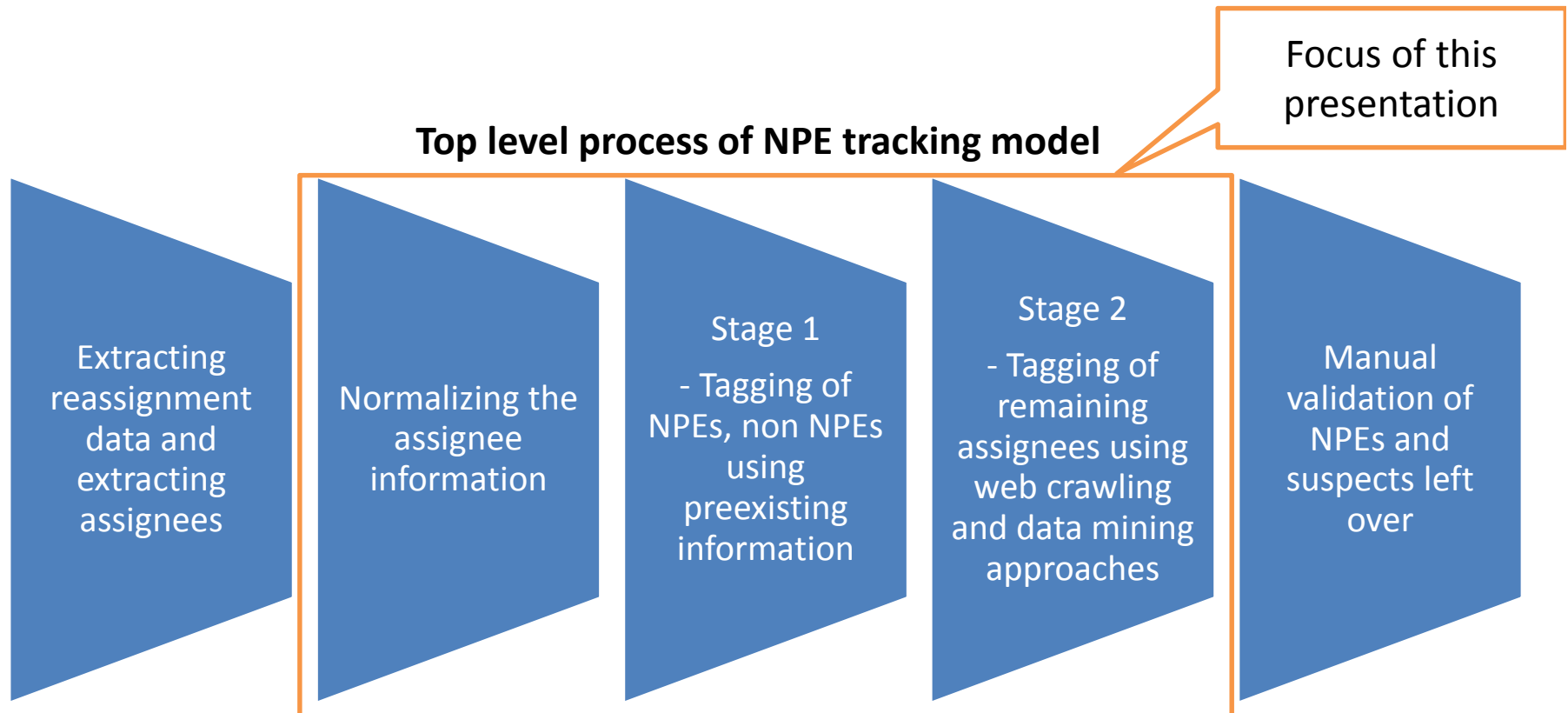# Automating Web Research using Customized Search Tools

Background

- Goal: Tag NPE suspects and identify patents that are reassigned to these entities in a timely, proactive manner.

- 'NPE' refers to entities that Dolcera research could not associate with having significant R&D or products. The goal of acquiring patents is believed to be for profit (litigation, licensing).

Top level process of NPE tracking model

Focus of this presentation

Extracting reassignment data and extracting assignees

Normalizing the assignee information

Stage 1
- Tagging of NPEs, non NPEs using preexisting information

Stage 2
- Tagging of remaining assignees using web crawling and data mining approaches

Manual validation of NPEs and suspects left over

- Test data set: 10 day USPTO reassignments (December 2011)
  - 11k transactions, 4k assignees

Category	Manual	Machine Assisted
Total transactions	11K+	11K+
Total assignees (normalized)	3977	3977
Automatically handled	1438	3776
Manually handled	2539	201
NPEs identified	58	61
False negatives	6	3

- For this demonstration, a sample of 48 assignees previously extracted from the USPTO assignment database has been used to walk through the logic of the model.
  - Sample biased towards suspects and NPEs to cover all aspects of model.

- NPE tracking model was developed iteratively (6 iterations to date)
  - Choice of filters very conservative
  - Using additional filters (assignor, technology etc.) and patent based information (IPC classes covered etc.) is likely to enhance and optimize model further.
  - Model needs testing on larger datasets.
- Leveraged previous Dolcera work
  - Dolcera databases

Database	Strength
Assignee normalization company names corpus	90000 company names from all corners of the world mapped to their ultimate parent through corporate tree
Assignee normalization university names corpus	20000 university names covering all geographies
Public companies database	47000 public companies names collected from various sources
Known-NPE database	3000 NPE names encountered to Dolcera team till date
Known non-NPE database	5000 known non-NPE companies database

  - Address match algorithm
- Potential reusability of developed scripts
  - Identify if a web address belongs to an entity.
  - Web crawling and keyword/context extraction and processing.
  - Use of multiple parameters (litigations, keywords, 3rd party information) for decision making

