August 4, 2014

Re: **CVDI Year 2 Intellectual Property Report – Response Due by Tuesday, November 4, 2014**

- **CVDI Project 1 - Semantic Information Extraction, Integration, and Visualization for Big Data Analytics**
- **CVDI Project 2 - Large-scale Social Media Analytical Tools with Application to Detecting Emerging Events**
- **CVDI Project 3 - Visual Analytic Approaches to Mining Large-scale Time-Evolving Graphs**
- **CVDI Project 4 – A Spatio-Temporal Data Mining Approach for Fraud Detection**
- **CVDI Project 5 – Scalable Visualization, Gap Analytics for Multiple Big Data Industry Sectors**

Dear IAB Members:

We are pleased to invite expressions of interest to use the output from the CVDI Year-2 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-2 CVDI Intellectual Property Reports
- Forms for indicating Year-2 IP licensing intentions for the Industry Member that you represent - **your response is due within 90 days – by 5pm on Tuesday, November 4, 2014**
- Project reports for Year-2 CVDI projects

Based on feedback received during the Year-1 IP selection cycle, we have modified our process, extending the time for response from 60 to 90 days, and now offering better flexibility in licensing CVDI IP:

1. **Expanded Non-Exclusive License Options.** The Non-Exclusive license option is modified to allow you to indicate interest in obtaining either an internal-use non-exclusive license or a commercial non-exclusive license. The internal-use license allows for internal-use of the licensed IP by an Industry Member as well as its subsidiaries and affiliates, is royalty-free, and as detailed below, *no longer requires payment of patent costs*. The commercial license, which is a royalty-bearing license, will allow for commercial distribution of the licensed IP.
2. **Extended Internal-Use License Availability.** The internal-use license now remains available beyond the 90-day time for response to this invitation, and may be requested by an Industry Member at any time. In the future, we intend to make project software even more easily accessible, via a "click-through" facility on the CVDI website.
3. **Relief on Patent Costs for internal-use.** The Industry Membership Agreement requires that members who exercise rights to a non-exclusive license pay associated patent costs. We have waived that requirement for the internal-use non-exclusive license option. Licensing of IP for commercial-use will still require payment of associated patent costs.

Although the internal-use license is now available beyond this 90-day window, we still require an initial response by the **November 4, 2014** 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 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-2 CVDI IP Report

The following tables list the software items developed for each Year-2 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 - Semantic Information Extraction, Integration, and Visualization for Big Data Analytics

Software Item	IP Description
S-1-1	**C-1-1**: **Prototype for Extracting Structured Knowledge (Gene Expression Relations) from Unstructured Text** A software prototype capable of extracting gene expression relations between genes and brain regions from biomedical literature. *Languages/technologies*: Stanford Natural Language Processing toolkit, MetaMap, SemRep, MultiR
	P-1-1: A method of extracting gene expression relations between genes and brain regions from biomedical literature that does not need manually labeled training examples for learning an extraction model.
	P-1-2: A method of extracting gene expression relations between genes and brain regions from biomedical literature that leverages the results of existing open information extraction systems and combines them with a distant learning method.
	P-1-3: A method of extracting gene expression relations between genes and brain regions from biomedical literature that takes into consideration of the interdependency characteristic of domain entities. With grouping strategy, the method achieves improved results compared with baselines at both sentential level and corpus-level.

CVDI Project 2 - Large-scale Social Media Analytical Tools with Application to Detecting Emerging Events

Software Item	IP Description
S-2-1	**C-2-1: Prototype for Performing Simultaneous Aspect and Sentiment Generation** A software prototype that studies the public opinions towards an entity using a hybrid HDP-LDA model, which can automatically determine the number of aspects, distinguish factual words from opinioned words, and further effectively extracts the aspect specific sentiment words, and performs sentiment analysis more efficiently than other influential models such as JST, AUSM, and MaxEnt-LDA. *Languages/technologies*: Java

	P-2-1: A method that includes an unsupervised LDA-based model adapted not to require the number of aspects to be specified in advance, making the system more objective.
	P-2-2: A method that includes an unsupervised LDA-based model adapted to distinguish between factual words and opinioned words.
	P-2-3: A method that includes an unsupervised LDA-based model configured to extract aspect-specific sentiment words, instead of considering an entire sentence to be of one aspect.

