
Deep Data: *Automatically accessed by HarvestIP Technologies*



3 Sources of ***Surface*** Data & ***Deep*** Data



Not Google…

Beyond "Surface Search", Crawl, & Mining

Google, Yahoo, Bing, Cuil *et al*: *(<u>Surface</u> <u>Search</u>)*

- Single user searches ONLY through one search engine or index
- Only one query can be issued at a time
- Content may be out-of-date
- Content may not exist within the index
- Often too much information is returned
- Often the valuable information is "hidden"

What is Validity Search Missing?

Validity Search utilizes static databases that cover patents only (such as the USPTO, Questel and Delphion) . They offer a subscription to "known" data sources and have little or no access to Non-Patent Literature (unstructured Deep Data) assigned to the Web.

Deep Data on the web represents an increasingly growing repository of documents and data now generally regarded as essential to the legal diligence process in today's world.

The HarvestIP Solution

(…at the scale of the Internet)

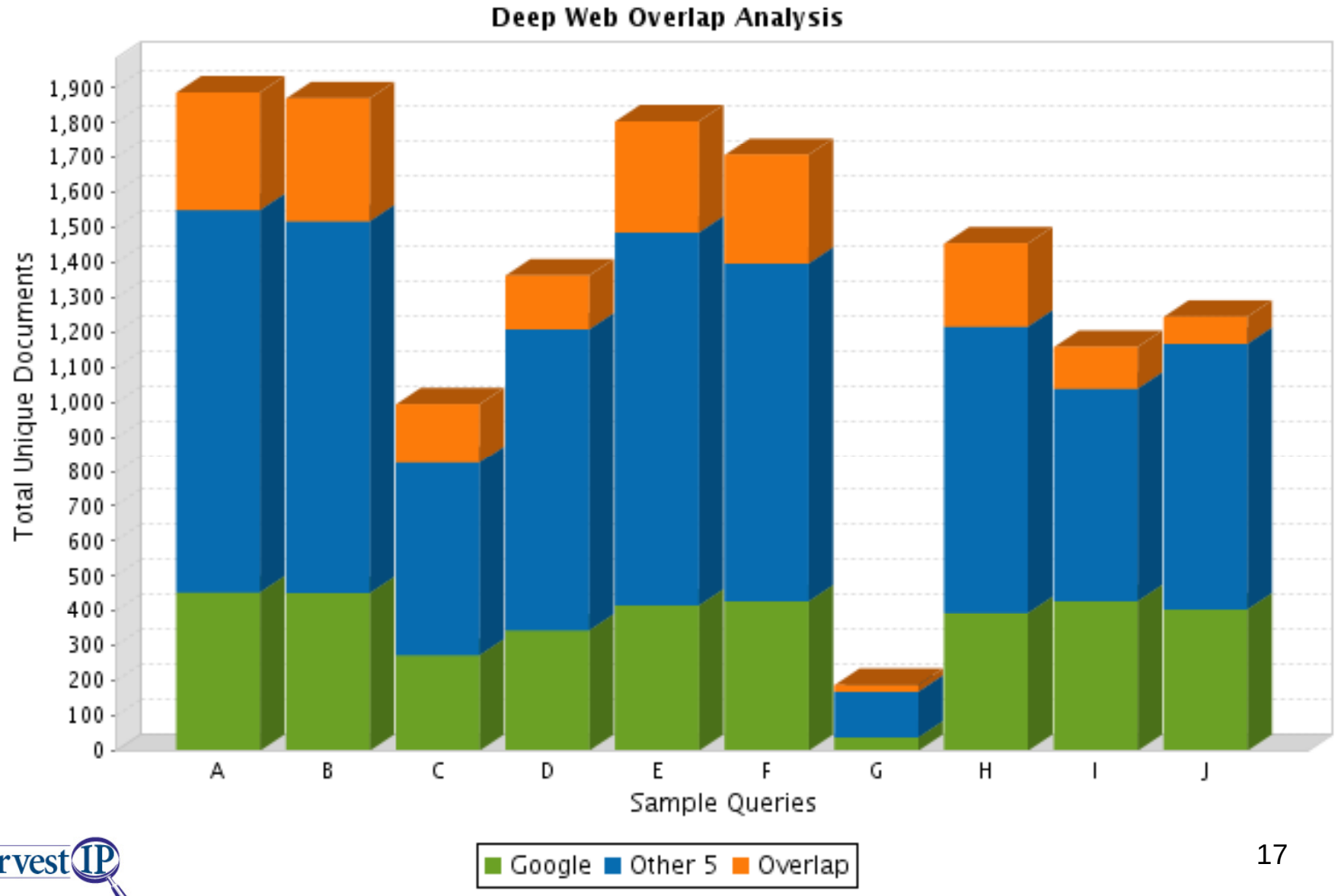
1. *HarvestIP **Harvests Documents** from the Deep Web and provides content not found through Surface Search technology.*
2. *HarvestIP uses a patented, custom configurable deep query technology that allows the harvesting of **"Relevant Content"** - documents search engines like Google cannot find.*
3. *HarvestIP directs the custom query manager to **Automatically Harvest "beyond the search box",** employing patented scripting technologies that find "unknown and hidden" Deep Data.*
4. *HarvestIP **Federates and Normalizes Content** so that it becomes relevant and purposeful to the analytic projects at hand.*

harvestIP

Top Ten Things…

1. *Engines have 30B to 1T indexed…just the surface*
2. *Does "The Biggest" mean "The Best"?*
3. *Popular yes, but did I find what I was looking for?*
4. *The results depend on how you ask the questions*
5. *Different Sites…Different Results…Different Rules*
6. *Now we see it… Now we don't*
7. *The trouble finding a good date*
8. *No such thing as a "Simple Search"*
9. *Results 1-10 of about 15,000,000*
10. *How do you know what you're missing?*

What Surface Search is Missing



Surface Search *(Google)*
vs.
Deep Web Harvest *(HarvestIP)*

An example...

