August 7, 2015

Re: **CVDI Year 3 Intellectual Property Report – Response Due by Tuesday, November 10, 2015**

Dear IAB Members:

We are pleased to invite expressions of interest to use the output from the CVDI Year-3 projects. You are eligible to license any copyrightable software or patentable inventions that are developed in the course of the project year, per the CVDI Industry Membership Agreement and By-Laws.

Enclosed you will find the following:

- Year-3 CVDI Intellectual Property Reports
    - CS14-1 Novel Methods for Hidden Relation and Error Detection, Dr. Yuan An
    - CS14-2 Multi-Level and Multi-Source Visual Analytics of Evidence-Based Knowledge Diffusion Processes, Dr. Chaomei Chen
    - CS14-3 A Predictive Analytics Framework for Spatio-Temporal Hotspots, Dr. Jian Chen
    - CS14-4 Analyzing, Modeling, and Summarizing Social Media and Linked Data Sets, Drs. Tony Hu & Ryan Benton
    - CS14-5 Visual Analytic Approaches for Mining Large-Scale Dynamic Graphs, Drs. Raju Gottumukkala & Christoph Borst
- Forms for indicating Year-3 IP licensing intentions for the Industry Member that you represent - **your response is due within 90 days – by 5pm on Tuesday, November 10, 2015**
- Project reports for Year-3 CVDI projects (already provided on July 28th)

As a reminder, based on feedback received during the Year-1 IP selection cycle, the following non-exclusive licensing options are now available for CVDI IP, in addition to the Exclusive License option:

1. **Non-Exclusive Internal Use License.** This license allows for internal-use of the licensed IP for non-commercial purposes by an Industry Member as well as its subsidiaries and affiliates, is royalty-free and no longer requires payment of patent costs. The internal-use license remains available beyond the 90-day time for response to this invitation, and may be requested by an Industry Member at any time.

2. **Non-Exclusive Commercial Distribution License.** This license allows for internal commercial use as well as commercial distribution. This license is a royalty-bearing license and requires reimbursement of related patent costs.

We would also like to convey two important notes regarding patent applications:

1. As a general practice, Drexel and UL will not pursue US or Foreign patent protection for CVDI projects, unless an Industry Member has indicated an interest in patent protection for a specific project.

2. If the Industry Member that you represent would like to see an Academic Member file US or Foreign patent applications for a particular project, please inform the Academic Member within the reply period.

Although the internal-use license is available beyond this 90-day window, we still require an initial response by the **November 10, 2015** due date, so that we can determine the disposition of each IP item with regard to commercial license availability. This will also enable us to make software available and execute patent filings in a timely manner. It is important that you consider your commercial interest during this time, to ensure you have the opportunity to obtain a commercial distribution license. If after the 90-day period has expired, no member has indicated interest in a commercial license, the commercial license options will remain available to CVDI-members, and will also be made available to non-CVDI entities on a first-come basis. Once an exclusive commercial license is executed, the IP will no longer be available for exclusive or non-exclusive commercial licensing.

After the 90-day window has expired, the Academic Member(s), will subsequently contact each Industry Member indicating interest in at least one IP asset to discuss the disposition and next steps.

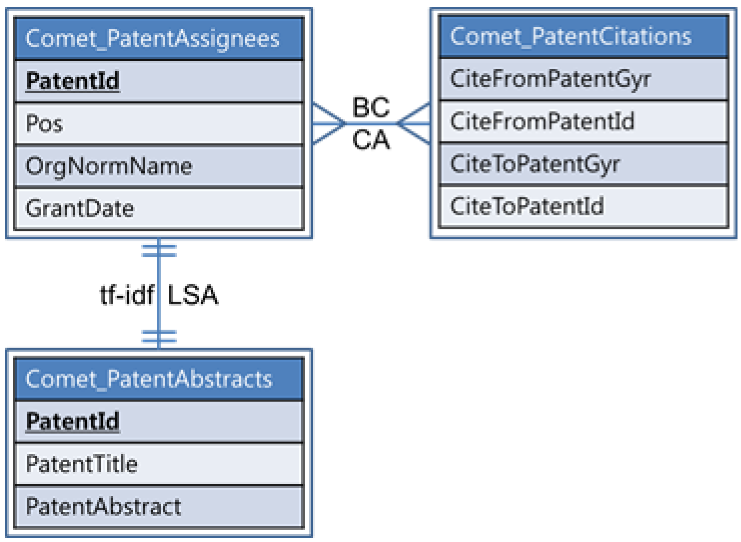
Year-3 CVDI IP Report

The following tables list the software items developed for each Year-3 project that are available for licensing, along with a description of each potentially patentable component. Project reports containing detailed information about the associated research results will be forthcoming. The ID numbers that appear in the tables are identified by the following patterns: Software Item: [Software]-[Project #]-[Software Item #], IP ID: [Patent | Copyright]-[Project #]-[IP #]

