



Minutes

1st PetaMedia SIG4 Meeting "Social Content Retrieval"

Delft, 23 September 2008

Mission

SIG4 is probably the most critical SIG within PetaMedia, as it draws relatively little from the previous research of the PetaMedia consortium and addresses entirely new challenges in the field of multimedia content access in the social/P2P network context. Successfully addressing these challenges requires out-of-the-box ideas and an intensive and broad interdisciplinary discussion that we hope to initiate during this meeting. The mission of this meeting is to come up with a clear set of research directions that will mark the PetaMedia contribution regarding social content retrieval. We will approach this mission by addressing the following main objectives:

Identifying R&D challenges falling thematically into SIG4,
Identifying the existing/planned research efforts across PetaMedia addressing these challenges, and filtering out those involving a sufficient number of partners
Creating a first set of ideas for mini-projects

Agenda

9:30 – 10:00 **Arrival & coffee**
10:00 – 10:10 **Welcome note and meeting overview (A. Hanjalic)**

Invited talk

10:10 – 10:40 Philips Perspective on Social Content Retrieval (M. Barbieri)

10:40 – 10:50 **Short coffee break**

Plenary technical session

10:50 – 11:00 SIG4-related research at Univ. Geneva (J. Kierkels/T. Pun)
11:00 – 11:10 SIG4-related research at TUB (N. Neubauer)
11:10 – 11:20 SIG4-related research at Univ. Yverdon (L. Raileanu)
11:20 – 11:30 SIG4-related research at Univ. Glasgow (I. Konstas)
11:30 – 11:40 SIG4-related research at EPFL (J.-S. Lee)
11:40 – 11:50 SIG4-related research at TUD (M. Clements,
A.P. de Vries,
H. de Ridder
J. Broekens)

11:50 – 13:00 **Lunch**

13:00 – 13:05 **Introduction to the Creativity Boosting Game**
13:05 – 14:45 **Creativity Boosting Game**

14:45 – 15:15 **Coffee/Tea break**

15:15 – 16:30 **Plenary discussion**

16:30 **Closing of the meeting**

19:00 **Social event (Dinner)**

Minutes

Presence list:

Jong-Seok Lee	EPFL (CH)
Nicolas Neubauer	TUB (D)
Laura Raileanu	University of Yverdon (CH)
Ioannis Konstas	University of Glasgow (UK)
Joep Kierkels	University of Geneva (CH)
Thierry Pun	University of Geneva (CH)
Mauro Barbieri	Philips Research (NL)
Wessel Kraaij,	TNO (NL)
Arjen De Vries	TUD/CWI (NL)
Maarten Clements	TUD (NL)
Huib de Ridder	TUD (NL)
Alan Hanjalic	TUD (NL)
Pien Rijnink	TUD (NL)

Welcome note and meeting overview (A. Hanjalic)

(for details please see the introductory presentation at www.petamedia.eu)

- Brief introduction to PetaMedia mission and partnership concept
 - Explanation of SIGs with emphasis on SIG4
 - The concept of mini-projects and the rules how to define and finance them
 - Presentation of the mission of the meeting
 - Presentation of the general timeline for setting up, submitting and approving mini-projects
-

Invited talk: *Personal Television Channels* (M. Barbieri)

(for details please see the presentation at www.petamedia.eu)

- Vision on the consumer (C) and market (M) needs for TV content consumption
 - o Good balance between freedom from choice/to choose (C)
 - o Individualization/Personalization (C)
 - o Ease of use (C)
 - o Value for money (C)
 - o Be user's preferred entry to available content offerings (M)
 - o Targeting of content and commercials (M)
 - o Effective measurements (M)
 - o Low cost of operation (M)
- APRICO & Personal TV
 - o Philips corporate venture led by IP&S
 - o Objective: Bring the right content seamlessly to the screen!
 - o Concept of "Personal TV Channel" proven successfully in extensive field trials
- Research directions
 - o Recommender technology
 - o Internet content
 - o Social aspects (social recommendations, privacy, new forms of advertising)
- Internship places available on topics related to (social) recommender systems