Software Item	IP Description
S-2-2	**C-2-2: Enhance Topic Evolution – Prototype for Learning Hierarchical Topic Models from Short and Noisy Social Media Messages** A software prototype that uses a new generative probabilistic model called Hash-Based Stream LDA for topic modeling in continuous social media streams, which outperformed the classical LDA approach and the stream-oriented On-line LDA and SparseLDA approaches in terms of average perplexity by more than 12% when using data crawled from Twitter, and by 21% when using data crawled from the IRC chat community. *Languages/technologies*: Java
	P-2-4: A topic modeling system having a generative probabilistic model configured to expose facilities to include inter-document similarity in topic modeling.
	P-2-5: A topic modeling system having an inference algorithm that relies on efficient estimation of document similarity with Locality Sensitive Hashing to retain the knowledge of past social discourse in a scalable way.
	P-2-6: A novel topic modeling system that utilizes historical knowledge of previous messages in inference to improve quality of topic discovery.

Software Item	IP Description
S-2-3	**C-2-3 - Affinity propagation (AP) clustering software prototype on Hadoop** A software prototype capable of clustering large text data efficiently and effectively with a distributed affinity propagation mechanism on Hadoop/MapReduce. *Languages/technologies*: Java, Hadoop/MapReduce
	P-2-7: A system of data clustering using Affinity propagation (AP) among data points, redesigned for distributed computation and parallel text clustering on Hadoop.
	P-2-8: A method of partial (pruned) affinity propagation (AP) which reduces time complexity of text clustering and makes the algorithm efficient even with large amounts of text data.
	P-2-9: A method of selecting a balanced value of k, the number of pair-wise similarities to retain in Pruned AP computation, leading to optimal clustering performance (inflection point).

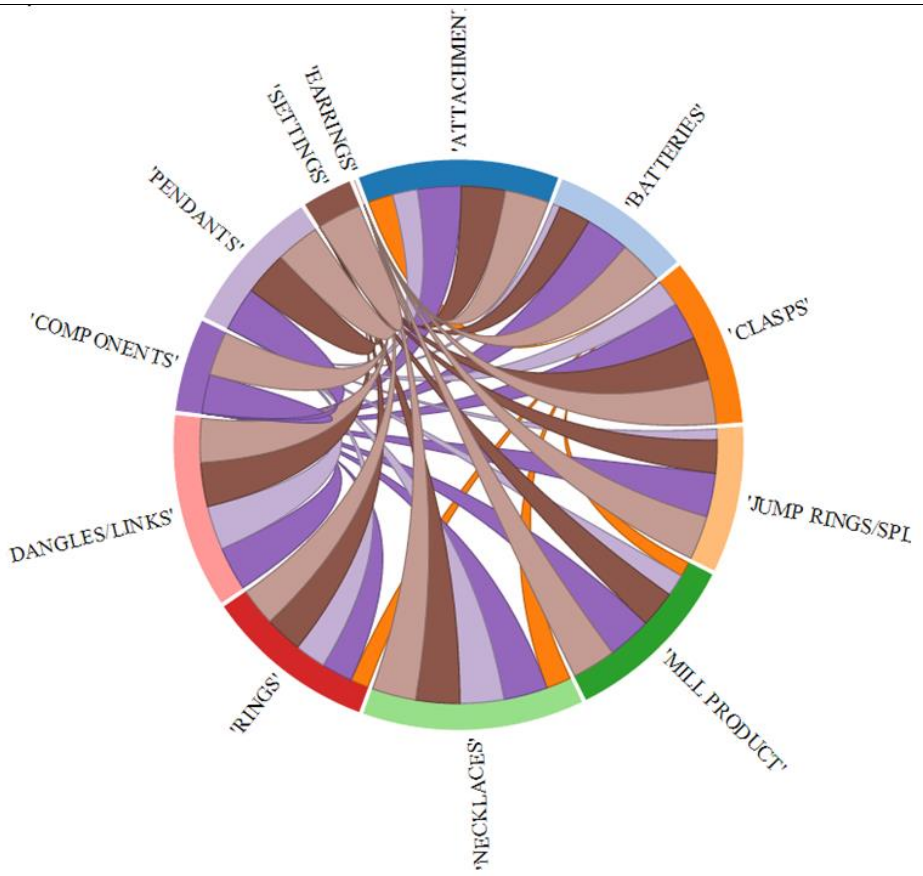
Software Item	IP Description
S-2-4	**C-2-4 - Emerging Event Detection** A software prototype that improves on the original Event Detection on Onset (EDO) by incorporating data from multiple streaming/social medias sources and detecting sub-events.
	P-2-10: A method of detecting emerging events and sub-events by combining data from multiple different media sources, with the ability to process and detect events within a minute.

