

<< Workshop proposal of **MoDAT**>>

MoDAT: Designing the Market of Data
- for Synthesizing Data in Sciences and Businesses -

ICDM 2013, IEEE International Conference on Data Mining, Dallas, Texas / one day of December 8-11, 2013.

<<Title>> **Designing the Market of Data**
(subtitle) - for Synthesizing Data in Sciences and Businesses -

<<Acronym>> **MoDAT**

<<Duration (full-day or half-day)>> Full day

<<Scope>> This workshop is not about data mining for marketing, but about how to create and design the market where data are reasonably dealt with, i.e., sold, opened free, or shared on negotiation. Our ultimate goal is to have each people on the earth feel free to share one's own data with others without fearing the loss of business opportunities.

In order to make a social environment where analysts and decision makers in active businesses and sciences can be provided with data they need, we discuss how to (re)design an environment called the Market of Data, where each user or provider of data can understand the value of each part of data so that one can buy/sell it in a reasonable condition, e.g., for a reasonable price. Here, the value of each part of data shall be visualized to aid users' considering its possible contribution to promoting/creating businesses and scientific findings.

Furthermore, data scientists often need to import mining techniques from others, but the techniques are not easy to learn from experts dealing with different kinds of data, because the similarity between data is not always obvious. The similarity between latent dynamics behind data is also desired to be visualized in such a way as computing distances among features of data for aiding the choice of success/failure cases of data mining to learn from.

For this workshop **we call for presentations and discussants** about what we can/should do for creating a marketplace where data and analysts' knowledge are shared, with reasonably determining the conditions for sharing. People in the market may sell/buy data for agreed prices, or communicate to decide to expose the data free i.e. as open-source, if the social responsibility of the data provider comes to be regarded as more important than the income for selling data. Thus the Market of Data (MOD) means a place where the value of data and knowledge are externalized. Relevant areas are (not restricted to):

Data/Text mining and visualization

- > Visualization of links among data, representing the possibility to combine them to discuss use scenarios of data
- > Visualization of links and distances among data, representing their similarities
- > Mining data or text for finding important events and attributes, in order to compute the links and distances
- > Extracting causalities, for externalizing links among data

Knowledge representation

- > Construction of dictionaries of variables, for reasonably linking among data
- > Representing the hierarchical structure of relevance among concepts and variables, used in the thoughts of analysts and users

Methods for creative communication

- > Data-based communication for evaluating the value of an event, i.e., chance discovery, and data which may include such an event
- > Visual interface for triggering meaningful thoughts of stakeholders

We also love to involve ones out of the data mining community – sociologists, stock dealers, biologists, .anyone from active communities are in MoD. Forming the interdisciplinary community, **we will conduct a gaming session of Innovators Marketplace®** where ideas to combine existing data and knowledge are reflected to the design of MoD.

<<How the organizers plan to attract quality submissions>>

- (1) We will mainly call for participants electronically, via
 - > Mailing lists of members who took part in workshops on data mining and relevant areas (marketing, product design, finance engineering, etc, and 1000 readers of the mail magazine of chance discovery, the Uncertainty AI mailing list, the ML of Japanese Society of AI, etc).
 - > Research Communities on marketing, management science, and other social sciences
- (2) We already started to call for interests in this topic by including slides and PR papers in symposia and workshops (e.g., AAI Spring Symposia 2013, at Stanford University).

<<Tentative Candidate of Keynote: contacted and agreed tentatively>>

<Speaker> Hiroe Tsubaki,

**<Keynote title> Application of Statistical Methods to Multidisciplinary Risk Evaluation
- foreseeing the way of data synthesis via analysis in the market of data -**

<<Program Committee: Tentative list, under expansion>>

Yukio Ohsawa, The University of Tokyo, Japan

Akinori Abe, Chiba University, Japan

Ajith Abraham, Machine Intelligence Research Labs, USA

Peter Bruza, Queensland University of Technology, Australia

Kuiyu Chang, Nanyang Technological University, Singapore

Koiti Hashida, The National Institute of Advanced Industrial Science and Technology, Japan

Chao-Fu Hong, Aletheia University, Taiwan

Huan Liu, Arizona State University, USA

Christel De Maeyer, Vrije Universiteit Brussel, Belgium

Dirk Van den Poel, Ghent University, Belgium

Dominik Slezak, Warsaw University, Poland

Hideaki Takeda, National Institute of Informatics, Japan

Henry Wang, Chinese Academy of Sciences, China

Michael Welge, University of Illinois at Urbana-Champaign, USA

Katsutoshi Yada, Kansai University, Japan

<<Contact: Organizers>>

Yukio Ohsawa

Professor, Dept. of Systems Innovation, School of Engineering,
The University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8583, JAPAN

Tel: +81-3-5841-2908

Email: (Ohsawa and Secretary) info@panda.sys.t.u-tokyo.ac.jp

Akinori Abe

Professor, Faculty of Letters, Chiba University

1-33 Yayoicho, Inageku, Chiba 263-8522, JAPAN

Email: ave@l.chiba-u.ac.jp