Sample dataset: 48 assignees

ACTAVIS GROUP PTC EHF

ACTIVEVIDEO NETWORKS INC

PRINCETON DIGITAL IMAGE CORP

A123 SYSTEMS INC

NASOPHLEX B.V

CITIZENS BANK OF PENNSYLVANIA

ACADEMIXDIRECT INC

MAYA-SYSTEMS INC

TERRA NOVA NURSERIES INC

FERGASON PATENT PROPERTIES LLC

INTER DIGITAL

ACCELLENT INC

INTELLECTUAL VENTURES HOLDING 32 LLC

HELYSSEN SARL

AIR PRODUCTS AND CHEMICALS INC

CPA GLOBAL PATENT RESEARCH LTD

MAPL VISION TECH INC

SIGHTSOUND TECHNOLOGIES LLC

ADAPTIVE MATERIALS INC

ABLECO FINANCE LLC A DELAWARE

MAQUIS TECHTRIX LLC

ADETHERAPEUTICS INC

OPEN INVENTION NETWORK LLC

ACERGY FRANCE SA

CLEMSON UNIVERSITY

UNILOC LUXEMBOURG S.A.

COMERICA BANK

INVENSAS CORP

ALNYLAM EUROPE AG

CROSSBAR INC

ADENIUM BIOTECH APS

KILOPASS TECHNOLOGY INC

AISIN SEIKI KABUSHIKI KAISHA

AGILENT TECHNOLOGIES INC

EYAL RESEARCH CONSULTANTS LTD

X2Y ATTENUATORS LLC

PERSONALWEB TECHNOLOGIES LLC

TALON RESEARCH LLC

AIDA ENGINEERING LTD

ADA ENVIRONMENTAL SOLUTIONS LLC

MEDIA SOURCERY INC

AMBERWAVE INC

HEADWATER PARTNERS I LLC

AD-FIN SPA

ZENTARIS LTD

MOTILITY SYSTEMS

TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

UNIVERSAL PREDICTION LTD

Normalizing the assignee information

Stage 1
Tag NPEs, non NPEs using pre existing data

Stage 2
Tag remaining assignees using web crawling and data mining approaches

Dolcera assignee normalization

- Coverts assignee to standard form and detects its corporate tree.
- Database contains 90K+ company names and 20K+ university names along with govt. institutions mapped to ultimate parents through corporate tree.
- Updated regularly – Dolcera analysts, learning module

ACTAVIS GROUP PTC EHF

Spelling error
MAPLE VISION TECHNOLOGIES INC

Standard name
AISIN SEIKI CO LTD

ACTIVEVIDEO NETWORKS INC

SIGHTSOUND TECHNOLOGIES LLC

AGILENT TECHNOLOGIES INC

PRINCETON DIGITAL IMAGE CORP

ADAPTIVE MATERIALS INC

EYAL RESEARCH CONSULTANTS LTD

A123 SYSTEMS INC

ABLECO FINANCE LLC A DELAWARE

X2Y ATTENUATORS LLC

NASOPHLEX B.V

MAQUIS TECHTRIX LLC

PERSONALWEB TECHNOLOGIES LLC

CITIZENS BANK OF PENNSYLVANIA

ADETHERAPEUTICS INC

TALON RESEARCH LLC

ACADEMIXDIRECT INC

OPEN INVENTION NETWORK LLC

AIDA ENGINEERING LTD

MAYA-SYSTEMS INC

ACERGY FRANCE SA

ADA ENVIRONMENTAL SOLUTIONS LLC

TERRA NOVA NURSERIES INC

CLEMSON UNIVERSITY

MEDIA SOURCERY INC **Standard name**

FERGASON PATENT PROPERTIES LLC

Ultimate Parent UNILOC

AMBERWAVE SYSTEMS CORP

Space typo INTERDIGITAL

COMERICA BANK

HEADWATER PARTNERS I LLC

ACCELLENT INC

INVENSAS CORP

AD-FIN SPA **Ultimate Parent**

INTELLECTUAL VENTURES HOLDING LLC

ALNYLAM EUROPE AG

AETERNA ZENTARIS LTD

HELYSSEN SARL

CROSSBAR INC

MOTILITY SYSTEMS

AIR PRODUCTS AND CHEMICALS INC

ADENIUM BIOTECH APS

TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

CPA GLOBAL PATENT RESEARCH LTD

KILOPASS TECHNOLOGY INC

UNIVERSAL PREDICTION LTD

Dolcera assignee normalization

Assignee processing

Cleaning assignee entities if they contain addresses usually identified using punctuation marks

Processing abbreviations e.g. Res gets converted to Research, Int or Intl gets converted to International

Looking for exact matches for given assignee name in corpus.

Looking for edit distance matches to adjust for spelling mistakes e.g. Dalcera Corp assignee name matches with Dolcera Corp with 1 edit distance

If match is detected in any of previous two steps, determine ultimate parent for given assignee.