CVDI Project 3 - Visual Analytic Approaches to Mining Large-scale Time-Evolving Graphs

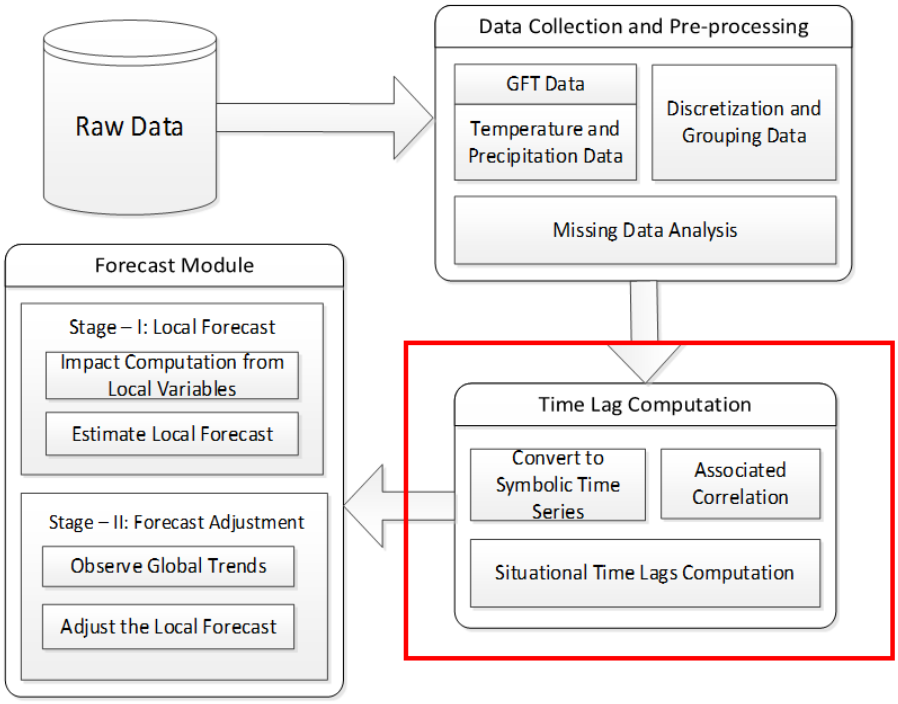
Software Item	IP Description
S-3-1	**C-3-1: Data Mining Tools for Cluster and Association Mining** Prototype software that implements well-know and vetted hierarchical clustering, and association mining to understand relationships between various entities especially as it relates to risky customers, and commercial product dependency

Software Item	IP Description
S-3-2	**C-3-2: Two-phase Forecasting System from Non-linear Time Series Graphs** Prototype software that generates a prediction model for non-linear time-evolving graphs and incorporation of a novel two-phase forecasting system.
	P-3-2: A novel method of incorporating inter-proximity between local forecasting to adjust each individual local forecast.

Software Item	IP Description
S-3-3	**C-3-3: Web-based Visual Analytics Dashboard** Prototype Software that generates a web-based dashboard where a user can generate multiple data visualizations



A wheel showing the dependencies between the products from the transactions of the customer.

Software Item	IP Description
S-3-4	**C-3-4: Symbolic Time Series to Compute Situational Time Lags** Prototype software that implements symbolic time series based time lag computation whereby converting the normal numeric time series into symbolic or categorical time series, and adjusting time lag per category.  **P-3-4:** A novel method of symbolic time-series based time lag computations whereby converting the normal numeric time series into symbolic or categorical time series adjusts time lag per category. In state of the art or most existing techniques the time lag between the time series is computed using cross correlation between the numeric values. The main disadvantage of this previous process is that it yields a single time lag between two time series. Using the current symbolic time series permits extraction of different time lags at different situations, categories.