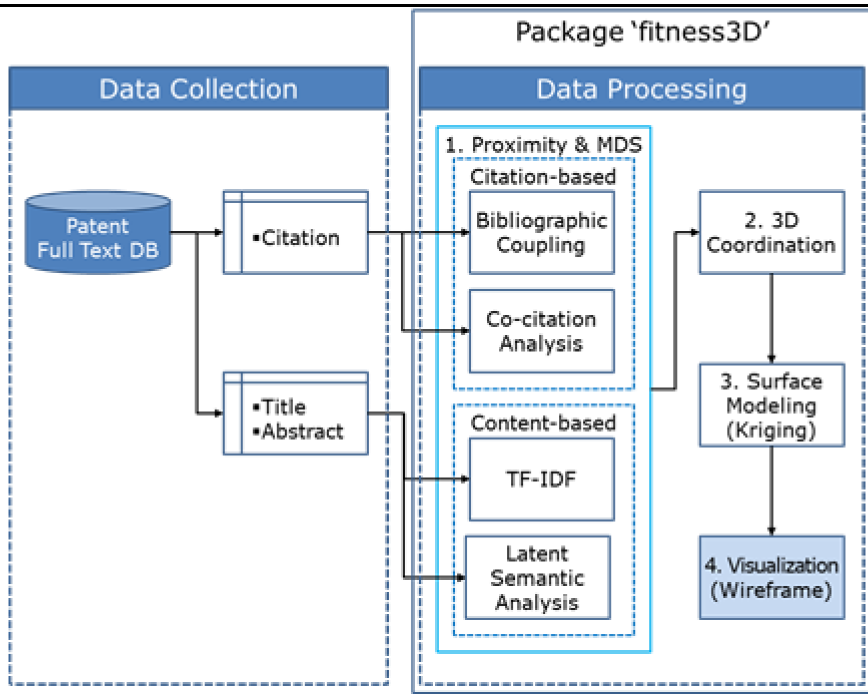
CVDI Project 1 - Novel Methods for Hidden Relation and Error Detection, Dr. Yuan An

Software Item	IP Description
S-1-1	**C-1-1**: **Novel Software for Hidden Relation and Error Detection (In Particular, Author Name Disambiguation in Citation Records)** This source code is capable of processing author citation dataset for the author name disambiguation task. Included is a fully automatic unsupervised method (DP-AP ensemble) through combining a recent developed Density Peaks (DP) clustering method with Affinity Propagation (AP), two advanced clustering methods that have not previously been applied to the author name disambiguation problem. Please see the Project Report for additional information. Languages/technologies: Java
	P-1-1: A new clustering method for automatic author name disambiguation that combines a recently developed Density Peaks (DP) clustering method with Affinity Propagation (AP).
	P-1-2: A new method for calculating similarity between citation records. For the title and abstract attributes, instead of only using the bag-of-words count based model for the feature representation, we explored to use embedding word vector trained from deep neural network. Word vectors consider semantic relations and the order of words, offering a better way to represent title and abstract information from publications.

CVDI Project 2 - Multi-Level and Multi-Source Visual Analytics of Evidence-Based Knowledge Diffusion Processes, Dr. Chaomei Chen

<table>
<tr><th>Software Item</th><th>IP Description</th></tr>
<tr><td>S-2-1</td><td>
C-2-1: A software prototype capable of generating a 3D fitness landscape with a company’s patenting activity.
Software description:
Input a patent’s citation information and contents (title and abstract) from the comet patent database. The tables and columns employed by the software are as follows.


Methods Bibliographic coupling (bc), co-citation analysis (ca), tf-idf, and latent semantic analysis (lsa)
Process Each step described below corresponds to the numbered procedure in the figure.
</td></tr>
</table>



Step 1. Proximity calculation and multidimensional scaling (MDS): This step computes the distance, *i.e.* similarity, between patents. Citation-based methods calculate the similarity with an intersection of cited patents, whereas titles and abstracts are used for the document-based similarity computation by the external "tm" package. Then, the built-in "lsa" function projects the high dimensional space to a two-dimensional space. Finally, it returns a patent distance matrix in the two-dimensional space.

Step 2. 3D coordination: Based on the distance matrix returned by Step 1, this step constructs a matrix of 3D coordinates including a patent's fitness, *i.e.* cited times.

Step 3. Surface modeling: Using an external package called "geoR", This step interpolates the missing values in the matrix generated from Step 2. Kriging is employed for the spatial estimation.

Step 4. Visualization: The wireframe function of the package "lattice" draws a three-dimensional fitness landscape. By default, patents cited more than or equal to one time are depicted with circles in a given color.

Output 3D fitness landscape and temporal trajectory of competitive intelligence in industry

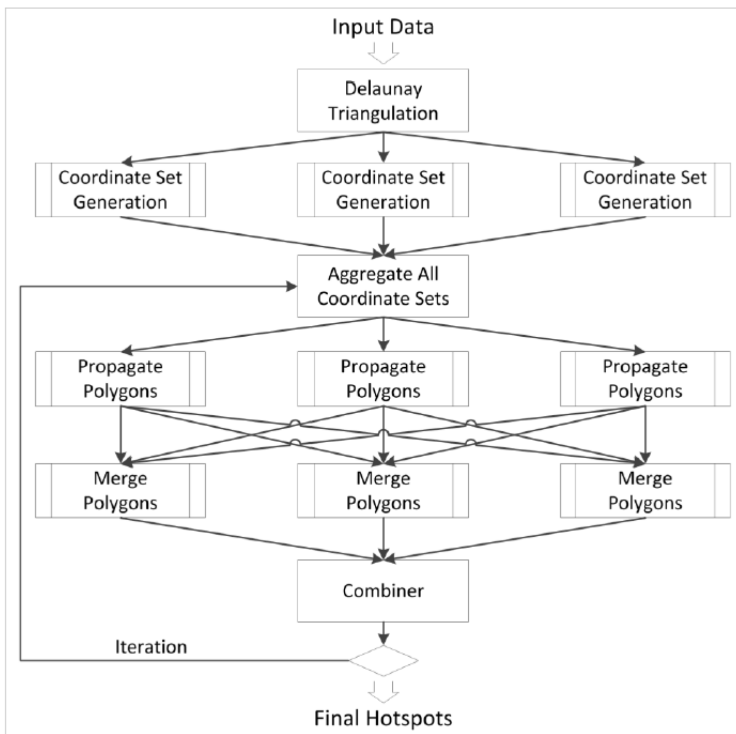
See project report for additional details.