If match is not found system stores assignee in its other database which is reviewed manually regularly to be added to database

Processing business entities and converting it into standard format. Dolcera Corporation is converted to Dolcera Corp
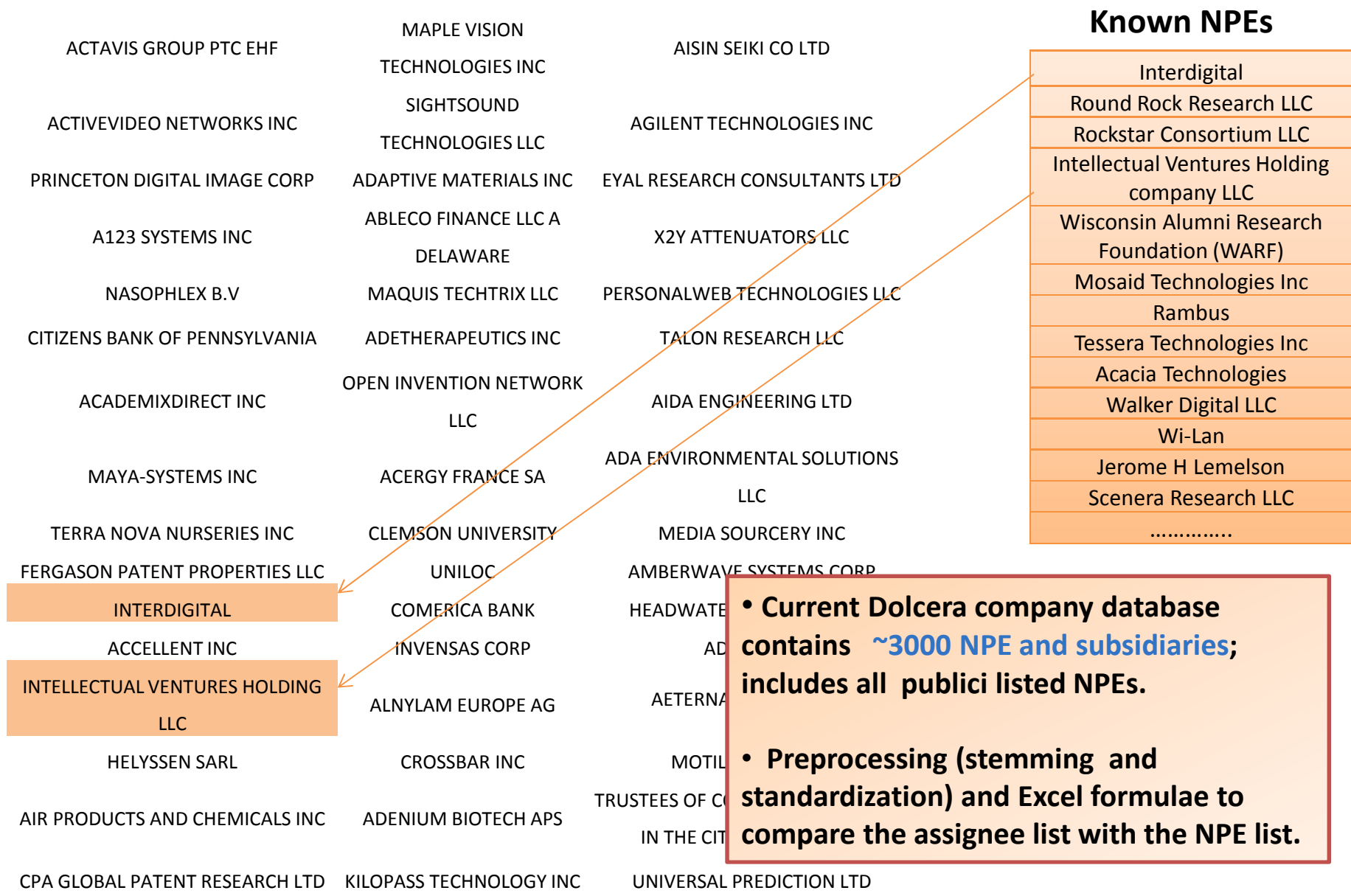
Normalizing the assignee information

Stage 1
- Tag NPEs, non NPEs using pre existing data

1. Use Dolcera company database / public sources to
   1. Tag known NPEs
   2. Tag known non NPEs
   3. Tag public companies, banks, universities
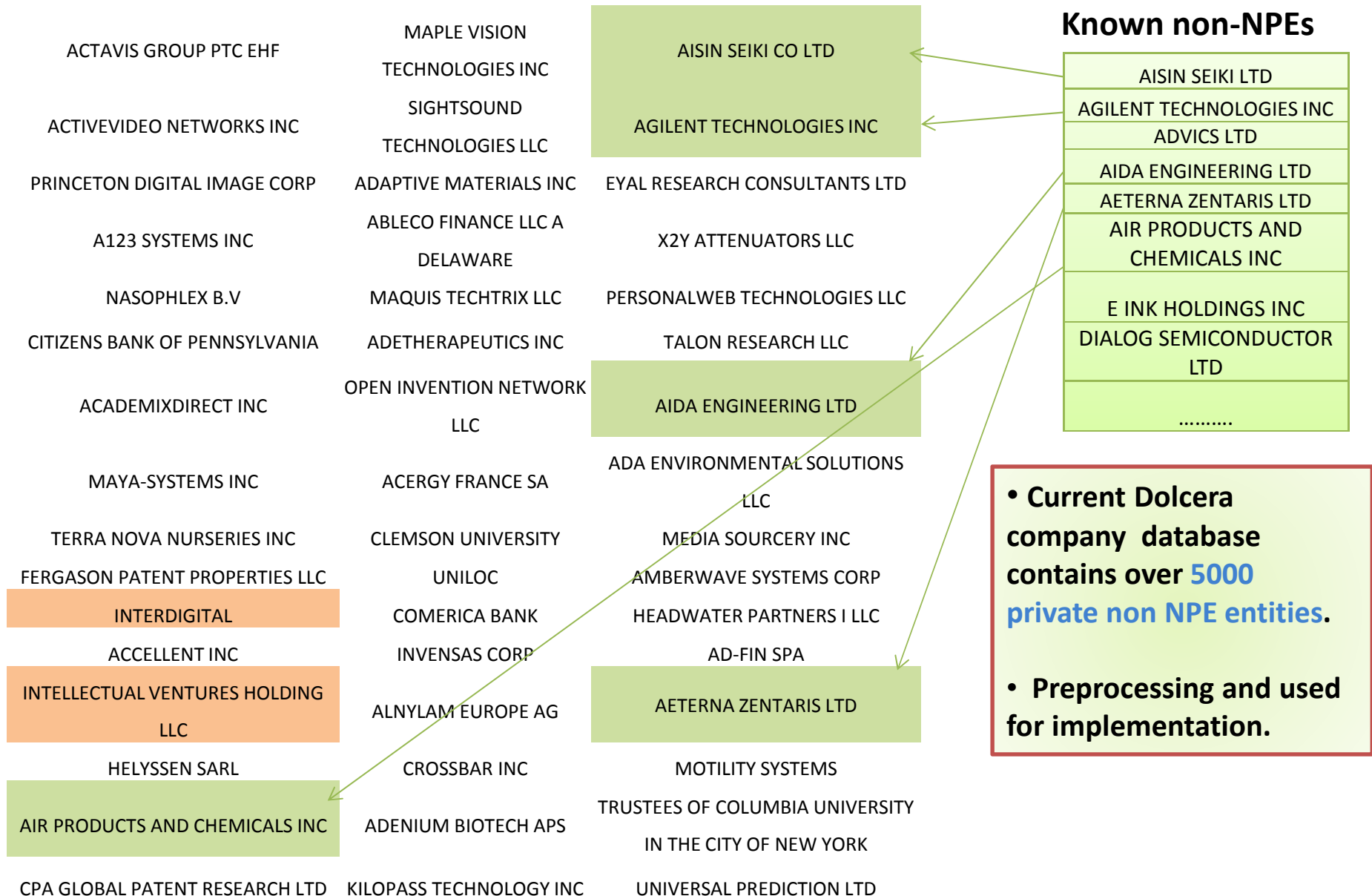2. Tag NPEs using address matching algorithm