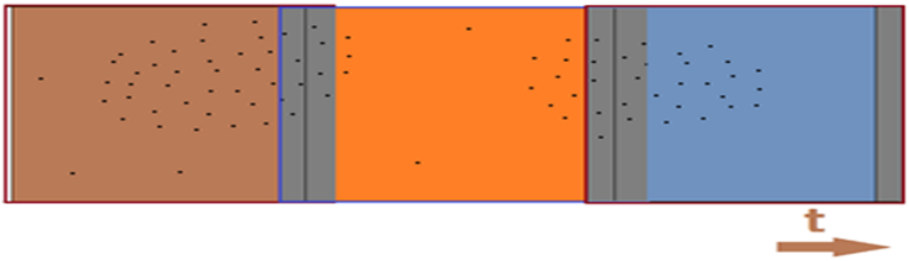
Software Item	IP Description
S-3-5	**C-3-5: Multi-touch Graph Interaction** Prototype software where multi-touch graph interaction allows users to visualize, explore data. While building on the Handymap application, graph interaction expands beyond the need for geo-located datasets.

	 A demonstration of the visualization application with multi touch support. The list-style menus allow analysis data mapped to visualization parameters.
	P-3-5: An improved method for user interface selection for cluttered VR environments using a tracked hand-held touch device.

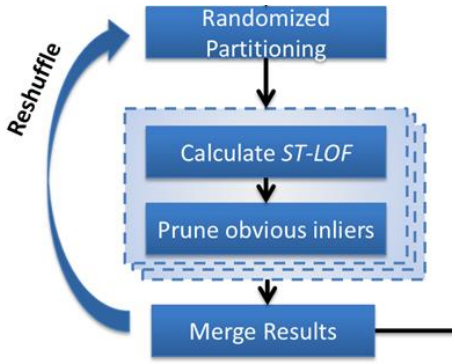
CVDI Project 4 - A Spatio-Temporal Data Mining Approach for Fraud Detection

Software Item	IP Description
S-4-1	**C-4-1: Software for Spatio-temporal local outlier factor (ST-LOF) Calculation** A software prototype that incorporates spatio-temporal context into the Local Outlier Factor (LOF). ST-LOF defines a spatio-temporal outlier to be a spatial-temporal object whose thematic attribute values are significantly different from those of other spatially and temporally referenced objects in its spatial or/and temporal neighborhoods. This software has a user interface for the user to enter the k-th nearest neighbor, number of neighboring points to consider and the normalized input file. As output, each object in the data set is assigned a ST-LOF outlier score and written to a single file. $$ST - LOF_{MinPts}(p) = \frac{\sum_{o \in ST-N_{MinPts}(p)} \frac{ST - lrd_{F-MinPts}(O_f)}{ST - lrd_{F-MinPts}(p_f)}}{|ST - N_{MinPts}(p)|}$$
	P-4-1: A novel method that improves on LOF by introducing spatio-temporal context to every data object and thereby capturing outliers or clusters based on spatio-temporal difference.

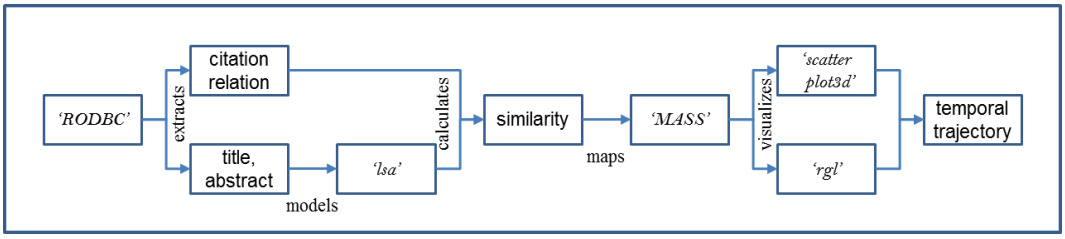
Software Item	IP Description
S-4-2	**C-4-2: Software for Spatio-temporal Local Density Based Clustering of Applications with Noise (ST-LBDSCAN)** Prototype software that incorporates spatio-temporal context into the Local Density Based Spatial Clustering of Applications with Noise (LDBSCAN). ST-LBDSCAN algorithm detects outliers based on spatio-temporal context and clusters the objects that have similar features spatially or/and temporally. The software has a user interface for the user to enter the k-th nearest neighbor, number of neighboring points to consider in ST-LOF and ST-LBDSCAN, local density reachability percentage and a normalized input file. The output is a file with ST-LOF score along with its dataset object ID and its cluster identity (cluster value of -1 implies outlier).
	P-4-2: A novel method that improves on LDBSCAN by introducing spatio-temporal context to every data object and thereby capturing outliers or clusters based on spatio-temporal difference.