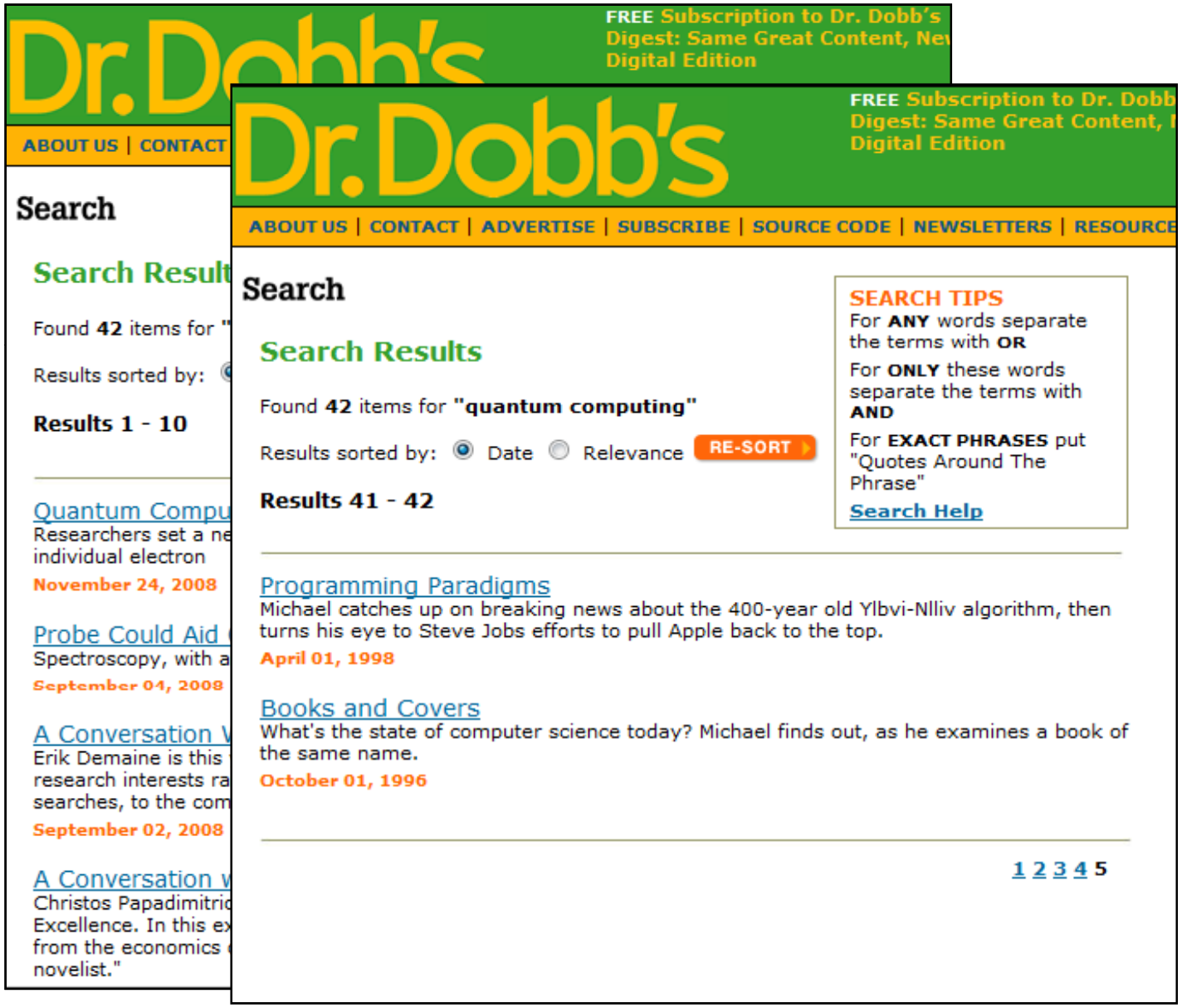
harvestIP

Google *(Surface Web)* vs. Deep Web

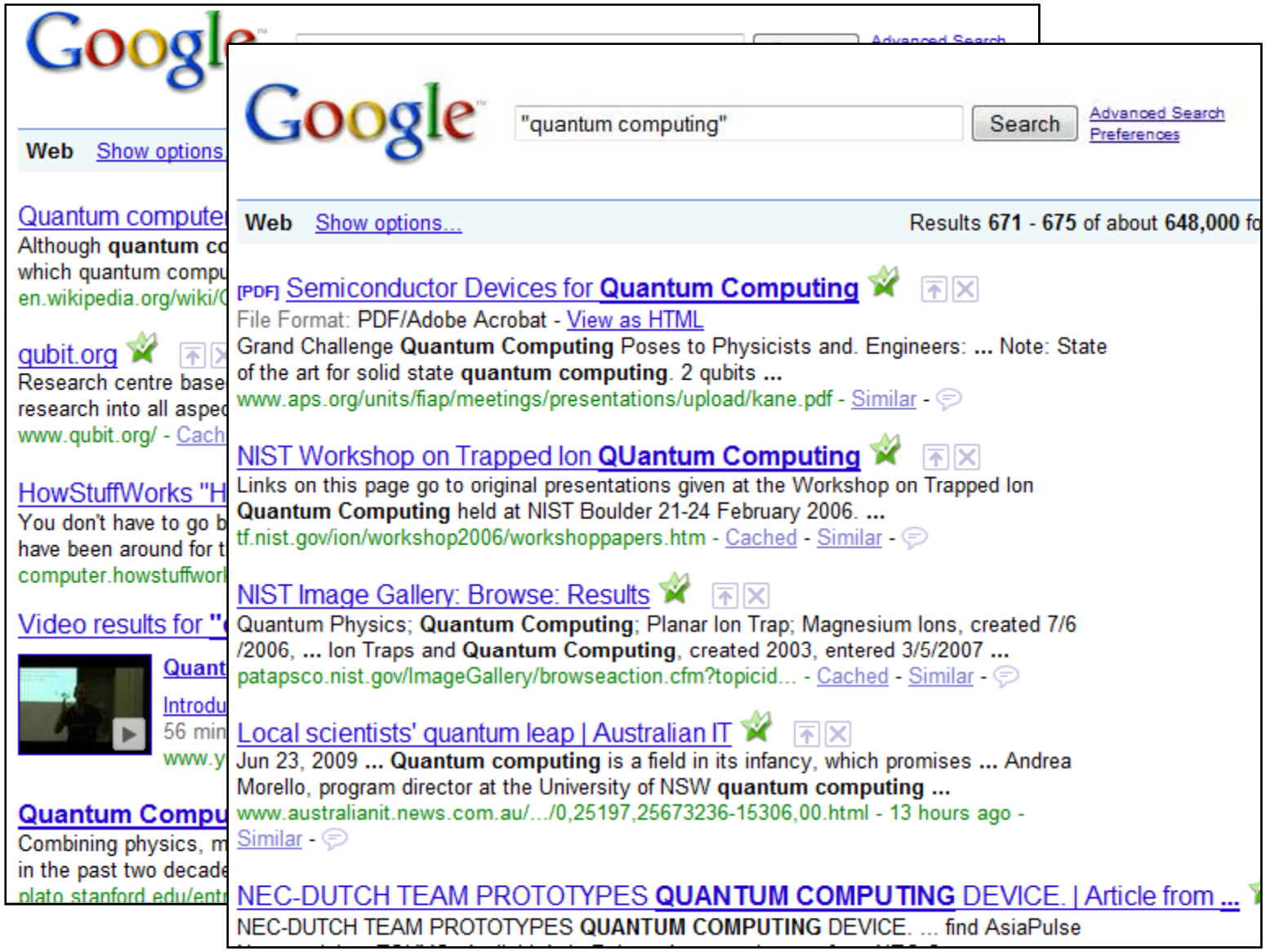
Case Study: Searching for "Quantum Computing"

- *A Search of Google with the query "quantum computing" returns "about 648,000" but the actual documents available are 675.*
- *A Deep Web Harvest of the Dr. Dobb's site with the identical query returns 42 documents, with all 42 available.*
- *A direct comparison of the two result sets showed no overlap from Google's result set to Dr. Dobb's result set.*
- ***Conclusion:** A researcher would not have found any of the documents contained within the Dr. Dobb's database by doing only a Google search. ("What other content was not available in the Google results for this query?")*

