EUROPEAN COMMISSION
Information Society and Media Directorate-General

Converged Networks and Services / Emerging Technologies & Infrastructures

Subject: Input received on the Question of Discovery and Search in the Future Internet

1. SEARCH AND DISCOVERY IN THE FUTURE INTERNET

Search is concerned with making the best use of available (human or machine generated) knowledge to provide the user with meaningful information even if his/her request might be poorly formulated and typically unanticipated. The value of a search engine depends on how efficiently the knowledge is managed (automatically acquired, enriched, structured, retrieved, filtered, interpreted) and how easily the information is accessed and understood by the end user. Keeping the user in the loop is vital, since a search engine's efficiency is maximised when it can learn and accommodate user preferences (through online and off-line learning) from user interactions.

In this document by the term "knowledge" we mean machine-readable metadata which characterises or describes some aspect of the information. By the term "information" we imply: digital content of any type such as (rich) media, real world information, resources, service components or applications. This information will need to be adaptable for multiple types of heterogeneous devices (including fixed and mobile devices) and usage for a diverse range of users.

In this context search involves a number of different disciplines within the Future Internet, including:

- **Future Content Networks:** a) content as the main ingredient of media, and b), networks in terms of improving the user's experience and satisfaction;
- **Internet of Things:** resource and information discovery from a sea of heterogeneous devices and sensors, including Extended Home Environments, Body Area Networks, Personal Area Networks, Social networks, and so on;
- **Internet of Services:** service discovery approaches range from keyword search over service directories to semantic approaches which delineate between a service capability (what the service does), non-functional properties, and descriptions of service behaviour.

2. PROBLEM STATEMENT

Today, we are experiencing a dramatic change in the way we perceive the Internet, which is evolving towards providing rich and immersive experiences, where the user interacts

seamlessly and transparently with digital and physical artefacts. Due to the widespread availability of digital recording devices, improved modelling tools, advanced scanning mechanisms as well as display and rendering devices, even within mobile environments, users are becoming more empowered to interact in highly immersive environments incorporating new-generation digital media and audiovisual content.

Moreover, the emergence of networked sensors, "embodied interaction" with information and the availability of services (including user-generated services), facilitated by the dramatic advances of multimodal/intelligent/natural interfaces, provides the user with further degrees of freedom and channels to access information, utilising techniques which combine full-body, non-verbal, expressive and social interaction mechanisms. It is now possible for users to rapidly move from a mainly textual-based to a networked information (media, resource/service)-based "embodied" Internet, where rich audiovisual content, 3D representations, virtual and mixed reality worlds, serious games, life-logging applications and multimodal affective utterances become a reality.

The significant changes in the nature of the digital space, described above, in which digital objects reside and interact coupled with the new modes of human-machine interaction create a radically different and challenging problem for search. The days of only matching keyword text queries against textual documents, will soon disappear.

Within the Future Internet content/context-based search is a topic which covers several different communities each with their own language, notations, tools and methodologies. Specifically, Future Internet search can be considered from the viewpoint of media, physical objects or services. One of our main concerns of the FIA Stockholm event will be to bring these communities together and to have an effective dialogue.

To initiate this common understanding, with respect to the Search Engines' problems that currently exist and span between the different "Internets" we have identified the following common themes:

I. Heterogeneity and diversity issues, concerning resources, information sources, infrastructure, user communities, etc;

II. Human factors, concerning "the user in the loop", feedback, personalisation, recommendations, interaction with the information, emotional characteristics, etc;

III. Real time issues, concerning content/media indexing and retrieval, visibility, aggregation, re-composition of services, etc;

IV. Security and privacy issues;

V. Evaluation and benchmarking issues.

It should be clearly stated that in the FIA Stockholm event we will try to provide answers to a set of questions that have already been identified (Section 2.1) and belong to the first three problems. The last two, which link the topic of discovery and search with security and privacy (relation with Trust and Identity FIA group) and with evaluation and benchmarking (relation with FIRE), will be tackled in the FIA Valencia event.