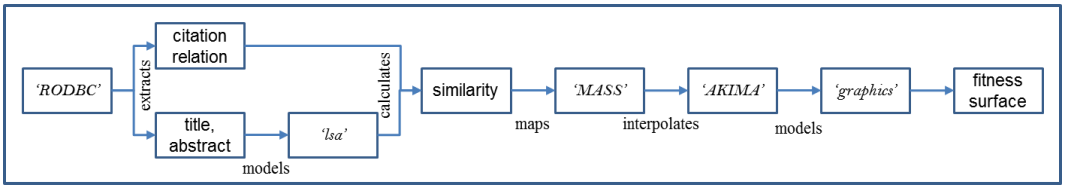
Software Item	IP Description
S-4-3	**C-4-3: Software for Overlapped Time Frame Partitioning and Parallel Application of LDBSCAN** Prototype software that is a parallelized program implementing LDBSCAN; a locality-aware density-based cluster algorithm that improves on the classical DBSCAN via incorporation of parallelization.  Figure 5.1 Overlapped Time Frames Partitoning
	P-4-3: An improved method of database scanning via parallelization which generates more time efficient and robust scanning as compared to other methods.

Software Item	IP Description
S-4-4	**C-4-4: Software for Data Pruning and Load-Balanced Parallelization of ST-LOF Calculation by Randomized Partitioning.** Prototype software that detects outliers by assigning a score for every object in the data set. It handles spatio-temporal outliers which are based on the spatio-temporal context of each object. The software performs multiple rounds, in each it randomly creates partitions of the data sets and runs the ST-LOF module on the partitioned objects. Objects that are definitive inliers are pruned. Results of each iteration are merged with previous ones. The whole loop is repeated a specified number of times depending on what level of accuracy is desired. All partitions can be run in ST-LOF in parallel.

	
	P-4-4: A novel process of data pruning and load-balanced parallelization of outlier calculation by randomized partitioning.

CVDI Project 5 - Scalable Visualization, Gap Analytics for Multiple Big Data Industry Sectors

Software Item	IP Description
S-5-1	**C-5-1: Temporal MDS Map Generator** (See Project Report, Figs 4 and 7) A software prototype capable of generating a temporal MDS map to investigate a company's intellectual and competitive intelligence. *Input*: Citation information and contents (title and abstract) from the patent database *Process*: As displayed below  *Output*: Temporal competitive trajectory in industrial intelligence *Languages/technologies*: R / packages 'RODBC', 'lsa', 'MASS', 'scatterplot3d', and 'rgl'
	P-5-1: A novel method that combines state-of-the-art techniques adapted to capture the temporal dynamics in competitive intelligence surveillance.

Software Item	IP Description
S-5-2	**C-5-2: Fitness Maps for Investigating IP and Competitive Intelligence** (See Project Report, Figs 12-14) A software prototype capable of landscaping a fitness map to investigate a company's intellectual property and competitive intelligence. *Input*: Citation information (including a patent's cited times) and contents (title and abstract) from the patent database *Process*: As illustrated below  *Output*: Fitness surface map in industrial intelligence *Languages/technologies*: R / packages 'RODBC', 'lsa', 'MASS', 'AKIMA', and 'graphics'
	P-5-2: A novel method that facilitates rapid business analysis by combining state-of-the-art algorithms to generate an industrial fitness landscape derived from non-geospatial data.

Software Item	IP Description
S-5-3	**C-5-3**: **Circle Pack Data Visualization Technique (See Quarterly Report, Figure 5)** A software prototype capable of generating an interactive Circle pack visualization of the hierarchical structure of entire US patents by International Patent Classification. The prototype could be extended to other dataset with similar hierarchical structure. The prototype is a web-based application and is configured to process large datasets (around 5 million records) stored as a CSV file. *Languages/technologies*: PHP, MySQL, D3.js
	P-5-3: A system capable of generating a Circle Pack visualization of large datasets having hierarchical structure.