Google *(Surface Web)* vs. Deep Web



Google *(Surface Web)* vs. Deep Web



Google *"Search"* vs. Deep Web Report

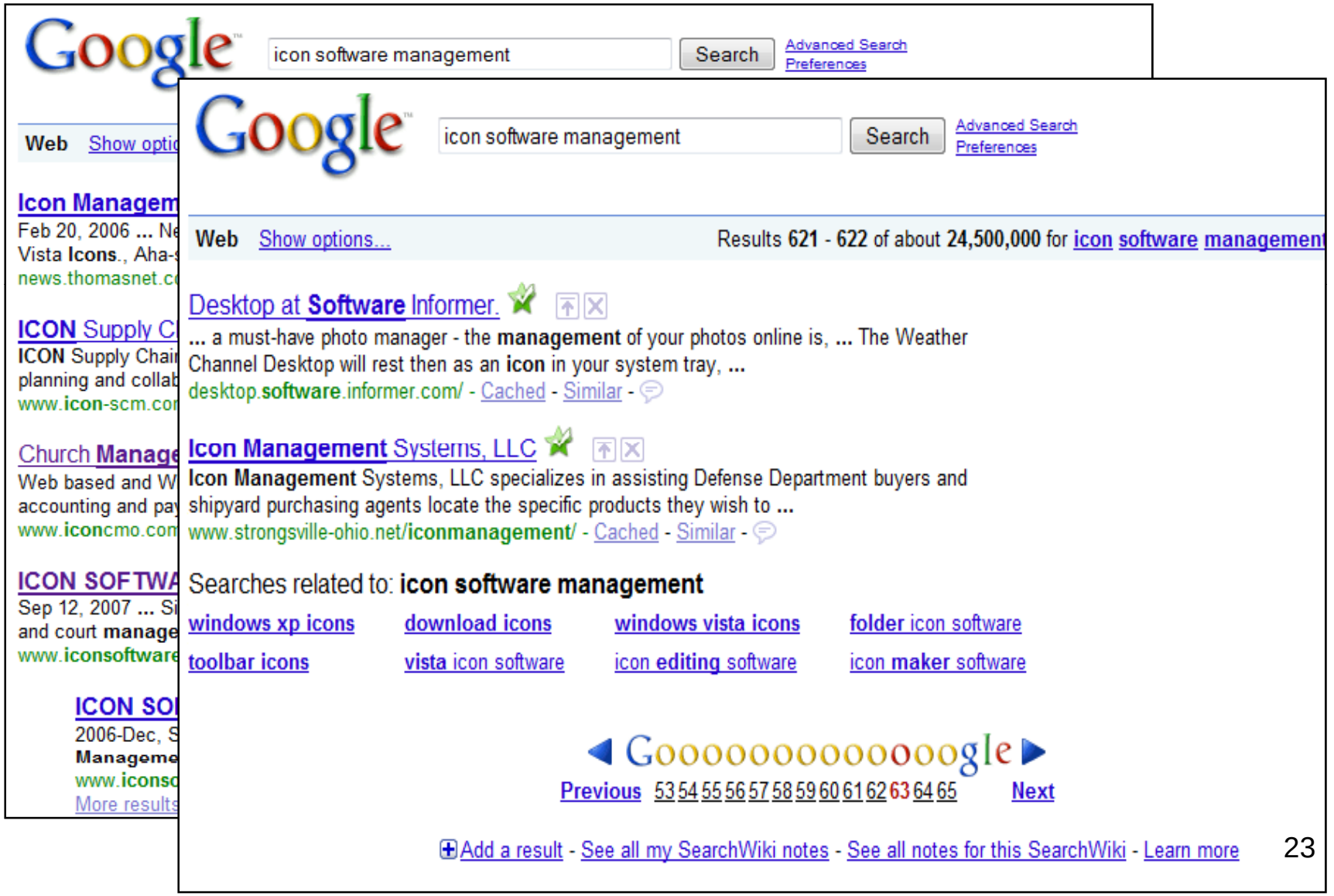
Example: Swartz GUI Patent Group - Prior Art Search

- Google Search
  - Search only one source with one query
  - Searching for "icon software management" returns "about 24,500,000 hits"
  - Only 62w can be viewed

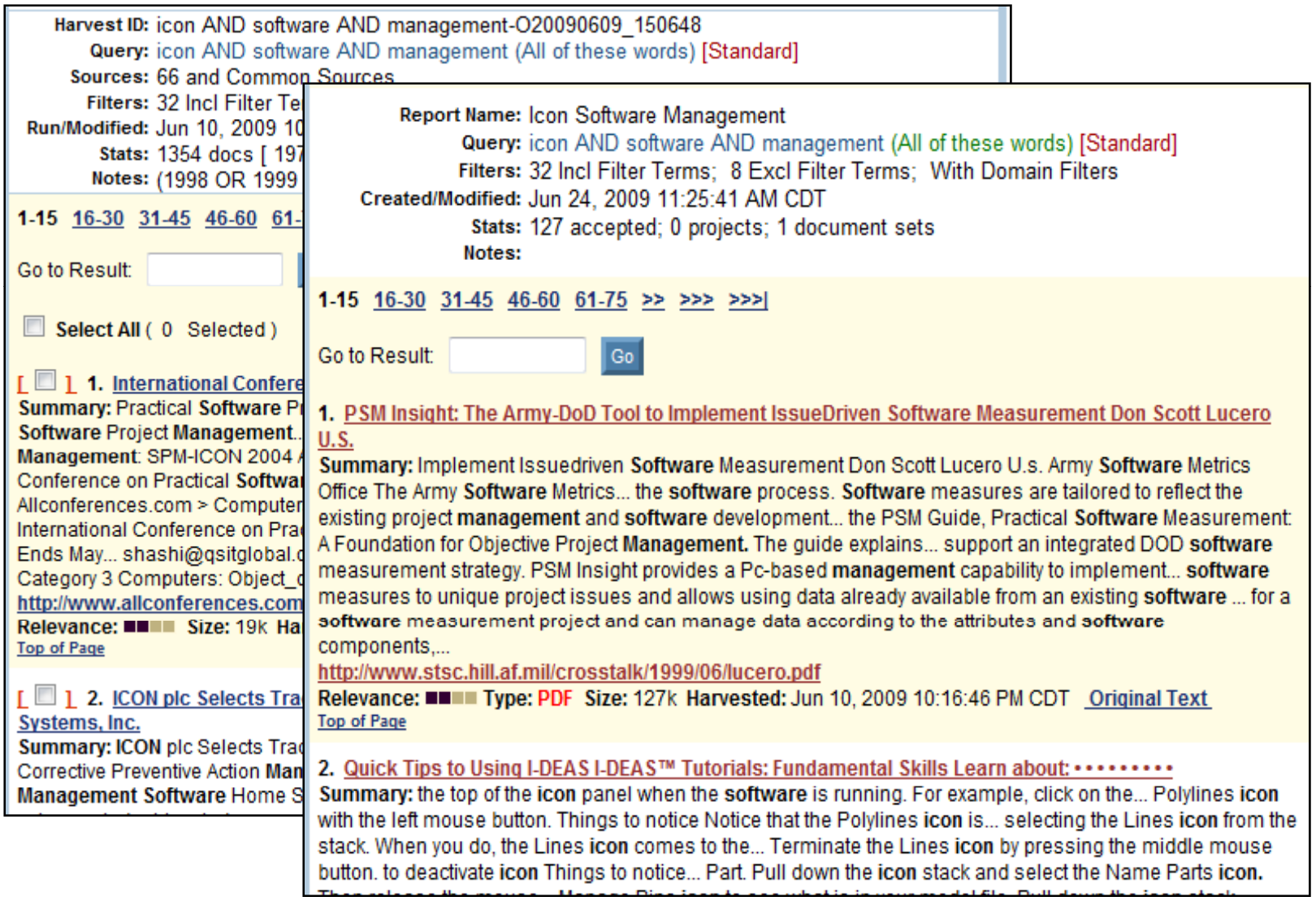
- HarvestIP Deep Web Report
  - Harvest from many sources
  - Searching for "icon AND software AND management" to 65+ Deep Web sources returns 1,354 qualified results
  - All are viewable and subset searchable
  - Further filtering to 127 highly qualified results

