CHARTER RENEWAL APPLICATION – SIAG/SDM (Data Mining & Analytics)

This CHARTER RENEWAL APPLICATION applies to the SIAM Activity Group on Data Mining & Analytics. The SIAG/SDM was originally formed under the aegis of SIAM on July 17, 2011 by the SIAM Council and by the SIAM Board of Trustees. Its initial operating period began July 17, 2011 and ends December 31, 2013. This is the first charter renewal to be renewed by the council and board thereafter. This SIAG had 261 members as of December 31, 2012; of these, 158 were students.

It is the purpose of the SIAM Activity Group on Data Mining to advance the mathematics of data mining, to highlight the importance and benefits of the application of data mining, and to identify and explore the connections between data mining and other applied sciences. Within the framework of SIAM, the SIAG will conduct activities that implement its purposes.

The SIAG on Data Mining and Analytics will organize activities in Data Mining. The SIAG is expected to:

1. Organize minisymposia at the SIAM Annual Meeting.

2. At least once every five years, either organize a track of at least six minisymposia at the SIAM Annual Meeting or have an activity group meeting held jointly with the annual meeting. The VP for Programs and the VP at Large will coordinate the scheduling with the SIAG chair.

3. Organize a yearly SIAM International Conference on Data Mining (SDM). The program director and chairperson of the SIAG are responsible for ensuring the continuity of SDM, including either serving as co-chairs of the meeting or naming appropriate designees. The organizing committee must be approved by the VP for Programs at least 16 months before the conference.

The SIAG has complemented SIAM's activities and supported its proposed functions. The answers to the questions below indicate how this was accomplished and what the officers propose as the future directions for the SIAG.

1. How is the field covered by the activity group doing? Is it growing, is the focus shifting? What have been the significant advances over the last [two/three] years?

The field of data mining and analytics has seen a tremendous increase in recent years as witnessed by the ubiquitous use of "Big Data" as the technology which will change everything as we know it, leading to better decisions in fields as diverse as medicine, finance, business, and the Internet. The field is growing, prompted not just by the increasing size and complexity of the data being collected, but also the positive results from its use in real applications (e.g., medicine and Google's self-driving car) and the realization that semi-automated techniques will be necessary for us to exploit the vast amounts of data we have collected. Some of the more significant advances in the recent past include the development of parallel tools such as Hadoop for analyzing vast amounts of data, the increasing deployment of automated tools in fields such as astronomy and medicine, the mining of data from multiple sources (e.g. text, images, and video on the web), and the development of new algorithms in areas such as streaming data, complex networks (e.g., social networks and the web), and cyber security.

2. How is the activity group doing? Is it remaining vibrant? Is the size of the SIAG stable or increasing? How is the SIAG keeping up with the changes in the field? How are the broader interests of SIAM reflected in the activities of the SIAG?

This is the first term of the SIAG, so it is still growing. The SIAG is keeping up with changes in the field by organizing panels, workshops, and tutorials on current topics at the SIAM Data Mining conference. These have included healthcare, crowdsourcing, social networks, Big Data, etc. Topics such as "Recent Advances in Matrix Technologies" reflect the broader interests of SIAM. We also held a mini-symposium during SDM12 on "Opportunities for Collaboration," jointly with the SIAG on Uncertainty Quantification.

3. Please list conferences/workshops the activity group has sponsored or co-sponsored over the past two years, and give a brief (one sentence or phrase) indication of the success or problems with each.

In 2012, the International Conference on Data Mining was held in Anaheim California, April of 2012 with 280 attendees.

The 2013 SIAM International Conference on Data Mining (SDM13) was held in Austin Texas on May 2-4, 2013. Attendance was down due to an unfortunate delay in receiving our NSF funds for student travel due to budget issues in Washington.

The SAMSI computational advertising program was held in cooperation with SIAG/DMA.

SDM, despite being a relatively new conference in data mining, has established itself as a premier meeting for the mathematical aspects of data mining. This is reflected in its ranking among the conferences in the "data mining" category on MSR academic search:

	#papers	#cites	#self-cites	year-start	year-end	avg cites	avg cites excl self
journal DMKD	584	19729	403	1994	2012	33.78253	33.09247
conf KDD	2062	69575	3140	1988	2011	33.7415	32.21872
journal TKDE	2747	61026	2040	1987	2012	22.21551	21.47288
conf ICDE	4016	67457	3915	1984	2012	16.79706	15.82221
conf SDM	**709**	**9101**	**228**	**2000**	**2011**	**12.83639**	**12.51481**
conf CIKM	2639	28624	1155	1992	2011	10.84653	10.40887
conf PKDD	1000	8881	222	1997	2011	8.881	8.659
journal KAIS	743	7682	1403	1998	2012	10.33917	8.450875
conf ICDM	2513	18401	1126	2000	2011	7.322324	6.874254
journal TKDD	111	721	9	?	2012	6.495495	6.414414
conf PAKDD	1253	6412	216	1996	2011	5.117318	4.944932
conf DASFAA	1251	4003	176	1989	2013	3.19984	3.059153

4. Please indicate the number of minisymposia directly organized by the activity group at the last two SIAM Annual Meetings. When did the SIAG last organize a track of minisymposia at an annual meeting?

SIAG/DMA is relatively new, this being its first term. We have not yet had the opportunity to organize a mini-symposium at an Annual Meeting. However, three of the keynote speakers at SIAM AN13 (Braverman, Santer, and Zack) were recommended by SIAG/DMA officers and there will be two minisymposia complementing these talks. These are all focused on the data aspect of the Math of Planet Earth. These could be considered as participation of SIAG/DMA at the Annual Meeting, though these activities were not explicitly identified as such.

5. Please indicate other activities sponsored by the activity group, to include newsletters, prizes and web sites. Have each of these been active and successful?

Since this is the first term of the SIAG, the focus has been on getting new members and ensuring a stable organization of SDM in the years ahead. It is expected that other activities will be sponsored by SIAG/DMA in the future.

6. What activities are planned and proposed for the next period of the charter? Please describe scheduled and suggested future activities in detail.

We are looking into having the first ever SDM conference outside the US in 2015 (Vancouver is one of the cities being considered). This was done at the request of several SDM participants over the last several years.

We would also like to explore mini-symposia at SDM with other relevant SIAGs, such as the ones on linear algebra, mathematical aspects of materials science, discrete mathematics, optimization, imaging science, and uncertainty quantification.

7. How can SIAM help the activity group achieve its goals?

SIAM currently provides excellent support in organizing the SDM conference. In future, we would like to explore better ways of reaching out to members, though wikis or social media. Since there will be a new set of officers in 2014, this could be something they could address.

8. How can the activity group help SIAM in its general role of promoting applied mathematics and computational science?

The activity group, through the organization of the SDM conference, helps SIAM to connect with participants interested in the mathematical aspects of data mining and the use of data mining in solving real problems. This is perhaps the major, if not the only, venue for reaching this community and making them aware of other related and relevant fields of applied and computational mathematics, such as those covered by the list of relevant SIAGs in Question 6. Since there are a relatively large number of industry participants (from Google, IBM Research, UTRC, and some of the smaller startups) at SDM and in the SIAG, it is also an opportunity for SIAM to reach those in industry who work in applied areas of mathematics.

This SIAG requests that the SIAM Council and Board of Trustees renew its charter for a two-year operating period beginning January 1, 2014.

Signed
Chandrika Kamath.

Chandrika Kamath, SIAG Chair
June 7, 2013