Stage 1: Tag known NPEs

Dolcera

ACTAVIS GROUP PTC EHF

ACTIVEVIDEO NETWORKS INC

PRINCETON DIGITAL IMAGE CORP

A123 SYSTEMS INC

NASOPHLEX B.V

CITIZENS BANK OF PENNSYLVANIA

ACADEMIXDIRECT INC

MAYA-SYSTEMS INC

TERRA NOVA NURSERIES INC

FERGASON PATENT PROPERTIES LLC

INTERDIGITAL

ACCELLENT INC

INTELLECTUAL VENTURES HOLDING LLC

HELYSSEN SARL

AIR PRODUCTS AND CHEMICALS INC

CPA GLOBAL PATENT RESEARCH LTD

MAPLE VISION TECHNOLOGIES INC

SIGHTSOUND TECHNOLOGIES LLC

ADAPTIVE MATERIALS INC

ABLECO FINANCE LLC A DELAWARE

MAQUIS TECHTRIX LLC

ADETHERAPEUTICS INC

OPEN INVENTION NETWORK LLC

ACERGY FRANCE SA

CLEMSON UNIVERSITY

UNILOC

COMERICA BANK

INVENSAS CORP

ALNYLAM EUROPE AG

CROSSBAR INC

ADENIUM BIOTECH APS

KILOPASS TECHNOLOGY INC

AISIN SEIKI CO LTD

AGILENT TECHNOLOGIES INC

EYAL RESEARCH CONSULTANTS LTD

X2Y ATTENUATORS LLC

PERSONALWEB TECHNOLOGIES LLC

TALON RESEARCH LLC

AIDA ENGINEERING LTD

ADA ENVIRONMENTAL SOLUTIONS LLC

MEDIA SOURCERY INC

AMBERWAVE SYSTEMS CORP

HEADWATE

AD

AETERNA

MOTIL

TRUSTEES OF C
IN THE CIT

UNIVERSAL PREDICTION LTD

Known NPEs

Interdigital
Round Rock Research LLC
Rockstar Consortium LLC
Intellectual Ventures Holding company LLC
Wisconsin Alumni Research Foundation (WARF)
Mosaid Technologies Inc
Rambus
Tessera Technologies Inc
Acacia Technologies
Walker Digital LLC
Wi-Lan
Jerome H Lemelson
Scenera Research LLC
....

- **Current Dolcera company database contains ~3000 NPE and subsidiaries; includes all publici listed NPEs.**

- **Preprocessing (stemming and standardization) and Excel formulae to compare the assignee list with the NPE list.**