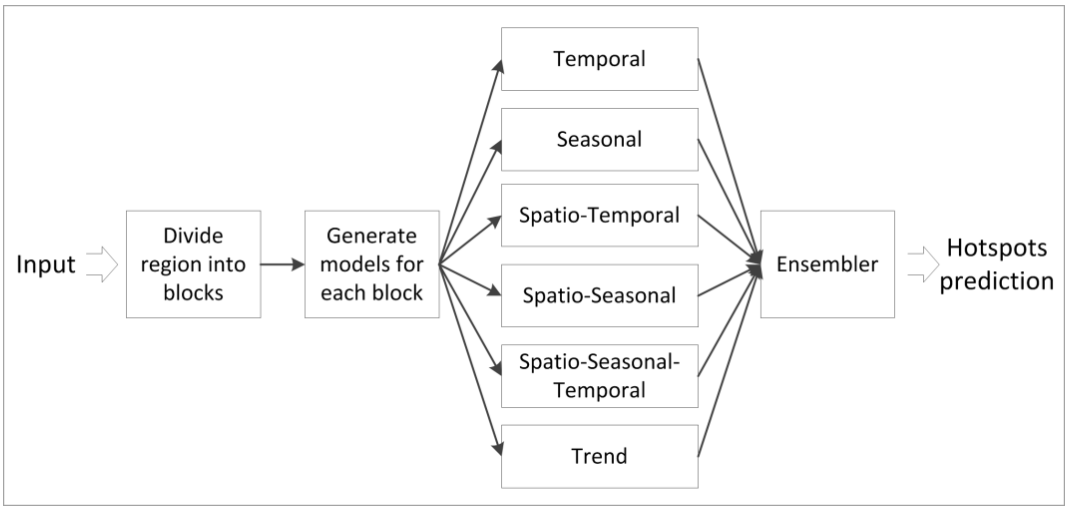
Languages/technologies: R (>=2.12.0), packages 'RODBC', 'tm', 'geoR', and 'lattice'

P-2-1: A new business analytics method that employs fitness landscapes in the field

	of patent analysis to support intuitive navigation of competitive intelligence and technological change in industry

CVDI Project 3 - A Predictive Analytics Framework for Spatio-Temporal Hotspots, Dr. Jian Chen

Software Item	IP Description
S-3-1	**Spatio-Temporal Hotspot Detection Using Polygon Propagation** The Spatio-Temporal Hotspot Detection Using Polygon Propagation identifies spatio-temporal significant clusters (hotspots). It is based on "A space–time permutation scan statistic for disease outbreak detection," (Kulldorff et al, 2005) but further considers the non-circular shaped clusters. The employed algorithm can be implemented on Map-reduce to run in a parallel fashion and yield better performance.  MapReduce framework for polygon propagation Please see the Project Report for additional information. **C-3-1 & P-3-1 :** Copyright and patentable subject matter contained in S-3-1 and Project Report

<table>
<tr><th>Software Item</th><th>IP Description</th></tr>
<tr><td>S-3-2</td><td>
An Approach to Predicting Crime Hotspots Using Ensemble of Spatial, Seasonal and Temporal Forecasting Models.
This ensemble approach identifies the most appropriate combination of models to predict the crime for a particual cluster. It is based on “A three-stage iterative procedure for space-time modeling” (Phillip et al, 1980) but further considers ensemble usage of multiple models where each optimizes the prediction for a certain local area. In addition, the employed algorithim is inherently parallel and could be implemented on Map-reduce to yield better performance.


Procedure for ensemble based prediction
Please see the Project Report for additional information.
C-3-2 & P-3-2 : Copyright and patentable subject matter contained in S-3-2 and Project Report
</td></tr>
</table>