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-2 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-2 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 4, 2014</u>** by e-mail to the senders 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
Drexel University

C. Dean Domingue
Director, Office of Innovation Management
University of Louisiana at Lafayette

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 2 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 - Semantic Information Extraction, Integration, and Visualization for Big Data Analytics

Software Item	IP Description
S-1-1	**C-1-1**: **Prototype for Extracting Structured Knowledge (Gene Expression Relations) from Unstructured Text** A software prototype capable of extracting gene expression relations between genes and brain regions from biomedical literature. *Languages/technologies*: Stanford Natural Language Processing toolkit, MetaMap, SemRep, MultiR
	P-1-1: A method of extracting gene expression relations between genes and brain regions from biomedical literature that does not need manually labeled training examples for learning an extraction model.
	P-1-2: A method of extracting gene expression relations between genes and brain regions from biomedical literature that leverages the results of existing open information extraction systems and combines them with a distant learning method.
	P-1-3: A method of extracting gene expression relations between genes and brain regions from biomedical literature that takes into consideration of the interdependency characteristic of domain entities. With grouping strategy, the method achieves improved results compared with baselines at both sentential level and corpus-level.

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 2 Intellectual Property Report: CVDI Project 2

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 2 - Large-scale Social Media Analytical Tools with Application to Detecting Emerging Events

Software Item	IP Description
S-2-1	**C-2-1: Prototype for Performing Simultaneous Aspect and Sentiment Generation** A software prototype that studies the public opinions towards an entity using a hybrid HDP-LDA model, which can automatically determine the number of aspects, distinguish factual words from opinioned words, and further effectively extracts the aspect specific sentiment words, and performs sentiment analysis more efficiently than other influential models such as JST, AUSM, and MaxEnt-LDA. *Languages/technologies*: Java
	P-2-1: A method that includes an unsupervised LDA-based model adapted not to require the number of aspects to be specified in advance, making the system more objective.
	P-2-2: A method that includes an unsupervised LDA-based model adapted to distinguish between factual words and opinioned words.
	P-2-3: A method that includes an unsupervised LDA-based model configured to extract aspect-specific sentiment words, instead of considering an entire sentence to be of one aspect.

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

Software Item	IP Description
S-2-2	**C-2-2: Enhance Topic Evolution – Prototype for Learning Hierarchical Topic Models from Short and Noisy Social Media Messages** A software prototype that uses a new generative probabilistic model called Hash-Based Stream LDA for topic modeling in continuous social media streams, which outperformed the classical LDA approach and the stream-oriented On-line LDA and SparseLDA approaches in terms of average perplexity by more than 12% when using data crawled from Twitter, and by 21% when using data crawled from the IRC chat community. *Languages/technologies*: Java
	P-2-4: A topic modeling system having a generative probabilistic model configured to expose facilities to include inter-document similarity in topic modeling.
	P-2-5: A topic modeling system having an inference algorithm that relies on efficient estimation of document similarity with Locality Sensitive Hashing to retain the knowledge of past social discourse in a scalable way.
	P-2-6: A novel topic modeling system that utilizes historical knowledge of previous messages in inference to improve quality of topic discovery.

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

Software Item	IP Description
S-2-3	**C-2-3 - Affinity propagation (AP) clustering software prototype on Hadoop** A software prototype capable of clustering large text data efficiently and effectively with a distributed affinity propagation mechanism on Hadoop/MapReduce. *Languages/technologies*: Java, Hadoop/MapReduce
	P-2-7: A system of data clustering using Affinity propagation (AP) among data points, redesigned for distributed computation and parallel text clustering on Hadoop.
	P-2-8: A method of partial (pruned) affinity propagation (AP) which reduces time complexity of text clustering and makes the algorithm efficient even with large amounts of text data.
	P-2-9: A method of selecting a balanced value of k, the number of pair-wise similarities to retain in Pruned AP computation, leading to optimal clustering performance (inflection point).

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

Software Item	IP Description
S-2-4	**C-2-4** - **Emerging Event Detection** A software prototype that improves on the original Event Detection on Onset (EDO) that incorporates data from multiple streaming/social medias sources and is now capable of detecting sub-events.
	P-2-10: A method of detecting emerging events by combining data from multiple different media sources and is capable of yielding sub-event detection.