Stage 1: Tag known non NPEs

Dolcera

ACTAVIS GROUP PTC EHF

ACTIVEVIDEO NETWORKS INC

PRINCETON DIGITAL IMAGE CORP

A123 SYSTEMS INC

NASOPHLEX B.V

CITIZENS BANK OF PENNSYLVANIA

ACADEMIXDIRECT INC

MAYA-SYSTEMS INC

TERRA NOVA NURSERIES INC

FERGASON PATENT PROPERTIES LLC

INTERDIGITAL

ACCELLENT INC

INTELLECTUAL VENTURES HOLDING LLC

HELYSSEN SARL

AIR PRODUCTS AND CHEMICALS INC

CPA GLOBAL PATENT RESEARCH LTD

MAPLE VISION TECHNOLOGIES INC

SIGHTSOUND TECHNOLOGIES LLC

ADAPTIVE MATERIALS INC

ABLECO FINANCE LLC A DELAWARE

MAQUIS TECHTRIX LLC

ADETHERAPEUTICS INC

OPEN INVENTION NETWORK LLC

ACERGY FRANCE SA

CLEMSON UNIVERSITY

UNILOC

COMERICA BANK

INVENSAS CORP

ALNYLAM EUROPE AG

CROSSBAR INC

ADENIUM BIOTECH APS

KILOPASS TECHNOLOGY INC

AISIN SEIKI CO LTD

AGILENT TECHNOLOGIES INC

EYAL RESEARCH CONSULTANTS LTD

X2Y ATTENUATORS LLC

PERSONALWEB TECHNOLOGIES LLC

TALON RESEARCH LLC

AIDA ENGINEERING LTD

ADA ENVIRONMENTAL SOLUTIONS LLC

MEDIA SOURCERY INC

AMBERWAVE SYSTEMS CORP

HEADWATER PARTNERS I LLC

AD-FIN SPA

AETERNA ZENTARIS LTD

MOTILITY SYSTEMS

TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

UNIVERSAL PREDICTION LTD

Known non-NPEs

AISIN SEIKI LTD
AGILENT TECHNOLOGIES INC
ADVICS LTD
AIDA ENGINEERING LTD
AETERNA ZENTARIS LTD
AIR PRODUCTS AND CHEMICALS INC
E INK HOLDINGS INC
DIALOG SEMICONDUCTOR LTD
....

- **Current Dolcera company database contains over 5000 private non NPE entities.**

- **Preprocessing and used for implementation.**

Stage 1: Tag public companies, banks, universities

Dolcera

- Public companies database has 47k+ entries.
- Banks and universities tagged after normalization using keyword matches.

ACTAVIS GROUP PTC EHF	MAPLE VISION TECHNOLOGIES INC	AISIN SEIKI CO LTD
ACTIVEVIDEO NETWORKS INC	SIGHTSOUND TECHNOLOGIES LLC	AGILENT TECHNOLOGIES INC
PRINCETON DIGITAL IMAGE CORP	ADAPTIVE MATERIALS INC	EYAL RESEARCH CONSULTANTS LTD
A123 SYSTEMS INC	ABLECO FINANCE LLC A DELAWARE	X2Y ATTENUATORS LLC
NASOPHLEX B.V	MAQUIS TECHTRIX LLC	PERSONALWEB TECHNOLOGIES LLC
CITIZENS BANK OF PENNSYLVANIA	ADETHERAPEUTICS INC	TALON RESEARCH LLC
ACADEMIXDIRECT INC	OPEN INVENTION NETWORK LLC	AIDA ENGINEERING LTD
MAYA-SYSTEMS INC	ACERGY FRANCE SA	ADA ENVIRONMENTAL SOLUTIONS LLC
TERRA NOVA NURSERIES INC	CLEMSON UNIVERSITY	MEDIA SOURCERY INC
FERGASON PATENT PROPERTIES LLC	UNILOC	AMBERWAVE SYSTEMS CORP
INTERDIGITAL	COMERICA BANK	HEADWATER PARTNERS I LLC
ACCELLENT INC	INVENSAS CORP	AD-FIN SPA
INTELLECTUAL VENTURES HOLDING LLC	ALNYLAM EUROPE AG	AETERNA ZENTARIS LTD
HELYSSEN SARL	CROSSBAR INC	MOTILITY SYSTEMS
AIR PRODUCTS AND CHEMICALS INC	ADENIUM BIOTECH APS	TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
CPA GLOBAL PATENT RESEARCH LTD	KILOPASS TECHNOLOGY INC	UNIVERSAL PREDICTION LTD

Stage 1: Address matching algorithm

- Dolcera's assignee normalization database stores the address information of all normalized assignees using latitude-longitude information.

- An address matching algorithm developed converts and matches address information of 'suspect' entities to known NPE's.

Maple Vision Technologies Inc has the same address as Wi-Lan

Results of stage 1 & input to stage 2

Dolcera

Stage 1 tagged entities

Remaining unknowns at start of stage 2

Stage 1 tagged entities		
AISIN SEIKI CO LTD	HELYSSEN SARL	ACTAVIS GROUP PTC EHF
AGILENT TECHNOLOGIES INC	SIGHTSOUND TECHNOLOGIES LLC	ACTIVEVIDEO NETWORKS INC
AIDA ENGINEERING LTD	ADAPTIVE MATERIALS INC	EYAL RESEARCH CONSULTANTS LTD
A123 SYSTEMS INC	ABLECO FINANCE LLC A DELAWARE	NASOPHLEX B.V
CLEMSON UNIVERSITY	MAQUIS TECHTRIX LLC	PERSONALWEB TECHNOLOGIES LLC
CITIZENS BANK OF PENNSYLVANIA	ADETHERAPEUTICS INC	TALON RESEARCH LLC
COMERICA BANK	OPEN INVENTION NETWORK LLC	PRINCETON DIGITAL IMAGE CORP
INVENSYS CORP	ACERGY FRANCE SA	ADA ENVIRONMENTAL SOLUTIONS LLC
AETERNA ZENTARIS LTD	X2Y ATTENUATORS LLC	MEDIA SOURCERY INC
TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK	UNILOC	AMBERWAVE SYSTEMS CORP
AIR PRODUCTS AND CHEMICALS INC	ACADEMIXDIRECT INC	HEADWATER PARTNERS I LLC
INTERDIGITAL	MAYA-SYSTEMS INC	AD-FIN SPA
INTELLECTUAL VENTURES HOLDING LLC	ALNYLAM EUROPE AG	TERRA NOVA NURSERIES INC
MAPLE VISION TECHNOLOGIES INC	CROSSBAR INC	MOTILITY SYSTEMS
ACCELLENT INC	ADENIUM BIOTECH APS	FERGASON PATENT PROPERTIES LLC
CPA GLOBAL PATENT RESEARCH LTD	KILOPASS TECHNOLOGY INC	UNIVERSAL PREDICTION LTD

Assignee types:
- Suspects (using website, litigation, Bloomberg description)
- Suspects (using 3rd party descriptions/references)
- Entities with no "suspicious" information on the web are non NPEs.