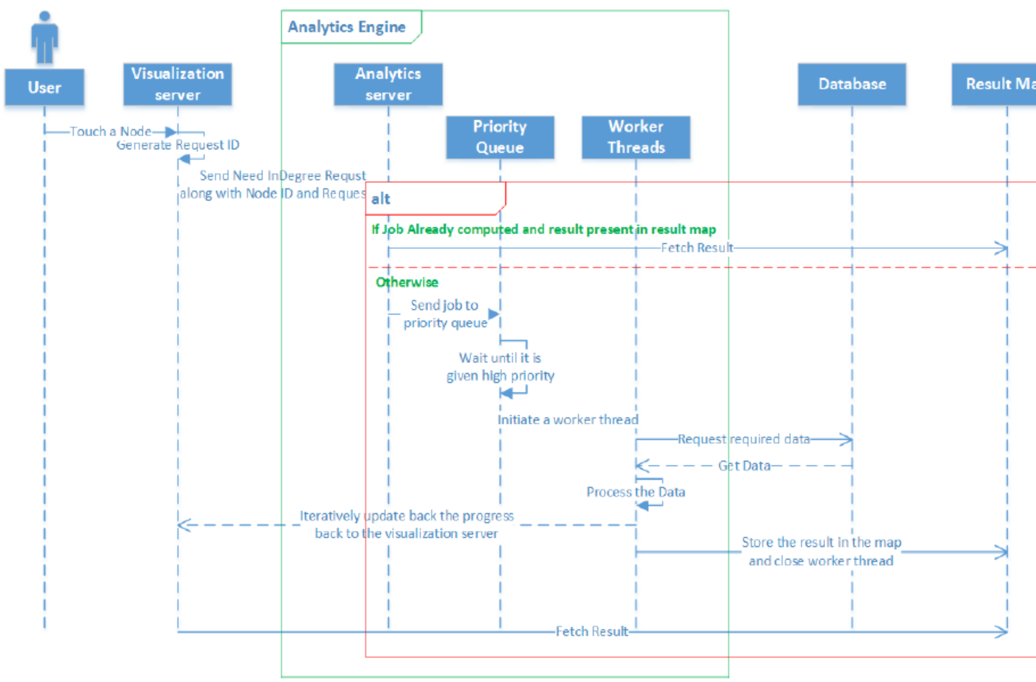
CVDI Project 4 - Analyzing, Modeling, and Summarizing Social Media and Linked Data Sets, Drs. Tony Hu & Ryan Benton (ULL)

Software Item	IP Description
S-4-1	**C-4-1: Software package for Analyzing, Modeling, and Summarizing Social Media and Linked Data Sets** A software package that builds on our earlier CVDI project, by adding support for the following methods: a new HDP-LDA based online review rating regression model named Topic-Sentiment-Preference Regression Analysis (TSPRA) and a Dual Sentimental Hawkes Process (DSHP) to cope with the inadequacies of feature oriented models in social media environments. *Languages/technologies*: Java, CVDI Year-2, Project 1 Software
	P-4-1: A new HDP-LDA based online review rating regression model named Topic-Sentiment-Preference Regression Analysis (TSPRA) Supplemental Information: TPSRA combines topics (i.e. product aspects), word sentiments and per-aspect user preferences as regression factors. TSPRA extends previous sentiment models by taking a key concept of CF models "user preference" into consideration, while it is distinct from current CF models by designing "user preference" and "sentiment" as independent variables that co-determine the review rating. TSPRA adopts Hierarchical Dirichlet Process (HDP) to automatically infer the number of topics.
	P-4-2: A Sentiment Propagation method for social video popularity prediction Supplemental Information: Video popularity prediction plays a foundational role in many aspects of life, such as recommendation systems and investment consulting. Because of its technological and economic importance, this problem has been extensively studied for years. However, four constraints have limited most related works' usability. First, most feature-oriented models are inadequate in the social media environment, because many videos are published with no specific content features, such as a strong cast or a famous script. Second, many studies assume that there is a linear correlation existing between view counts from early and later days, but this is not the case in every scenario. Third, numerous works just take view counts into consideration, but discount associated sentiments. Nevertheless, it is the public opinions that directly drive a video's success/failure. Also, many related approaches rely on a network topology, but such topologies are unavailable in many situations. Here, we

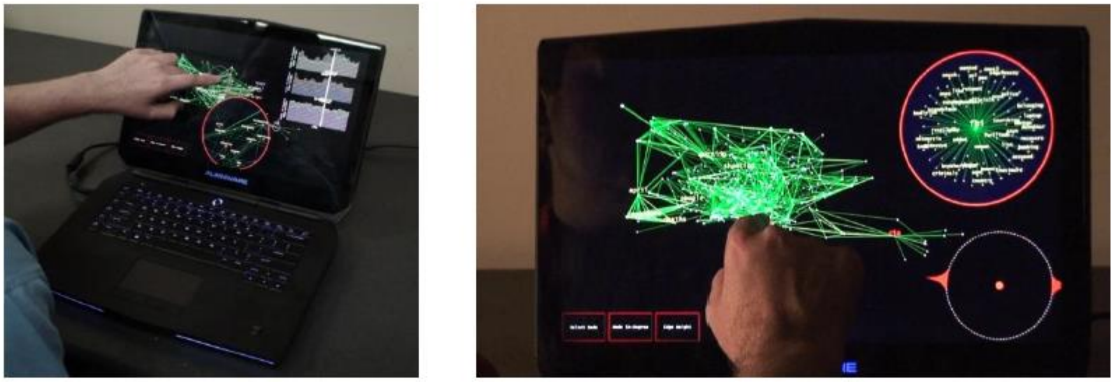
	propose a Dual Sentimental Hawkes Process (DSHP) to cope with all the problems above. DSHP's innovations are reflected in three ways: (1) it breaks the "Linear Correlation” assumption, and implements Hawkes Process; (2) it reveals deeper factors that affect a video's popularity; and (3) it is topology free.