harvestIP

Google *"Search"* vs. Deep Web Report



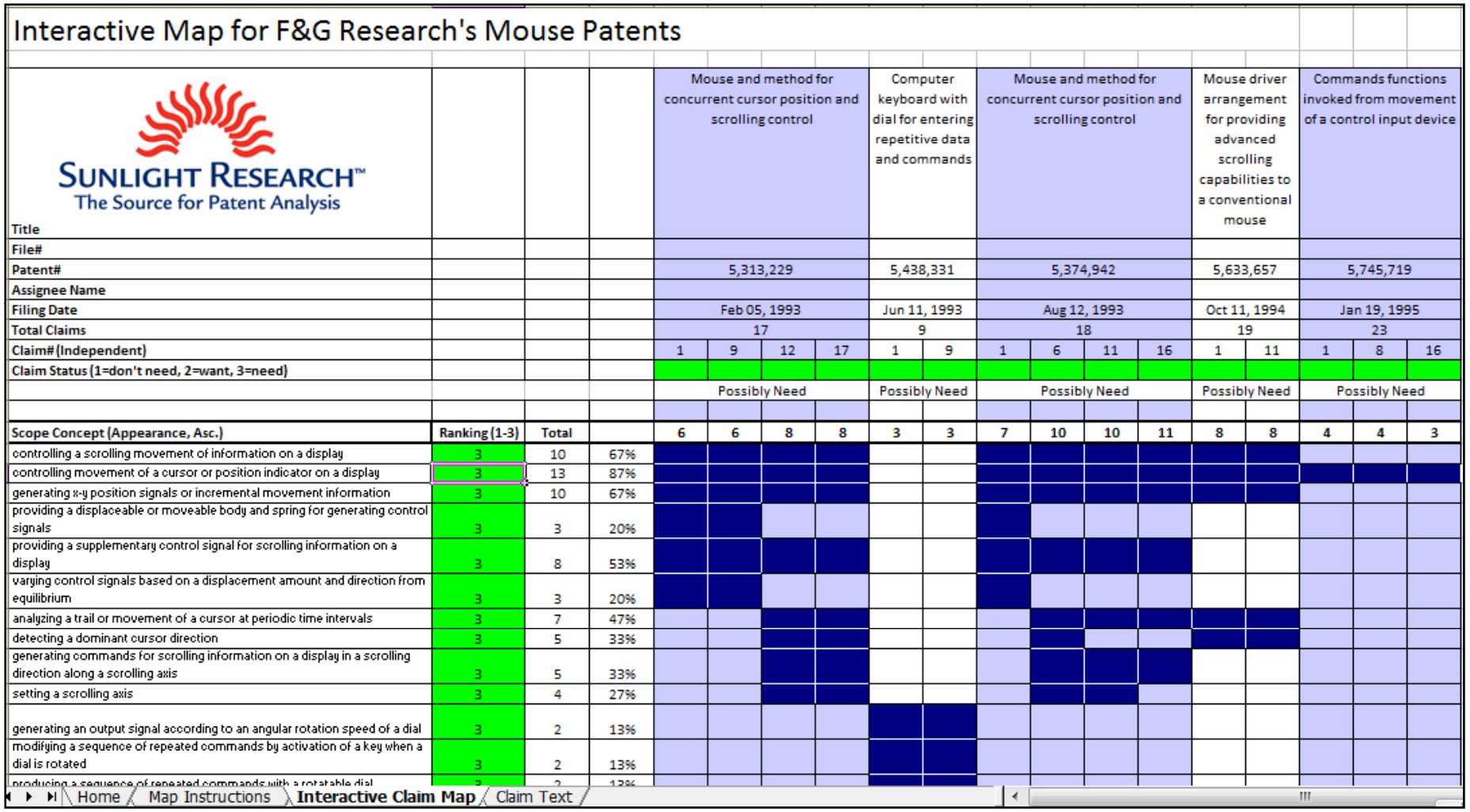
Google *"Search"* vs. Deep Web Report



Examples of a HarvestIP Prior Art Search

Deep Web Harvest Reports
&
HarvestIP Searchable Portals

Sunlight Research Portfolio Claim Map



Interactive Map for F&G Research's Mouse Patents

SUNLIGHT RESEARCH™
The Source for Patent Analysis

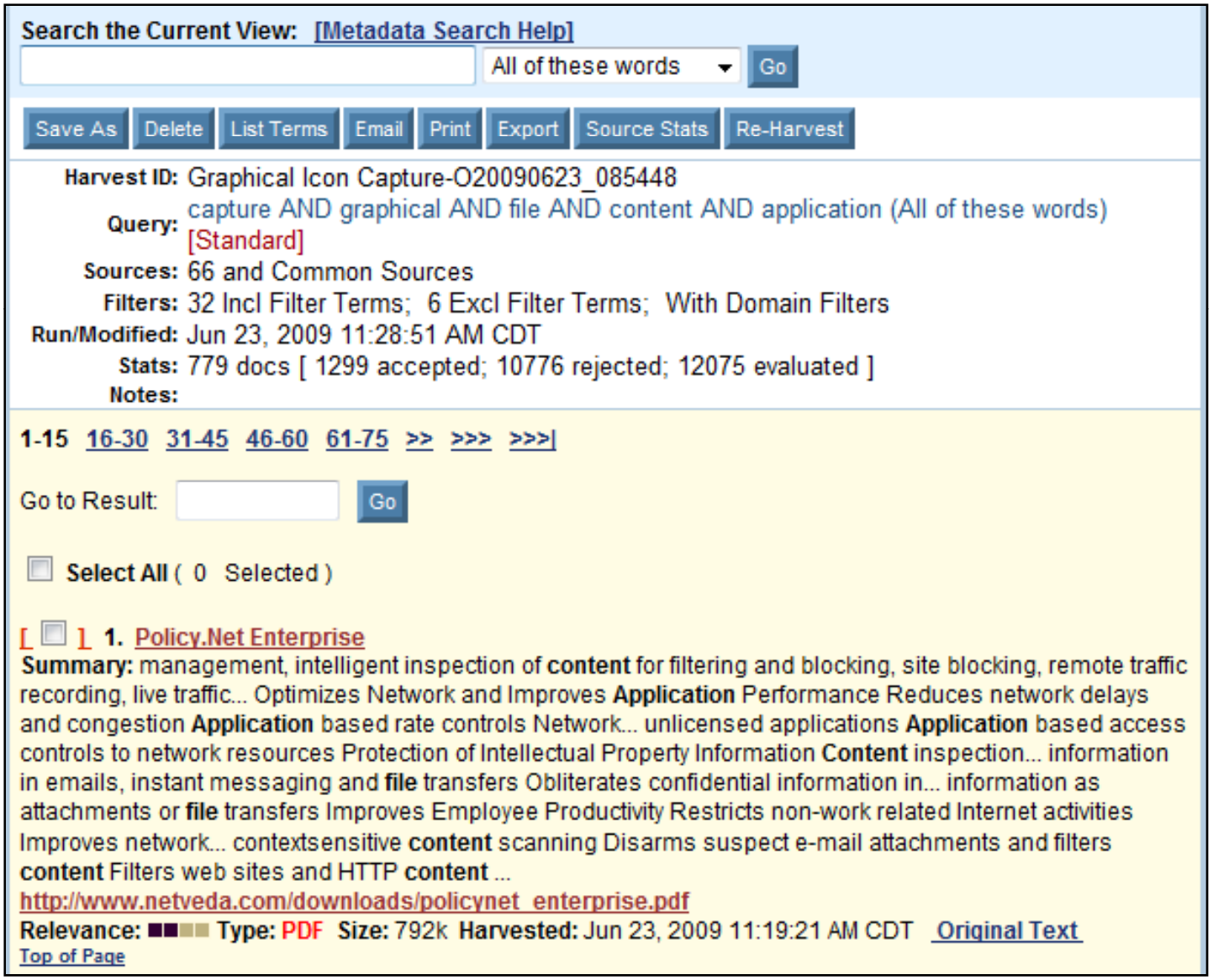
Title				Mouse and method for concurrent cursor position and scrolling control				Computer keyboard with dial for entering repetitive data and commands		Mouse and method for concurrent cursor position and scrolling control				Mouse driver arrangement for providing advanced scrolling capabilities to a conventional mouse		Commands functions invoked from movement of a control input device		
File#																		
Patent#				5,313,229				5,438,331		5,374,942				5,633,657		5,745,719		
Assignee Name																		
Filing Date				Feb 05, 1993				Jun 11, 1993		Aug 12, 1993				Oct 11, 1994		Jan 19, 1995		
Total Claims				17				9		18				19		23		
Claim# (Independent)				1	9	12	17	1	9	1	6	11	16	1	11	1	8	16
Claim Status (1=don't need, 2=want, 3=need)																		
				Possibly Need				Possibly Need		Possibly Need				Possibly Need		Possibly Need		
Scope Concept (Appearance, Asc.)	Ranking (1-3)	Total		6	6	8	8	3	3	7	10	10	11	8	8	4	4	3
controlling a scrolling movement of information on a display	3	10	67%	■	■	■	■			■	■	■	■	■	■			
controlling movement of a cursor or position indicator on a display	3	13	87%	■	■	■	■			■	■	■	■	■	■			■
generating x-y position signals or incremental movement information	3	10	67%	■	■	■	■			■	■	■	■	■	■			
providing a displaceable or moveable body and spring for generating control signals	3	3	20%	■	■					■								
providing a supplementary control signal for scrolling information on a display	3	8	53%	■	■	■	■			■	■	■	■					
varying control signals based on a displacement amount and direction from equilibrium	3	3	20%	■	■					■								
analyzing a trail or movement of a cursor at periodic time intervals	3	7	47%			■	■				■	■	■	■	■			
detecting a dominant cursor direction	3	5	33%			■	■				■			■	■			
generating commands for scrolling information on a display in a scrolling direction along a scrolling axis	3	5	33%			■	■				■	■	■					
setting a scrolling axis	3	4	27%			■	■				■	■						
generating an output signal according to an angular rotation speed of a dial	3	2	13%					■	■									
modifying a sequence of repeated commands by activation of a key when a dial is rotated	3	2	13%					■	■									
producing a sequence of repeated commands with a rotatable dial	3	2	13%					■	■									