Stage 2
 - Tag remaining assignees using web crawling and data mining approaches

Dolcera

- Identify the correct homepage, if it exists
  - Retrieve first 10 links from web search.
  - Use assignee 'name matches' in link and title in addition to count of backslashes, words like index, home etc. (if needed), to find the 'homepage' of the assignee.
    - http://www.accelent.com versus http://www.accelent.com/solutions
  - Ignore results from alibaba.com etc. (htttp://www.accelent .co.en.alibaba.com/)

The first link is the homepage for X2Y Attenuators LLC using name, backslash rule

The third link is the homepage for Talon

X2Y ATTENUATORS LLC

About 6,760 results (0.35 seco...

X2Y Attenuators - Home
www.x2y.com/
X2Y Attenuators, LLC is an Intellectual Property Company that develops ... Patented X2Y® Technology is an enabler to this end; ultra-low inductance X2Y ...

X2Y® S-parameter Extraction
DISCLAIMER: Information and
suggestions furnished in this ...

IP Portfolio
X2Y® improves performance,
component reduction and cost ...

Common Mode Filters
Occurs when individual filters are not
matched. • Differential signal ...

News
X2Y® improves performance,
component reduction and cost ...

DC Motor EMI Suppression
2. DC Motor EMI Suppression. 1.
Why Do DC Motors Need EMI ...

Contact Us
contact us. Fields marked with * are
required. Nature of ...

More results from x2y.com »

X2Y Attenuators, LLC v. Intel Corporation et. al. patent lawsuit

TALON RESEARCH LLC

About 1,120,000 results (0.33 seconds)

Talon Research, LLC v. Hynix Semiconductor Ame... et al ...
dockets.justia.com › ... › Intellectual Property › Patent
14 Oct 2011 – California Northern District Court - Intellectual Pro... atent - **Talon**
Research, LLC v. Hynix Semiconductor America Inc. et al - Jus...

Talon Research, LLC v. Toshiba America Electronic... ponents ...
dockets.justia.com › ... › Intellectual Property › Patent
29 Sep 2011 – California Northern District Court - Intellectual Pro... y - Patent - **Talon**
Research, LLC v. Toshiba America Electronic Components, Inc. al ...

Talon Research
www.talonresearch.com/
Talon Research Inc. brings innovative products to market. Our research develops
solutions for unique product applications in the global marketplace. We serve ...

[PDF] 3 DAVID ALBERTI (SBN 220625) W
cdnet.stpi.org.tw/.../Complaint_pclass_11_A252_1_**Talon**%20Resear...
File Format: PDF/Adobe Acrobat - Quick View
Telephone: 650.618.4360 7 Facsimile: 650.618.4368. 8 Attomeys for Plaintiff Hi 2,1}
S'/\b. 9 **TALON RESEARCH**, LLC g. 10 UNITED STATES DISTRICT COURT ...

Talon Research, LLC v. Hynix Semiconductor America Inc. et. al ...
news.priorsmart.com/**talon-research**-v-hynix-semiconductor-america-...
14 Oct 2011 – **Talon Research, LLC** patent infringement lawsuit against Hynix

- Record occurrences of keywords like patent, licensing, license, intellectual property, attorney etc. if found on the company's website.

- Stop words and context (Ex: keywords in association with others like patent protect* occur equally on all websites) may be used to enhance model (needs additional pattern research).

X2Y Attenuators may be a licensing company: keywords like intellectual property

- Number of links on the website (NPEs tended to have fewer links in our sample).

- Words such as 'distribution', 'careers', 'retail', etc. typically indicate non-NPEs (possible enhancement).

- 'University' and its context (words like laboratory) on the web site is recorded to detect a university spinoff.

NPE/Suspects	Links
Headwater Partners I LLC	10
Motility Systems	9
Amberwave Systems	4

HELYSSEN SARL: University (EPFL), spin off, etc. indicate a non NPE entity

```
commercial spinoff of technology developed in the <a href="http://ie     s.edu/">integratio
partner researchers to pilot applications using multiple modules in th     esearch.</p>
00px; font-size: .9em;">

="http://barobo.com/wp-content/uploads/2011/03/ucd-logo.png" alt="NSF" width="300" height="110"
technology developed in the Integration Engineering Laboratory at the University of California
t/uploads/2011/03/NSF-Logo.png" alt="NSF" width="200" height="110" />
d by the National Science Foundation under Grant No. IIP-1047441. Any opinions, findings, and c
```

Stage 2: Litigations

- Record any litigations associated with the assignee on RFC Express (both as plaintiff, defendant).

X2Y Attenuators LLC is present before "v." in title of the case; recorded as a plaintiff. For X2Y Attenuators, plaintiff count is 2, defendant count is 0.

- Search query format:

    "X2Y Attenuators site:investing.businessweek.com"

- Validation and picking the correct link (using name and pattern matches)

- Search and records keywords like patents, IP, licensing etc.

Your search for **X2Y Attenuators** returned 0 private company results.

Company	Country	Capital IQ Industry
No matches found for X2Y Attenuators in the Any region.		

- No Bloomberg description for X2Y Attenuators.

Company Overview of Adaptive Materials, Inc.

Snapshot	People

Adaptive Materials Inc, another entity on our list has a description, but no patent related keywords. - No litigations, website has no patent or licensing keyword.