CVDI Project 5 - Visual Analytic Approaches for Mining Large-Scale Dynamic Graphs, Drs. Raju Gottumukkala & Christoph Borst

Software Item	IP Description
S-5-1	**Visual Analytic Framework for Mining Large-Scale Dynamic Graphs** Here we have developed a visual analytics framework for large-scale time-evolving graphs. The novel framework includes (1) design of a visual-analytics framework to facilitate seamless communication between a visualization station (with a large-scale touch display) and analytical engine components, (2) development of a visual data browser with multi-touch browsing capabilities (the visualization station) and tools to query the graphs, and (3) performance evaluation of the framework by experimenting with various data models and applying various graph mining algorithms.



Process flow and message sequence for a sample use case



Visualization station running on a touch-enabled laptop

Please see the Project Report for additional information.

C-5-1 & P-5-1 : Copyright and patentable subject matter contained in S-5-1 and Project Report

Using the enclosed sheets please indicate the licensing intentions for the Industry Member that you represent, checking all license-types that are of interest for each item.

Please note that if only one Industry Member wishes to obtain a non-exclusive commercial license, that Industry Member will have the option to obtain an exclusive commercial license. For that reason, please be sure to indicate whether you might like to obtain an exclusive commercial-distribution license.

If Academic Member(s) do not receive a response by the due date, the non-response will be interpreted as an indication that "No License" is desired for any Year-3 IP item at this time. **A non-response will constitute a formal forfeiting of entitlement rights only to a <u>commercial</u> license of CVDI Year-3 IP**.

Please return a signed copy of this letter and the attached sheets (**on which you must check your licensing interests**) by 5 pm, **<u>Tuesday, November 10, 2014</u>** by e-mail to the sender of this letter or by mail to: Drexel University Office of Technology Commercialization, The Left Bank Building, 3180 Chestnut Street, Suite 104, Philadelphia, Pennsylvania 19104

Sincerely,

Robert B. McGrath
Senior Associate Vice Provost

C. Dean Domingue
Director, Office of Innovation Management

Read and Accepted.

______________________ ______________________ _________

Name Signature Date

Drexel University
Office of Technology Commercialization
The Left Bank Building
3180 Chestnut Street
Suite 104
Philadelphia, Pennsylvania 19104

Re: CVDI Year 3 Intellectual Property Report: CVDI Project 1

Dear Dr. McGrath:

Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 1 - Novel Methods for Hidden Relation and Error Detection, Dr. Yuan An

Software Item	IP Description
S-1-1	**C-1-1**: **Novel Software for Hidden Relation and Error Detection (In Particular, Author Name Disambiguation in Citation Records)** This source code is capable of processing author citation dataset for the author name disambiguation task. Included is a fully automatic unsupervised method (DP-AP ensemble) through combining a recent developed Density Peaks (DP) clustering method with Affinity Propagation (AP), two advanced clustering methods that have not previously been applied to the author name disambiguation problem. Please see the Project Report for additional information. Languages/technologies: Java
	P-1-1: A new clustering method for automatic author name disambiguation that combines a recently developed Density Peaks (DP) clustering method with Affinity Propagation (AP).
	P-1-2: A new method for calculating similarity between citation records. For the title and abstract attributes, instead of only using the bag-of-words count based model for the feature representation, we explored to use embedding word vector trained from deep neural network. Word vectors consider semantic relations and the order of words, offering a better way to represent title and abstract information from publications.

Non-Exclusive Internal-Use License	**Non-Exclusive Commercial License**	**Exclusive Commercial License**	**No License**
☐	☐	☐	☐

_______________ _______________________ _________
Name Signature Date

Drexel University
Office of Technology Commercialization
The Left Bank Building
3180 Chestnut Street
Suite 104
Philadelphia, Pennsylvania 19104

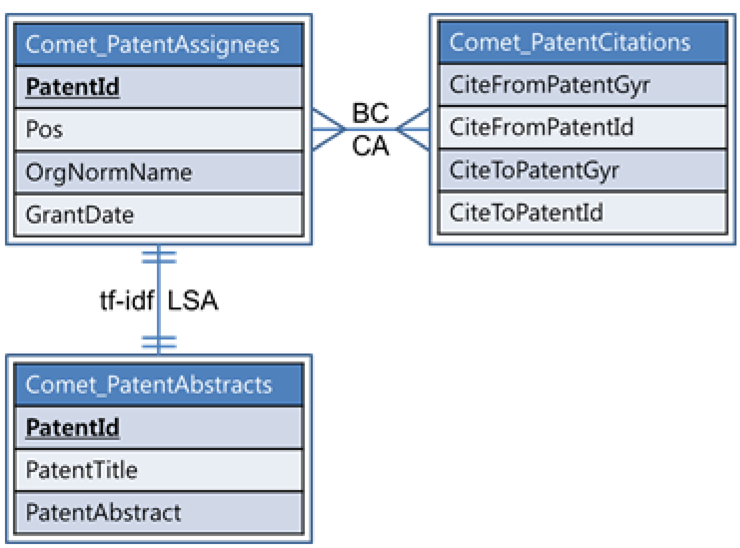
University of Louisiana at Lafayette
Office of Innovation Management
PO Box 43610
Lafayette, LA 70504