2.1. Key Challenges

A non-exhaustive list of potential challenges categorised into the above mentioned 5 main categories, can be found in the table below. These have been reported by experts belonging in different FIA groups (Section 4).

Heterogeneity & Diversity	Human Factors	Real Time Issues	Security & Privacy	Evaluation & Benchmarking
Q1.1: How do we best support the formulation of search queries over possibly rich Future Internet artefacts, over a set of different types of devices and modalities?	Q2.1: how do we support discovery using social/emotional cues, individual vs group (is the group an average over the individuals?). For group search issues related to identity, social group, empathy, arise	Q3.1: How do we manage service & discovery in the FI in real-time when everything (producers, consumers, networks) is recomposing and moving in real-time?	Q4.1: How do we manage privacy, trust and security within search?	Q5.1: How can we evaluate our SEs in FIRE facilities?
Q1.2: Given the Polymorphic facets of the Internet (e.g. communication-centric, information-centric, context-centric, resource-centric, content – centric, service/computation-centric, device-centric, object-centric and management-centric Future Internet), how a SE should look like? Are the open, federated search services a/the solution	Q2.2: How do we best take into account the diversity of the user community?	Q3.2: How do we manage the visibility of services in real-time?		
Q1.3: Do we need specific infrastructures and representations to support search by	Q2.3: How do we address ranking, user feedback about the relevance of	Q3.3: How do balance visibility and real-time		

agents?	search results, and quality of the recovered information?	requirements?		
Q1.4: Can we chain (interoperate) search and discovery engines?	Q2.4: How do we best incorporate human computation into search and discovery engines?	Q3.4: How do we manage changing relationships?		
Q1.5: How can we implement/architect search and discovery over a range of information artifacts including services, sensor networks and rich multimedia content? What types of indexing/caching could support FI search and discovery?	Q2.5: How can we provide better recommendation systems based on SEs?	Q3.5: Should we discern between different types of urgency?		
Q1.6: What types of representations (meta-data) could we use to support Future Internet search across the heterogeneous resources? How would these be created and maintained?				

3. CARETAKERS

Petros Daras (FCN) and John Domingue (FISO)

4. PARTICIPANTS

Michel Riguidel (FISO), Pierre-Yves Danet (FCN), Tasos Gavras (FIRE), Theodore Zahariadis (FCN), Srdjan Krco (RWI), Stamatis Karnouskos (RWI), Harald Sundmaeker (RWI), Dieter Fensel (FISO), Alex Gluhak (RWI), Nozha Boujemaa (FCN)

5. POINTS OF AGREEMENT

<List here the points of agreement and a brief explanation of why/how a consensus has been reached. Include as well significant options that have been left behind.>

- Definition of search and information retrieval for the different facets of the FI.
- Heterogeneity of infrastructures/platforms, resources, protocols, representations.

6. POINTS OF DISCUSSION

<List here the sub-points which are still under discussion as well as a brief explanation of the open options for each>

- Types of representations or meta-data models that could be used to support FI search. It is expected to be covered in FIA Stockholm.
- Architectural considerations (centralised vs distributed search, fixed & mobile devices, etc.). It is expected to be covered in FIA Stockholm.
- Standards. It is expected to be covered in FIA Valencia.
- Evaluation framework. It is expected to be covered in FIA Valencia.

7. FOLLOW UP ACTIONS

<List of agreed actions to do after Stockholm and before Valencia. These actions are to be undertaken by the 7 groups>

- Clarification of questions/answers

8. CONSOLIDATION OF THE DESCRIPTION OF THE MILESTONES AND ROADMAP OF RESEARCH RESULTS REFERENCES

<References to external documents should be included here with a view to keep the overall text not longer than 2-4 pages. Include as well the presentations made during the conference>

9. DOES THIS TOPIC REQUIRE A FOLLOW-UP DISCUSSION IN VALENCIA?

Yes: **Empowering discovery and search in the FI**