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 2 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 - Visual Analytic Approaches to Mining Large-scale Time-Evolving Graphs

Software Item	IP Description
S-3-1	**C-3-1: Data Mining Tools for Cluster and Association Mining** Prototype software that implements well-know and vetted hierarchical clustering, and association mining to understand relationships between various entities esp as it relates to risky customers, and commercial product dependency

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

Software Item	IP Description
S-3-2	**C-3-2: Two-phase Forecasting System from Non-linear Time Series Graphs** Prototype software that generates a prediction model for non-linear time-evolving graphs and incorporation of a novel two-phase forecasting system.
	P-3-2: A novel method of incorporating inter-proximity between local forecasting to adjust each individual local forecast.

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

Software Item	IP Description
S-3-3	**C-3-3: Web-based Visual Analytics Dashboard** Prototype Software that generates a web-based dashboard where a user can get generate and view multiple data visualizations

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

Software Item	IP Description
S-3-4	**C-3-4: Symbolic Time Series to Compute Situational Time Lags** Prototype software that implements symbolic time series based time lag computation whereby converting the normal numeric time series into symbolic or categorical time series, and adjusting time lag per category.
	P-3-4: A novel method of symbolic time series based time lag computation whereby converting the normal numeric time series into symbolic or categorical time series, and adjusting time lag per category. In state of the art or most existing techniques the time lag between the time series is computed using cross correlation between the numeric values. The main disadvantage of this process is that we are going to get a single time lag between these two time series. Using a symbolic time series permits extraction of different time lags at different situations, categories.

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

Software Item	IP Description
S-3-5	**C-3-5: Multi-touch Graph Interaction** Prototype software where multi-touch graph interaction allows users to visualize, explore data. While building on the Handymap application, graph interaction expands beyond the need for geo-located datasets.
	P-3-5: An improved method for user interface selection for cluttered VR environments using a tracked hand-held touch device.

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 2 Intellectual Property Report: CVDI Project 4

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 4 - A Spatio-Temporal Data Mining Approach for Fraud Detection

Software Item	IP Description
S-4-1	**C-4-1: Software for Spatio-temporal local outlier factor (ST-LOF) Calculation** A software prototype that incorporates spatio-temporal context in to the Local Outlier Factor (LOF). ST-LOF defines a spatio-temporal outlier to be a spatial-temporal object whose thematic attribute values are significantly different from those of other spatially and temporally referenced objects in its spatial or/and temporal neighborhoods. This software has a user interface for the user to enter the k-th nearest neighbor, number of neighboring points to consider and the normalized input file. As output, each object in the data set is assigned a ST-LOF outlier score and written to s file.
	P-4-1: A novel method that improves on LOF by introducing spatio-temporal context to every data object and thereby capturing outliers or clusters based on spatio-temporal difference.

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

Software Item	IP Description
S-4-2	**C-4-2: Software for Spatio-temporal Local Density Based Clustering of Applications with Noise (ST-LBDSCAN)** Prototype software that incorporates spatio-temporal context into the Local Density Based Spatial Clustering of Applications with Noise (LDBSCAN). ST-LBDSCAN algorithm detects outliers based on spatio-temporal context and clusters the objects that have similar features spatially or/and temporally. The software has a user interface for the use to enter the k-th nearest neighbor, number of neighboring points to consider in ST-LOF and ST-LBDSCAN, local density reachability percentage and a normalized input file. The output is a file with ST-LOF score along with its dataset object ID and its cluster identity (cluster value of -1 impies outlier).
	P-4-2: A novel method that improves on LDBSCAN by introducing spatio-temporal context to every data object and thereby capturing outliers or clusters based on spatio-temporal difference.

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

Software Item	IP Description
S-4-3	**C-4-3: Software for Overlapped Time Frame Partitioning and Parallel Application of LDBSCAN** Prototype software that is a parallelized program implementing LDBSCAN; a locality-aware density-based cluster algorithm that improves on the classical DBSCAN via incorporation of parallelization.
	P-4-3: An improved method of database scanning via parallelization which generates more time efficient and robust scanning as compared to other methods.

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

Software Item	IP Description
S-4-4	**C-4-4: Software for Data Pruning and Load-Balanced Parallelization of ST-LOF Calculation by Randomized Partitioning.** Prototype software that detects outliers by assigning a score for every object in the data set. It handles spatio-temporal outliers which are based on the spatio-temporal context of each object. The software performs multiple rounds, in each it randomly creates partitions of the data sets and runs the ST-LOF module on the partitioned objects. Objects that are definitive inliers are pruned. Results of each iteration are merged with previous ones. The whole loop is repeated a specified number of times depending on what level of accuracy is desired. All partitions can be run in ST-LOF in parallel
	P-4-4: A novel process of data pruning and load-balanced parallelization of outlier calculation by randomized partitioning.