Re: CVDI Year 3 Intellectual Property Report: CVDI Project 2

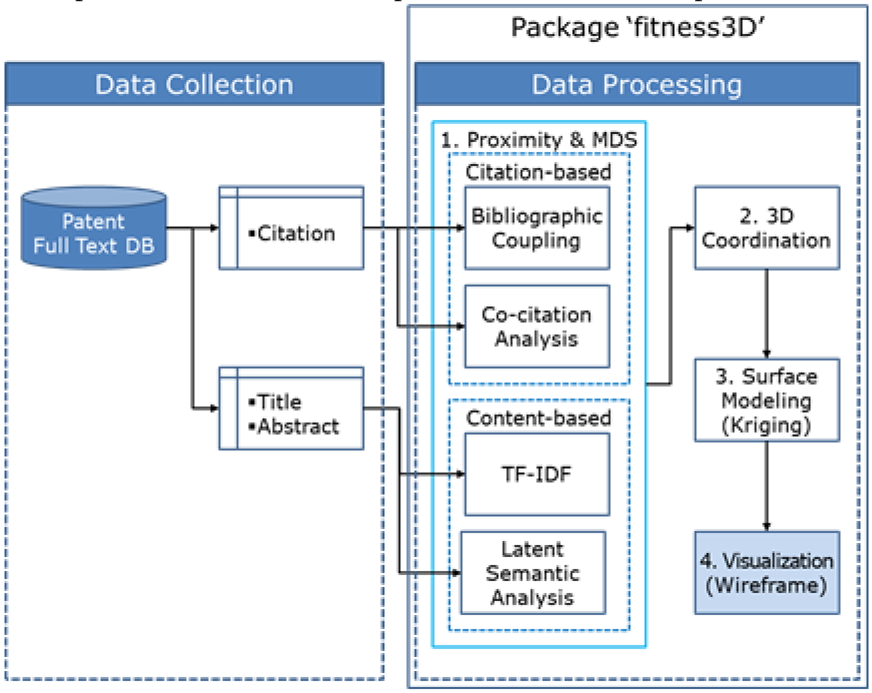
Dear Dr. McGrath:

Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 2 - Multi-Level and Multi-Source Visual Analytics of Evidence-Based Knowledge Diffusion Processes, Dr. Chaomei Chen

Software Item	IP Description
S-2-1	**C-2-1: A software prototype capable of generating a 3D fitness landscape with a company's patenting activity.** Software description: **Input** a patent's citation information and contents (title and abstract) from the comet patent database. The tables and columns employed by the software are as follows.  **Methods** Bibliographic coupling (bc), co-citation analysis (ca), tf-idf, and latent semantic

	analysis (lsa) **Process** Each step described below corresponds to the numbered procedure in the figure.



Step 1. Proximity calculation and multidimensional scaling (MDS): This step computes the distance, *i.e.* similarity, between patents. Citation-based methods calculate the similarity with an intersection of cited patents, whereas titles and abstracts are used for the document-based similarity computation by the external "tm" package. Then, the built-in "lsa" function projects the high dimensional space to a two-dimensional space. Finally, it returns a patent distance matrix in the two-dimensional space.

Step 2. 3D coordination: Based on the distance matrix returned by Step 1, this step constructs a matrix of 3D coordinates including a patent's fitness, *i.e.* cited times.

Step 3. Surface modeling: Using an external package called "geoR", This step interpolates the missing values in the matrix generated from Step 2. Kriging is employed for the spatial estimation.

Step 4. Visualization: The wireframe function of the package "lattice" draws a three-dimensional fitness landscape. By default, patents cited more than or equal to one time are depicted with circles in a given color.

Output 3D fitness landscape and temporal trajectory of competitive intelligence in industry

See project report for additional details.

Languages/technologies: R (>=2.12.0), packages 'RODBC', 'tm', 'geoR', and 'lattice'

	P-2-1: A new business analytics method that employs fitness landscapes in the field of patent analysis to support intuitive navigation of competitive intelligence and technological change in industry

Non-Exclusive Internal-Use License	**Non-Exclusive Commercial License**	**Exclusive Commercial License**	**No License**
☐	☐	☐	☐

________ Name

_____________________ Signature

_________ Date

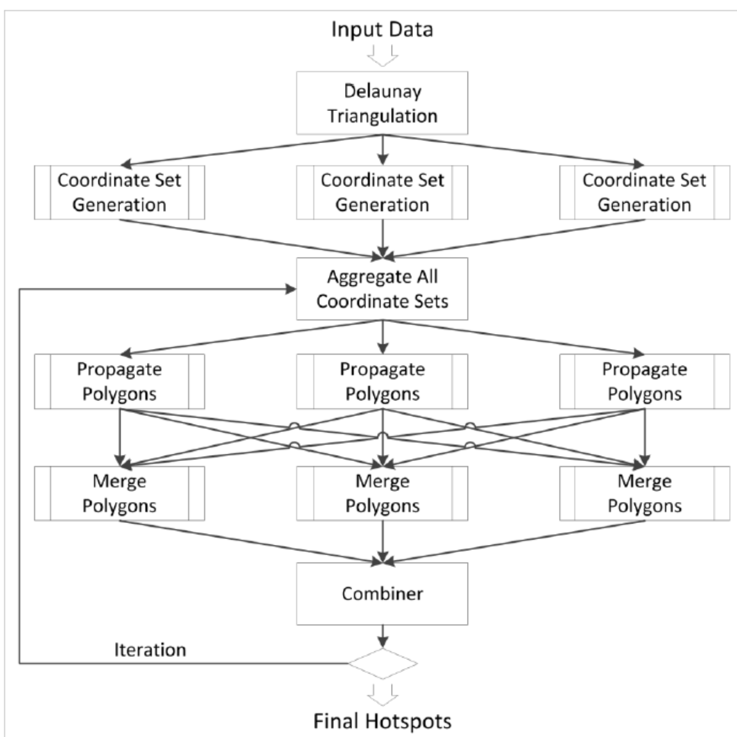
University of Louisiana at Lafayette
Office of Innovation Management
PO Box 43610
Lafayette, LA 70504