Home | Map Instructions | Interactive Claim Map | Claim Text

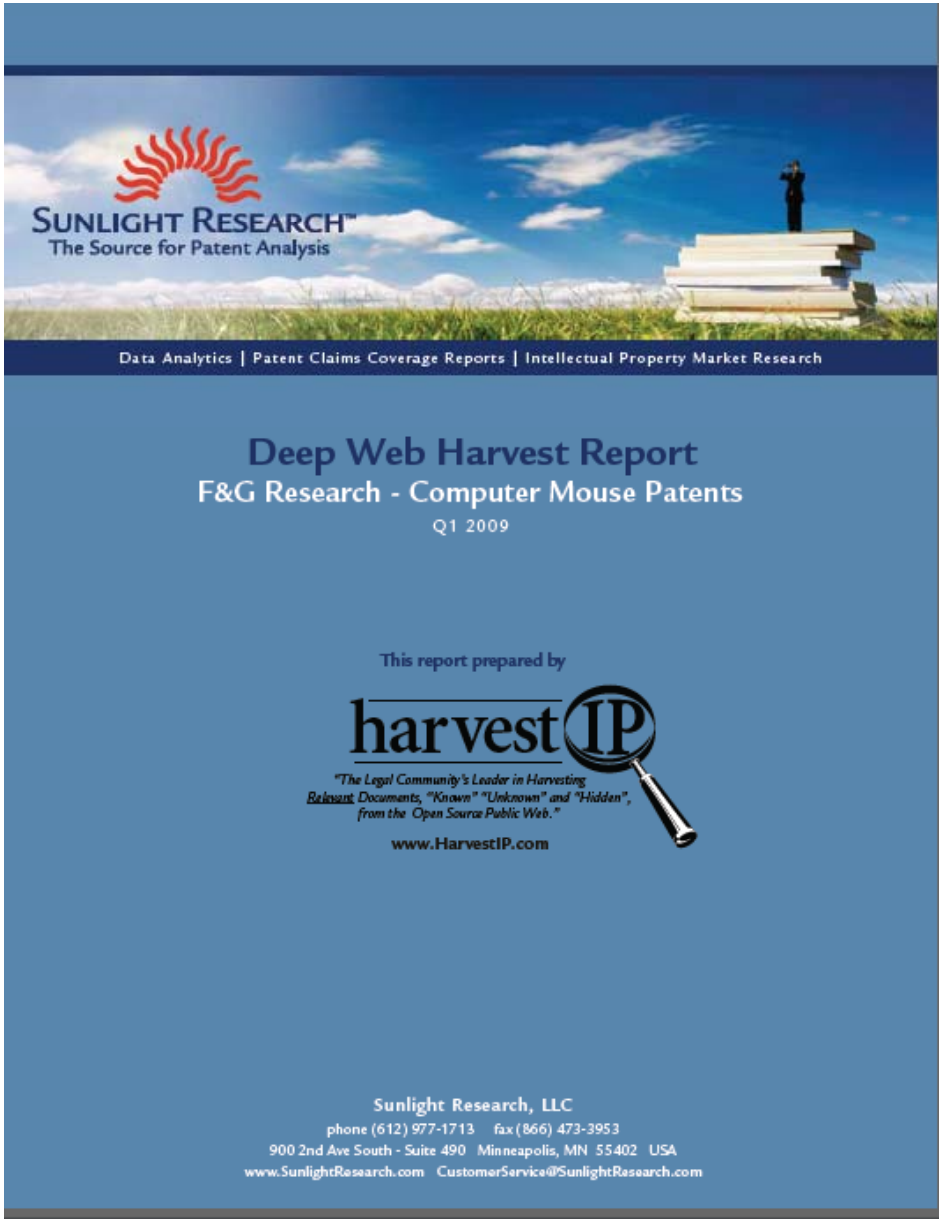
harvestIP

Validity Search for Portfolio Claim Map

(Initial Harvest Results)



Deep Web Harvest Report



Deep Web Harvest Report

Deep Web Harvest Report
F&G Research - Computer Mouse Patents

SUNLIGHT RESEARCH
The Source for Patent Analysis

Quarter 1 - March 2009

Document Summaries

F&G Research - Computer Mouse Patents Document Summaries

Set forth below for each document obtained in the Deep Web Harvest are Document Summaries, which include the document Title, an automatically generated document description, last harvested date, source that the document is located, and a URL link. Each document has been hand-vetted for potential relevancy to this portfolio. All documents referenced in this Deep Web Harvest Report are included in the Harvested Documents folder as a PDF file.