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

Name _____ Signature _____ Date _____

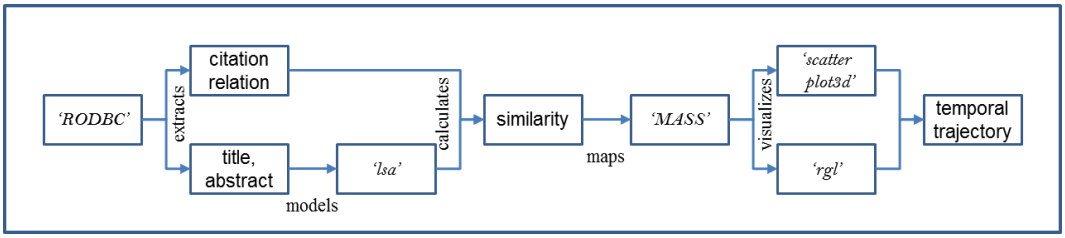
Drexel University
Office of Technology Commercialization
The Left Bank Building
3180 Chestnut Street
Suite 104
Philadelphia, Pennsylvania 19104

Re: CVDI Year 2 Intellectual Property Report: CVDI Project 5

Dear Dr. McGrath:

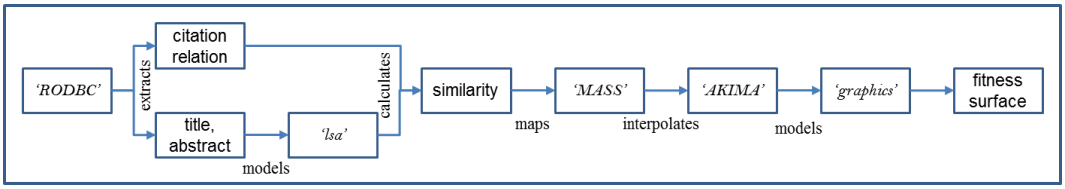
Please see the below table, where we have indicated our licensing interests for the IP assets associated with the subject CVDI project:

CVDI Project 5 - Scalable Visualization, Gap Analytics for Multiple Big Data Industry Sectors

Software Item	IP Description
S-5-1	**C-5-1: Temporal MDS Map Generator** (See Project Report, Figs 4 and 7) A software prototype capable of generating a temporal MDS map to investigate a company's intellectual and competitive intelligence. *Input*: Citation information and contents (title and abstract) from the patent database *Process*: As displayed below  *Output*: Temporal competitive trajectory in industrial intelligence *Languages/technologies*: R / packages 'RODBC', 'lsa', 'MASS', 'scatterplot3d', and 'rgl'
	P-5-1: A novel method that combines state-of-the-art techniques adapted to capture the temporal dynamics in competitive intelligence surveillance.

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

Software Item	IP Description

S-5-2	**C-5-2: Fitness Maps for Investigating IP and Competitive Intelligence** (See Project Report, Figs 12-14) A software prototype capable of landscaping a fitness map to investigate a company's intellectual property and competitive intelligence. *Input*: Citation information (including a patent's cited times) and contents (title and abstract) from the patent database *Process*: As illustrated below  *Output*: Fitness surface map in industrial intelligence *Languages/technologies*: R / packages 'RODBC', 'lsa', 'MASS', 'AKIMA', and 'graphics'
	P-5-2: A novel method that facilitates rapid business analysis by combining state-of-the-art algorithms to generate an industrial fitness landscape derived from non-geospatial data.

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

Software Item	IP Description
S-5-3	**C-5-3**: **Circle Pack Data Visualization Technique (See Quarterly Report, Figure 5)** A software prototype capable of generating an interactive Circle pack visualization of the hierarchical structure of entire US patents by International Patent Classification. The prototype could be extended to other dataset with similar hierarchical structure. The prototype is a web-based application and is configured to process large datasets (around 5 million records) stored as a CSV file. *Languages/technologies*: PHP, MySQL, D3.js
	P-5-3: A system capable of generating a Circle Pack visualization of large datasets having hierarchical structure.

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

Name _____ Signature _____ Date _____