Plenary technical session

The objective was to generate an overview of SIG4-related research and possible contributions of the partners. All presentations can be downloaded via the PetaMedia website (www.petamedia.eu). The main items discussed per presentation are listed below:

Presentation 1 (T. Pun)

- Cross-modal information retrieval and management
- Semantic information discovery and propagation
- Multimedia collection exploration
- Emotion assessment for implicit tagging
- Affective characterization of movie scenes (affective profiling)
- Social content retrieval: affective profile propagation, user groups based on taste

Presentation 2 (N. Neubauer)

- Click-log data for Tribler
- Tagging in virtual worlds
- Machine learning for tag processing
- Image search by relevance feedback

Presentation 3 (L. Raileanu)

- Knowledge enrichment in a collaborative multimedia tagging system
 - o Enrich existing tags
 - o Browsing/retrieval based on users relationships and resource popularity
- Strong emphasis on platform implementation!

Presentation 4 (I. Konstas)

- Retrieval algorithms and models
 - o Approach to multimedia retrieval using the traditional IR concepts
- Mining media
 - o Automatic image/video annotation
 - o Event mining
 - o Adaptive search models
- Search interfaces
 - o Ostensive browsing scheme
 - o Explorative search interface for video search
 - o Affective search interfaces for IR

Presentation 5 (J.-S. Lee)

- Visual attention model as link between content analysis and tags
- Content tags vs. context tags
- Improving annotation by modeling user preferences
- Scene-object-mood paradigm

Presentation 6 (Maarten Clements)

- Social content systems: enhancing the distribution and retrieval of content through exploiting social aspects of networks (e.g. by modeling friendship and trust among people)
- Taste buddies versus social network friends
- Taking into account affective user profiles for content recommendation
- Implicit rating by the "Query&view" paradigm: is it as good as explicit rating (tags)?

Presentation 7 (Huib de Ridder)

- Product intelligence
 - o Embedding, context awareness, personalization, adaptivity, anticipation
 - o From user-friendliness to human-centered design
 - o Transparent interfaces
- Bridging people through technology
 - o Paradigms: Social distance, user profiling, power of collectivity, social visibility
 - o Social networks: trust and incentives for cooperation, social exchange theory, social distance or degrees of separation, reciprocity versus mutuality

Presentation 8 (Wessel Kraaij)

- Content distribution (IPTV, standardization)
- Media mining (text mining, metadata extraction, recommender systems)
- "ConnecTV" (extensive field trials performed!)

Creativity boosting game

In order to converge towards the definitive scope of activities within the SIG4, to check how the research at different partner institutions maps onto this scope, and to investigate which research challenges can be identified to be approached jointly within SIG4 using the concept of mini-projects, a brainstorm session is organized in the form of a game with the following rules:

- "Speeddating" concept
 - Playing in pairs, random pair composition
 - 10 minuts per round
 - Players "interview" each other on what they did/do/will do on research, which projects that are involved in, etc.
 - Get inspired by each other's research and think of as many ideas as possible related to social content retrieval
 - Write down ideas on yellow notes
 - Winner: the player with the largest number of different (uncorrelated!) ideas
-

Plenary discussion

In this session, all the ideas that emerged from the creativity boosting game were collected, plenary discussed and assigned to different clusters, with the objectives to

- (a) identify the scope of the SIG4 based on the obtained clusters
- (b) define a number of ideas for mini-projects to be realized in SIG4 based on the activities envisioned per cluster

For each of these ideas, the original planning was to generate a preliminary list of partners involved (based on those partners present in the meeting), and to select a coordinator. Once the idea is worked out, it was planned to be distributed by the coordinator via the SIG4 mailing list to investigate whether

1. there are similar efforts planned by some other SIG4 partners
2. there are partners elsewhere in PetaMedia interested to join the project
3. there are adjustments needed
4. the project could/should be merged with another project idea proposed elsewhere in PetaMedia.