Re: CVDI Year 3 Intellectual Property Report: CVDI Project 3

Dear Mr. Domingue:

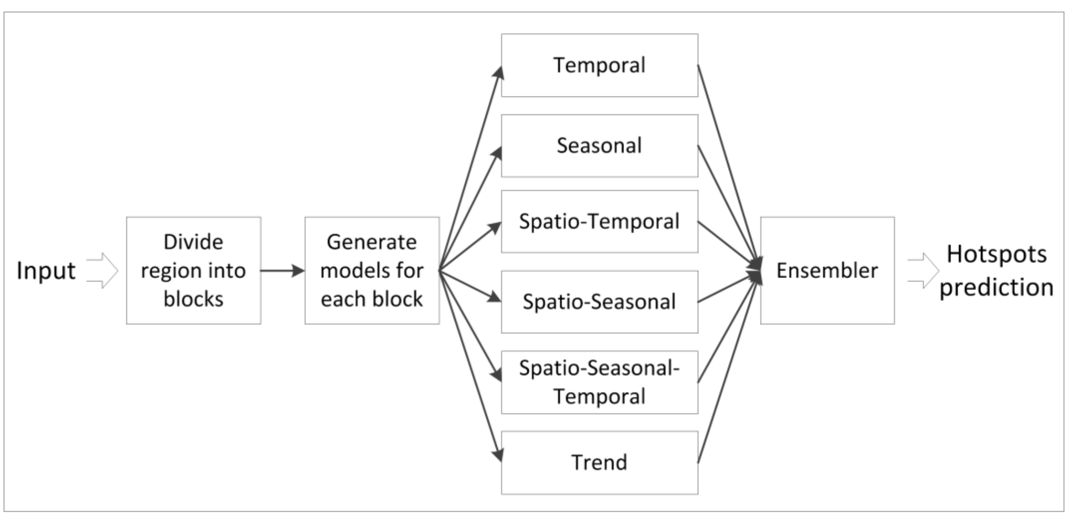
Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 3 -

Software Item	IP Description
S-3-1	**Spatio-Temporal Hotspot Detection Using Polygon Propagation** The Spatio-Temporal Hotspot Detection Using Polygon Propagation identifies spatio-temporal significant clusters (hotspots). It is based on "A space–time permutation scan statistic for disease outbreak detection," (Kulldorff et al, 2005) but further considers the non-circular shaped clusters. The employed algorithm can be implemented on Map-reduce to run in a parallel fashion and yield better performance.  MapReduce framework for polygon propagation

	Please see the Project Report for additional information. **C-3-1 & P-3-1 :** Copyright and patentable subject matter contained in S-3-1 and Project Report

Non-Exclusive Internal-Use License	Non-Exclusive Commercial License	Exclusive Commercial License	No License
☐	☐	☐	☐

Software Item	IP Description
S-3-2	**An Approach to Predicting Crime Hotspots Using Ensemble of Spatial, Seasonal and Temporal Forecasting Models.** This ensemble approach identifies the most appropriate combination of models to predict the crime for a particual cluster. It is based on "A three-stage iterative procedure for space-time modeling" (Phillip et al, 1980) but further considers ensemble usage of multiple models where each optimizes the prediction for a certain local area. In addition, the employed algorithim is inherently parallel and could be implemented on Map-reduce to yield better performance.  *Procedure for ensemble based prediction*

	Please see the Project Report for additional information. **C-3-2 & P-3-2 :** Copyright and patentable subject matter contained in S-3-2 and Project Report

Non-Exclusive Internal-Use License	**Non-Exclusive Commercial License**	**Exclusive Commercial License**	**No License**
☐	☐	☐	☐

Name ________________________ Signature __________ Date _______

Drexel University
Office of Technology Commercialization
The Left Bank Building
3180 Chestnut Street
Suite 104
Philadelphia, Pennsylvania 19104

University of Louisiana at Lafayette
Office of Innovation Management
PO Box 43610
Lafayette, LA 70504

Re: CVDI Year 3 Intellectual Property Report: CVDI Project 4

Dear Dr. McGrath and Mr. Domingue:

Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 4 - Analyzing, Modeling, and Summarizing Social Media and Linked Data Sets, Drs. Tony Hu & Ryan Benton (ULL)

Software Item	IP Description
S-4-1	**C-4-1: Software package for Analyzing, Modeling, and Summarizing Social Media and Linked Data Sets** A software package that builds on our earlier CVDI project, by adding support for the following methods: a new HDP-LDA based online review rating regression model named Topic-Sentiment-Preference Regression Analysis (TSPRA) and a Dual Sentimental Hawkes Process (DSHP) to cope with the inadequacies of feature oriented models in social media environments. *Languages/technologies*: Java, CVDI Year-2, Project 1 Software

	P-4-1: A new HDP-LDA based online review rating regression model named Topic-Sentiment-Preference Regression Analysis (TSPRA) TPSRA combines topics (i.e. product aspects), word sentiments and per-aspect user preferences as regression factors. TSPRA extends previous sentiment models by taking a key concept of CF models "user preference" into consideration, while it is distinct from current CF models by designing "user preference" and "sentiment" as independent variables that co-determine the review rating. TSPRA adopts Hierarchical Dirichlet Process (HDP) to automatically infer the number of topics.