HIP F&G0001

ISSUES FOR PROXIMATE USER INTERFACES Ken Pier Information Sciences and Technology Laboratory Xerox

Summary: display cursor. The user could then invoke an operation from that menu by releasing the mouse ... cursor in response to a particular mouse button press, or a mouse button press without cursor ... equally close to the cursor when the menu appears, and very little cursor motion is needed... the mouse, so that existing applications for the Windows environment may simply substitute pen for mouse. ... the location of the mouse cursor to provide input focus and to direct user actions such... cursor lies. When a user moves a mouse cursor and then lets go of a mouse ...

http://www.cs.berkeley.edu/~landay/research/publications/proximate.pdf

Harvested: Tue Jan 27 11:21:25 CST 2009

Source: UC Berkeley EECU

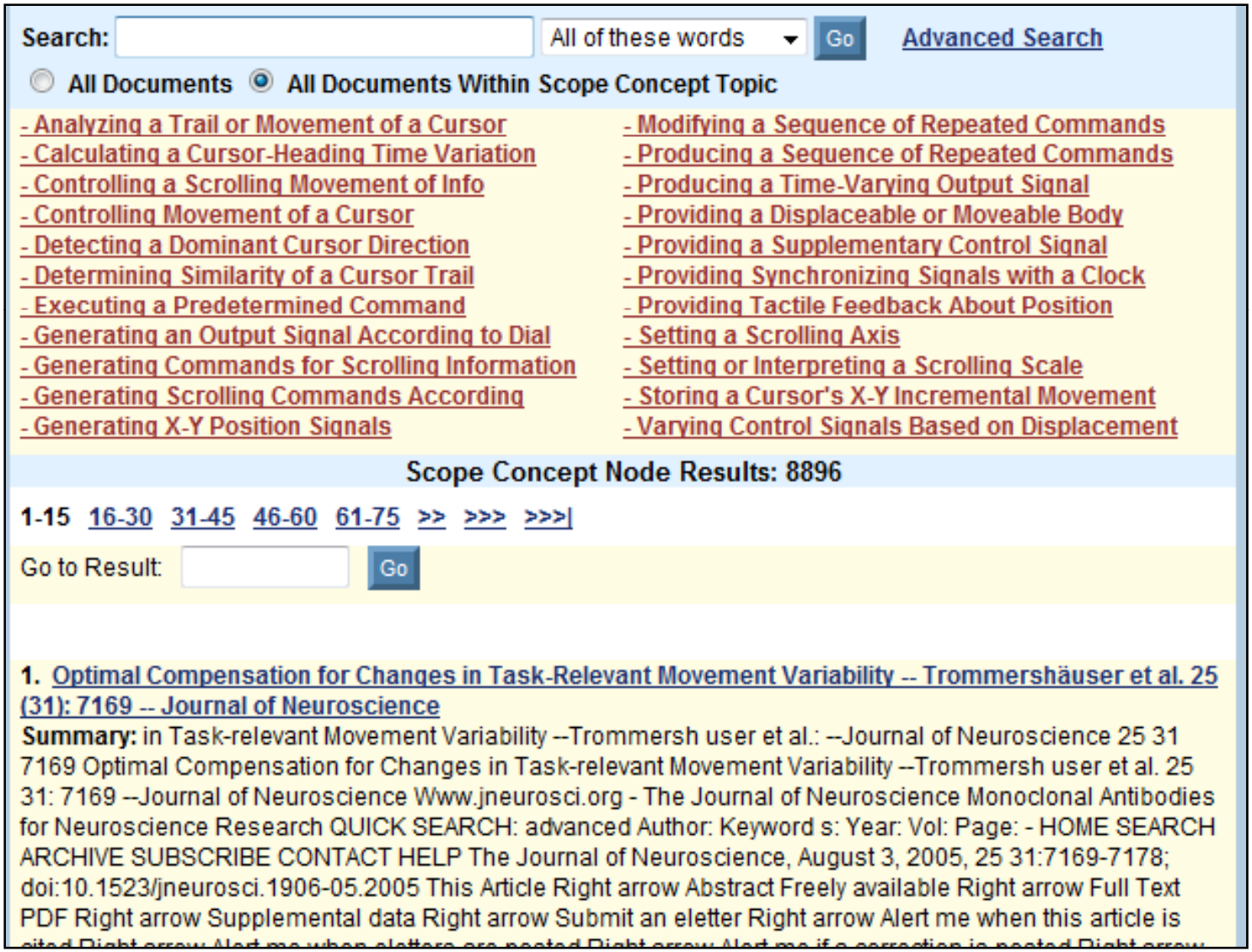
http://www.cs.berkeley.edu/

HIP F&G0002

A Taxonomy of See-Through Tools Eric A. Bier, Maureen C. Stone, Ken Fishkin, William Buxton†, Thoma

Summary: other hand controls a mouse cursor. The user clicks through a tool onto application objects, simultaneously... by positioning a cursor using a mouse in the right hand. Together, the cursor, toolglass sheet,... into place with the mouse cursor and then apply it. These tools have many other potential... s gaze and the mouse cursor to remain in the work area. They can be

HarvestIP Searchable Portal



Harvesting "Unknown and Hidden" Relevant, Current Content for:

- *Validity Search*
- *Patentability*
- *Portfolio Management*
- *Litigation Support*
- *Claims Coverage*
- *Portfolio Management*
- *Competitive Intelligence*
- *Licensing*
- *Freedom to Operate*
- *...and much more.*

harvestIP

harvest IP

http://harvestip.com

Deep Web Harvest Reports can also be found at:

http://sunlightresearch.com

Q & A

Thanks!

Steve Pederson

stevep@harvestIP.com / +1 (612) 840-6888

Will Bushee

willb@harvestIP.com / +1 (605) 212-6396

Troy Mentzer

troym@harvestIP.com / +1 (605) 261-0450

Nora Best

norab@harvestIP.com / +1 (815) 549-6613

harvestIP