Based on the discussion within SIG4, project ideas can be finalized and formulated in a more detailed fashion, specifying the challenge to be addressed, expected results, positioning within PetaMedia (links to other mini-projects, relevant SIGs, timing), the team, the coordinator and the funding required from PetaMedia. This proposal will then be submitted to the PetaMedia SB for approval and will be presented during the plenary meeting in Berlin. If approved, the ideas can be further adjusted based on the discussion during the plenary meeting and made ready for realization.

Based on the results of the creativity boosting game and the discussions around the proposed mini-projects, it was agreed to adjust the planning above regarding some of its elements.

Summary of the results:

The following clusters of research ideas/issues/challenges emerged from the creativity boosting game were defined ¹:

1. Social Network

- How to stimulate people in a social network to behave altruistic?
- How to combine user/use data and profiles if a user participates in different social networks?
- Use physiological responses for comparing users
- Explicit vs. affective user groups
- Finding Emotional buddies (e.g. Finding other people with water-fear)
- How to create and evaluate common profiles?
- How to create and how to spot a trend in a social network?
- Merging / composing information on user groups from explicit input data
- How to validate the constitution of a group?

¹ **Note:** some of the clusters and ideas per cluster were identified as out-of-scope for SIG4. These clusters and ideas were discarded and are not listed here. Furthermore, some cleaning, fine-tuning and/or combining of (related) ideas was applied where required.

Privacy → what to exchange with friends, friend of friend, ...?
What is a social network and social distance?
Cross-over between social networks in real and virtual world
How to use egoistic users for the good of the system?
Role of social content appreciation and ratings

2. **Social (Collaborative) Tagging**

Tag propagation for "smoothed" taste & expertise estimation in social networks
Resolving ambiguities in local implicit tagging using the social network context (inferring tags from more than one user, establishing relations of implicit tags of different users for the same content)
Enhancing distributed tag sets (and delete spam tags?): associations rules, relationship between tags (via vocabulary mapping, synonyms, thesaurus)
Tag set mapping to different abstraction layers, quality label for tags
Use of gaming to create groups and create/filter/weigh tags
Annotation incentives and a natural and intuitive user-system interaction environment needed for a successful game scenario
Use of "negative" tags. What is the content 'not' about?
What do users input for retrieval: a specific content description or an abstract content category?
Opinion mining from descriptive tags
Classifying tags as objective vs. subjective e.g using information about taggers
Success of del.icio.us is 'bottom-up'. Can a successful tagging system be designed 'top-down'?

3. **Dynamics of social networks and tags**

Use of context (How to model it? How to benefit from it?)
Context-dependent social network model (age, intellectual background)
Temporal validity of social relations ("I'm not your friend any more")
Temporal/context-dependent validity of tags for one and the same content

4. **Social/device network enhanced user profiling and content filtering/recommendation**

Propagate/ update profiles according to users relationships
Integrate social information into recommendation system
Predict user emotion based of affective profile + multimedia features
Cross-domain recommendations
Combine different platforms → cross-domain recommendation
Evaluation of a social recommendation system
Integrating affective content analysis and collaborative filtering to enhance each other for optimizing individual profiles in a social network
Integration of local/network resources (semantic inference at different nodes for different queries/profiles/contexts, synergy between networked MCA and distributed tag sets) to improve automated content indexing

5. **Interfaces**

Use physiological features to validate other implicit and explicit feedback
Natural and intuitive in-video tagging
Advanced tags suggestions, dialogue-based
Beyond keyword-based search and access of online video
Tag clicking to support relevance detection, validated using EEG or BCI

The following list of project ideas emerged from the clusters listed above and the related plenary discussion:

1. Social network analysis/modeling (PI1)²

- Coordinator: TUD (De Vries)
- Covers the issues related to the mechanisms underlying the fundamental content exchange and collaboration processes in a social network. Examples of issues are incentives for tagging and collaboration, relationship: reciprocity versus mutuality, friendship and trust, degree of separation, etc.
- Current participants: Yverdon, a number of the participants from the Netherlands

2. Network-enhanced tagging and user2user relations (PI2)

- Coordinator: TBD
- Current participants: TUD (De Ridder), Geneva, Glasgow
- Covers the issues resulting from expanding the ideas from implicit tagging onto the social network context. Also falling into this scope is knowledge enrichment in a collaborative multimedia tagging system. Examples of issues are explicit tags enhancement in a network context, implicit tags enhancement (e.g. via implicit tags of multiple users or via explicit tags) and propagation, the concept of "emo buddies", etc.