Company Overview

Adaptive Materials, Inc. designs, tests, and manufactures solid oxide fuel cell power systems for military, leisure, and emergency use. The company offers lightweight fuel cells that deliver portable power for ground sensors, unmanned aerial vehicles, and robots; and 250 and 300 watt fuel cells for applications requiring more power or longer durations than batteries. It offers portable power to the United States Armed Forces, as well as leisure, remote monitoring, and medical device industries. The

Key Executives For Adaptive Materials, Inc.

Ms. Michelle Crumm
Co-founder and Chief Business Officer

Mr. Aaron Crumm
Co-founder and Chief Visionary Officer

Mr. John Halloran
Co-founder and Chief Technical Officer

Mr. Jon Rice
Director of Field Operations & Sales

Stage 2: Combining parameters for tagging suspects

Dolcera

Condition/Assignee	X2Y ATTENUATORS LLC	MAYA-SYSTEMS INC	PERSONAL WEB TECHNOLOGIES LLC	EYAL RESEARCH CONSULTANTS LTD	CROSSBAR INC	SIGHTSOUND TECHNOLOGIES LLC
'Patent' + 'IP' keywords on homepage	FALSE	FALSE	FALSE	**TRUE**	FALSE	FALSE
'IP' keyword on homepage + litigation plaintiff > defendant	**TRUE**	FALSE	FALSE	FALSE	FALSE	FALSE
Number of links < 15 + 'Patent' keyword	FALSE	**TRUE**	FALSE	FALSE	FALSE	FALSE
Number of links < 15 + 'licens*' in Bloomberg description	FALSE	FALSE	FALSE	FALSE	**TRUE**	FALSE
Keyword 'Patent' in Bloomberg description	FALSE	FALSE	FALSE	FALSE	FALSE	**TRUE**
Litigation as plaintiff > (defendant +2)	FALSE	FALSE	**TRUE**	FALSE	FALSE	FALSE

Stage 2: Combining parameters for tagging suspects

Dolcera

Suspects tagged using the website, litigation, Bloomberg data related parameters combined

AISIN SEIKI CO LTD	HELYSSEN SARL	ACTAVIS GROUP PTC EHF
AGILENT TECHNOLOGIES INC	SIGHTSOUND TECHNOLOGIES LLC	ACTIVEVIDEO NETWORKS INC
AIDA ENGINEERING LTD	ADAPTIVE MATERIALS INC	EYAL RESEARCH CONSULTANTS LTD
A123 SYSTEMS INC	ABLECO FINANCE LLC A DELAWARE	NASOPHLEX B.V
CLEMSON UNIVERSITY	MAQUIS TECHTRIX LLC	PERSONALWEB TECHNOLOGIES LLC
CITIZENS BANK OF PENNSYLVANIA	ADETHERAPEUTICS INC	TALON RESEARCH LLC
COMERICA BANK	OPEN INVENTION NETWORK LLC	PRINCETON DIGITAL IMAGE CORP
INVENSYS CORP	ACERGY FRANCE SA	ADA ENVIRONMENTAL SOLUTIONS LLC
AETERNA ZENTARIS LTD	X2Y ATTENUATORS LLC	MEDIA SOURCERY INC
TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK	UNILOC	AMBERWAVE SYSTEMS CORP
AIR PRODUCTS AND CHEMICALS INC	ACADEMIXDIRECT INC	HEADWATER PARTNERS I LLC
INTERDIGITAL	MAYA-SYSTEMS INC	AD-FIN SPA
INTELLECTUAL VENTURES HOLDING LLC	ALNYLAM EUROPE AG	TERRA NOVA NURSERIES INC
MAPLE VISION TECHNOLOGIES INC	CROSSBAR INC	MOTILITY SYSTEMS
ACCELLENT INC	ADENIUM BIOTECH APS	FERGASON PATENT PROPERTIES LLC
CPA GLOBAL PATENT RESEARCH LTD	KILOPASS TECHNOLOGY INC	UNIVERSAL PREDICTION LTD

Dolcera

- Web search query used for this step is <company name > + patents OR intellectual property licensing

- Links in top 10 results which don't have patent in link or in title are visited (usually patent research websites).

- Algorithm looks for cases where assignee name is mentioned in vicinity (300 chars) of words like patent, IP, known large NPE's etc.

- Every word in dataset related to NPE business has been given some weight. Every assignee gets score based on words observed in the vicinity and thresholds applied.

The U.S. Patent & Trademark Office has no record of Thalveg Data Flow. Nor does California's business registry. The only mention of the company is in Net Perceptions' SEC filings.

Thalveg, it turns out, is a shell company, sister to others with names like Orange Computer, Purple Mountain Server and Maquis Techtrix. They are all controlled by a venture capital firm called Intellectual Ventures, based in Bellevue, Wash. Donnelly never questioned Detk... credentials or his interest in the patents, but lately a lot of people in high-tech circles are beginning to wonder what Intellectual Ventures is up to.

The answer lies inside the mop-topped head of polymath Nathan Myhrvold, the chief executive of Intellectual Ventures, who gained fame and wealth in his 14 years at Microsoft, retiring as chief technology officer in 2000. Myhrvold started Intellectual Ventures with his friend Edward Jung right after they left Microsoft. Three years ago they hired