Non-Exclusive Internal-Use License	**Non-Exclusive Commercial License**	**Exclusive Commercial License**	**No License**
☐	☐	☐	☐

____________________ ________________ ________

Name Signature Date

University of Louisiana at Lafayette
Office of Innovation Management
PO Box 43610
Lafayette, LA 70504

Re: CVDI Year 3 Intellectual Property Report: CVDI Project 5

Dear Mr. Domingue:

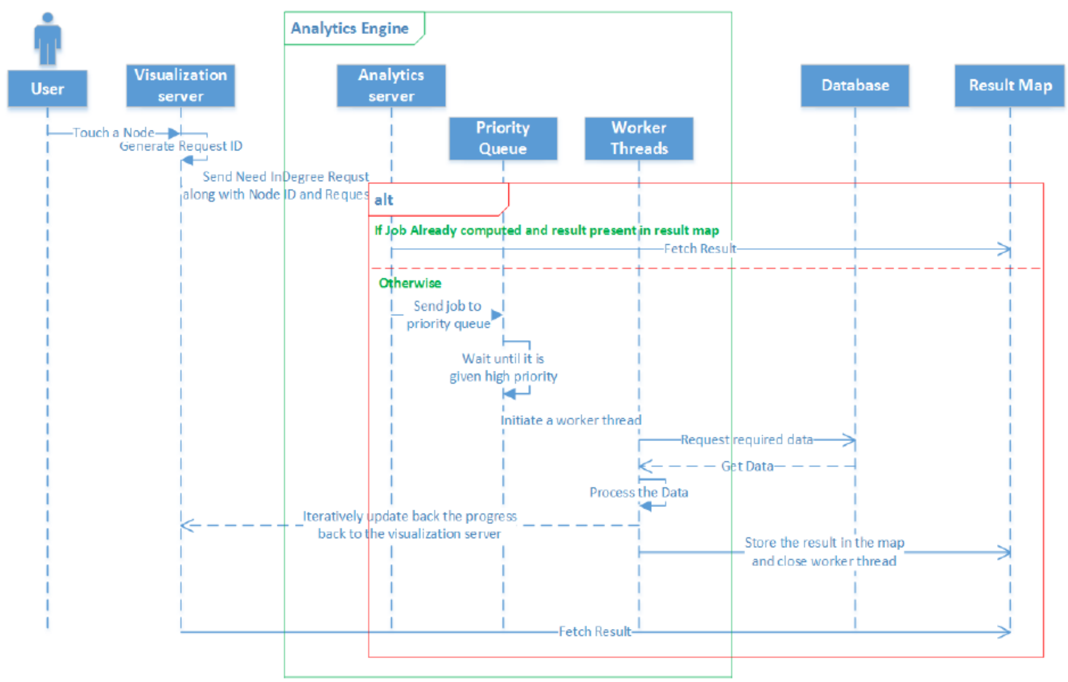
Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 5 –

<table>
<tr><th>Software Item</th><th>IP Description</th></tr>
<tr><td>S-5-1</td><td>

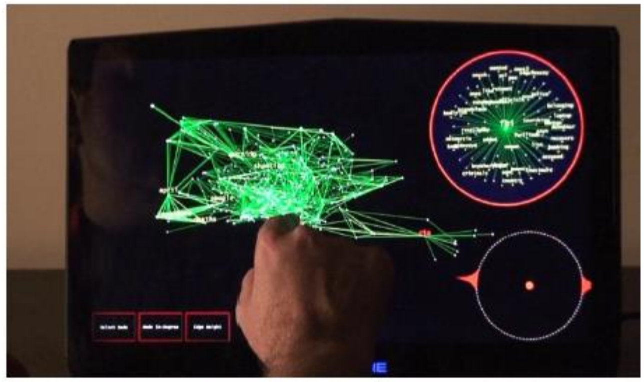
Visual Analytic Framework for Mining Large-Scale Dynamic Graphs

Here we have developed a visual analytics framework for large-scale time-evolving graphs. The novel framework includes (1) design of a visual-analytics framework to facilitate seamless communication between a visualization station (with a large-scale touch display) and analytical engine components, (2) development of a visual data browser with multi-touch browsing capabilities (the visualization station) and tools to query the graphs, and (3) performance evaluation of the framework by experimenting with various data models and applying various graph mining algorithms.



Process flow and message sequence for a sample use case

</td></tr>
</table>

Visualization station running on a touch-enabled laptop

Please see the Project Report for additional information.

C-5-1 & P-5-1 : Copyright and patentable subject matter contained in S-5-1 and Project Report

Non-Exclusive Internal-Use License	Non-Exclusive Commercial License	Exclusive Commercial License	No License
☐	☐	☐	☐

Name Signature Date