3. Data acquisition and use scenario definition project (PI3)

- Coordinating team: TUD (De Vries), Geneva (Pun), Yverdon (Raileanu)
- Envisioned participants: Peta-Media wide, representatives of all SIGs
- This project is identified as the basis of all other mini-projects, not only in SIG4 but also PetaMedia wide. For the project rationale and directions for shaping the project, please see the discussion below.

NOTE: It is worth mentioning that a mini-project addressing the cluster 5 (Interfaces) was defined during the Dutch PetaMedia Branch (DPB) meeting, Delft, 22 September 2008. Please see minutes of this meeting for project details.

Conclusions

1. Following the planning explained on page 5, it was agreed to generate 2-page draft project descriptions for projects PI1 and PI3 and submit them for discussion. For the project PI2, this will be done as well as soon as the coordinator has been identified.
2. Although originally planned to be conducted within SIG4, the discussion about the project ideas should be led at the entire PetaMedia level (i.e. using the all@petamedia.eu mailing list). This is because the "new" scope of SIG4 described by the idea clusters listed above may attract new partners to SIG4. An expansion of the partner list of SIG4 would also benefit the search for the coordinator of the project PI2.
3. The second step involving the refinement of the project description, the partner list and the financial planning was recommended only for the project PI3.

² **Note:** This project idea was also defined during the DPB meeting, and the following list of participants was defined from the Dutch side: TUD (De Ridder) Telematics Institute, University of Twente, TNO (pending). For details about the Dutch participants, please see the Minutes of the 1st DPB Meeting, Delft, 22 September 2008

4. All project ideas will be presented during the plenary meeting in Berlin: PI1 and PI2 only conceptually, and PI3 in detail.
5. SIG4 will advise the PetaMedia SB to reserve funds for later realization of PI1 and PI2, based on the results of the “enabling” project PI3.

Explanation of the conclusion No.3³

The discussion about the objectives and the context for realization of the projects listed above revealed the need for resolving the following issues in order to be able to plan and realize some of the mini-projects.

- a. definition of the data set
- b. definition of the use scenario
- c. definition of the evaluation methodology
- d. effective benchmarking effort

While the above issues may hold for all PetaMedia SIGs, at least one of the following issues needs to be resolved for some of the SIG4 mini-projects:

- a. Building a large-enough social network of people equipped with e.g. EEG or BCI
- b. Obtaining the EEG/BCI measurements for all people already involved in a social network
- c. Having some of the people in an existing social network equipped by EEG/BCI equipment: How many of them is enough to observe their influence on the recommendation/item rating performance?

Action points:

1. Coordinators (De Vries, Pun, Raileanu) of project ideas PI1 and PI3 to draft each a 2-page project idea and submit the ideas to the PetaMedia community for discussion. The same is to be done for the idea PI2 as soon as the coordinator has been identified. **(until 10 October)**
2. Coordinators (De Vries, Pun, Raileanu) of project ideas PI3 to finalize the project proposal, the team and the financial construction based on the inputs from the PetaMedia-level discussions **(until 15 October)**
3. Coordinator of the SIG4 (Hanjalic) to initiate the search for a capable/interested coordinator of the project PI2.
4. Coordinators of all project ideas to present the projects during the plenary meeting in Berlin: projects PI1 and PI2 only conceptually, and PI3 in detail **(20-21 October)**
5. Coordinator of the SIG4 (Hanjalic) to present the issues related to conclusion No. 3 to the PetaMedia community during the plenary meeting in Berlin **(20-21 October)**

³ **NOTE:** The same conclusion came up during the meeting of the Dutch PetaMedia Branch, Delft, 22 September 2008. There, an appeal was made to the PetaMedia community to start discussing the above issues in order to enable proper definition and smooth realization of mini-projects