There is no website for Maquis Techtrix LLC, no Bloomberg description and no litigation case. Algorithm detected a source relating it with known NPE, IV. (http://www.forbes.com/free_forbes/2005/1114/166.html)
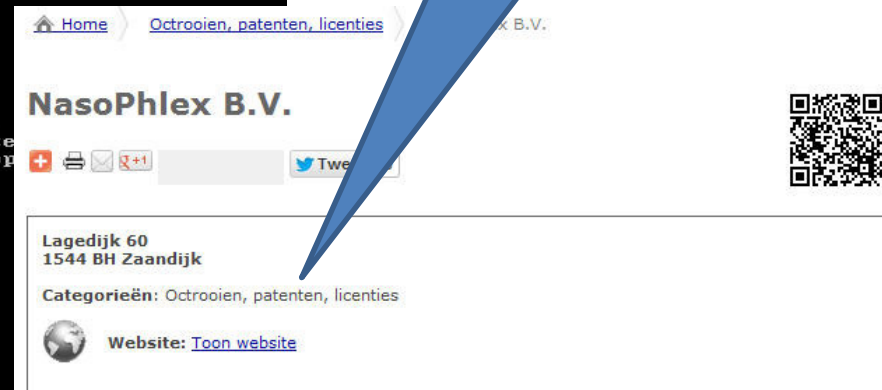
Processing assignees with no direct information

- Non English websites are processed in the native language, to check for keywords like patent, IP etc.


```
For  Nasophlex Bv  looking for websites!
Number of Results:  10
 Finding Bloomberg Information for  Nasophlex Bv
Nasophlex  looking for litigations on RFC express
reached here
0   0
No Keyword Search
Nasophlex
Bv
We believe company is Dutch.. Looking for Dutch words
Nasophlex  Looking for Information on web....
0
Nasophlex Bv " intellectueel eigendom" OR licenties OR patenten  SearchQuery
Going through link 1
Google Search Result: "NasoPhlex B.V. - Octrooien, patenten, licenties in Zaanda
m ..."
length of lengths 6
Word Spotted in vicinity of company name  patenten
Word Spotted in vicinity of company name  patent
Word Spotted in vicinity of company name  licenties
Word Spotted in vicinity of company name  patenten
Word Spotted in vicinity of company name  patent
Word Spotted in vicinity of company name  licenties
Word Spotted in vicinity of company name  patenten
Word Spotted in vicinity of company name  patent
Word Spotted in vicinity of company name  licenties
Word Spotted in vicinity of company name  patenten
Word Spotted in vicinity of company name  patent
Word Spotted in vicinity of company name  licenties
Word Spotted in vicinity of company name  patenten
Word Spotted in vicinity of company name  patent
Word Spotted in vicinity of company name  licenties
Going through link 2
Google Search Result: "Octrooien patenten licenties in Zaandam : Pagina 1 - zaan
dam-mkb.nl"
Going through link 3
Google Search Result: "Nationale Bedrijvengids - Octrooien in Amsterdam"
length of lengths 2
Word Spotted in vicinity of company name  patent
Going through link 4

length of lengths 2
Word Spotted in vicinity of company name  patent
Going through link 5
Google Search Result: "NasoPhlex B.V. uit Zaandijk"

Going through link 6
Google Search Result: "Bedrijven in Zaandijk"

Going through link 7
Google Search Result: "Bulletin 2011/17"
['patenten,patent,licenties,patenten,patent,licenties,patente
patenten,patent,licenties,patenten,patent,licenties,', 'Nasop
www.zaandam-mkb.nl/bedrijven/NasoPhlex_B.V/6776']
----------SUMMARY for Nasophlex Bv ----------
Company website record
['Nasophlex Bv', 'Site not found']
Score: 0.0
Bloomberg information score
No Infomation on Bloomberg Database
Score: 0.0
Litigation information score
['Nasophlex Bv', 'Nasophlex', 0, 0]
Score: 0.0
Keyword search information score
Score: 2.0
```

NASOPHLEX B.V. when subjected to keyword search, algorithm could detect Dutch words 'patenten' and 'licenties' (http://www.zaandam-mkb.nl/bedrijven/NasoPhlex_B.V/6776)

Results

NPEs

INTERDIGITAL

INTELLECTUAL VENTURES HOLDING LLC

MAPLE VISION TECHNOLOGIES INC

Non NPEs

AISIN SEIKI CO LTD	ADAPTIVE MATERIALS INC
AGILENT TECHNOLOGIES INC	ABLECO FINANCE LLC A DELAWARE
AIDA ENGINEERING LTD	ADETHERAPEUTICS INC
A123 SYSTEMS INC	ACERGY FRANCE SA
CLEMSON UNIVERSITY	ACADEMIXDIRECT INC
CITIZENS BANK OF PENNSYLVANIA	ADENIUM BIOTECH APS
COMERICA BANK	ACTAVIS GROUP PTC EHF
AETERNA ZENTARIS LTD	ACTIVEVIDEO NETWORKS INC
TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK	ADA ENVIRONMENTAL SOLUTIONS LLC
AIR PRODUCTS AND CHEMICALS INC	AD-FIN SPA
ACCELLENT INC	HELYSSEN SARL

Suspects

INVENSYS CORP	TALON RESEARCH LLC
CPA GLOBAL PATENT RESEARCH LTD	PRINCETON DIGITAL IMAGE CORP
SIGHTSOUND TECHNOLOGIES LLC	MEDIA SOURCERY INC
MAQUIS TECHTRIX LLC	AMBERWAVE SYSTEMS CORP
OPEN INVENTION NETWORK LLC	HEADWATER PARTNERS I LLC
UNILOC	TERRA NOVA NURSERIES INC
MAYA-SYSTEMS INC	MOTILITY SYSTEMS
ALNYLAM EUROPE AG	FERGASON PATENT PROPERTIES LLC
CROSSBAR INC	UNIVERSAL PREDICTION LTD
KILOPASS TECHNOLOGY INC	X2Y ATTENUATORS LLC
EYAL RESEARCH CONSULTANTS LTD	NASOPHLEX B.V
PERSONALWEB TECHNOLOGIES LLC	

Subject to manual analysis
(<5% of cases of original test data set)