JRC
EUROPEAN COMMISSION

ipts
Institute for Prospective Technological Studies

Joint Research Centre (JRC)

"Socio-Economic and Legal Aspects of Audiovisual Search Engines"

Ramón Compañó

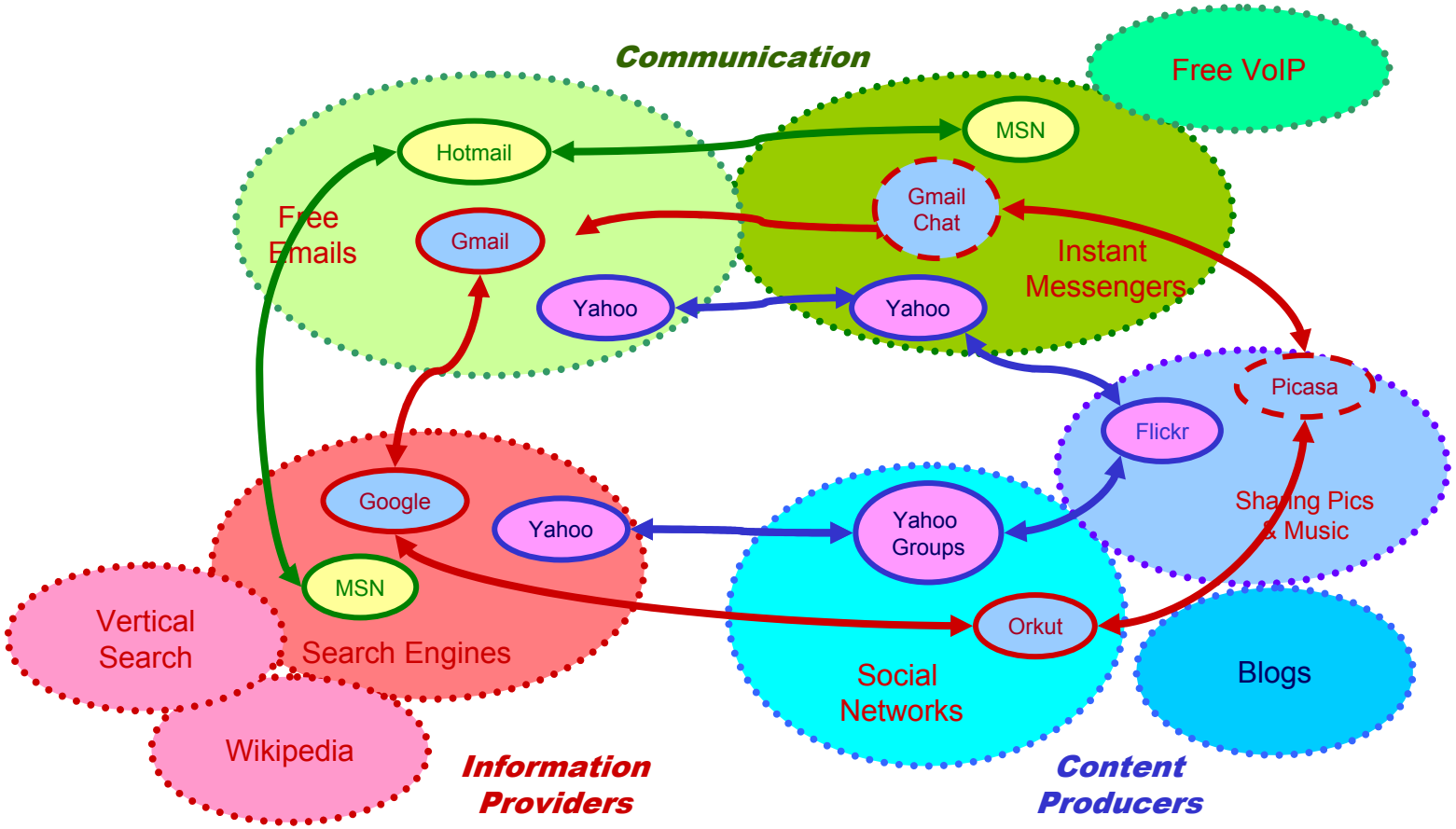
Brussels, 27th April 2009

IPTS - Institute for Prospective Technological Studies

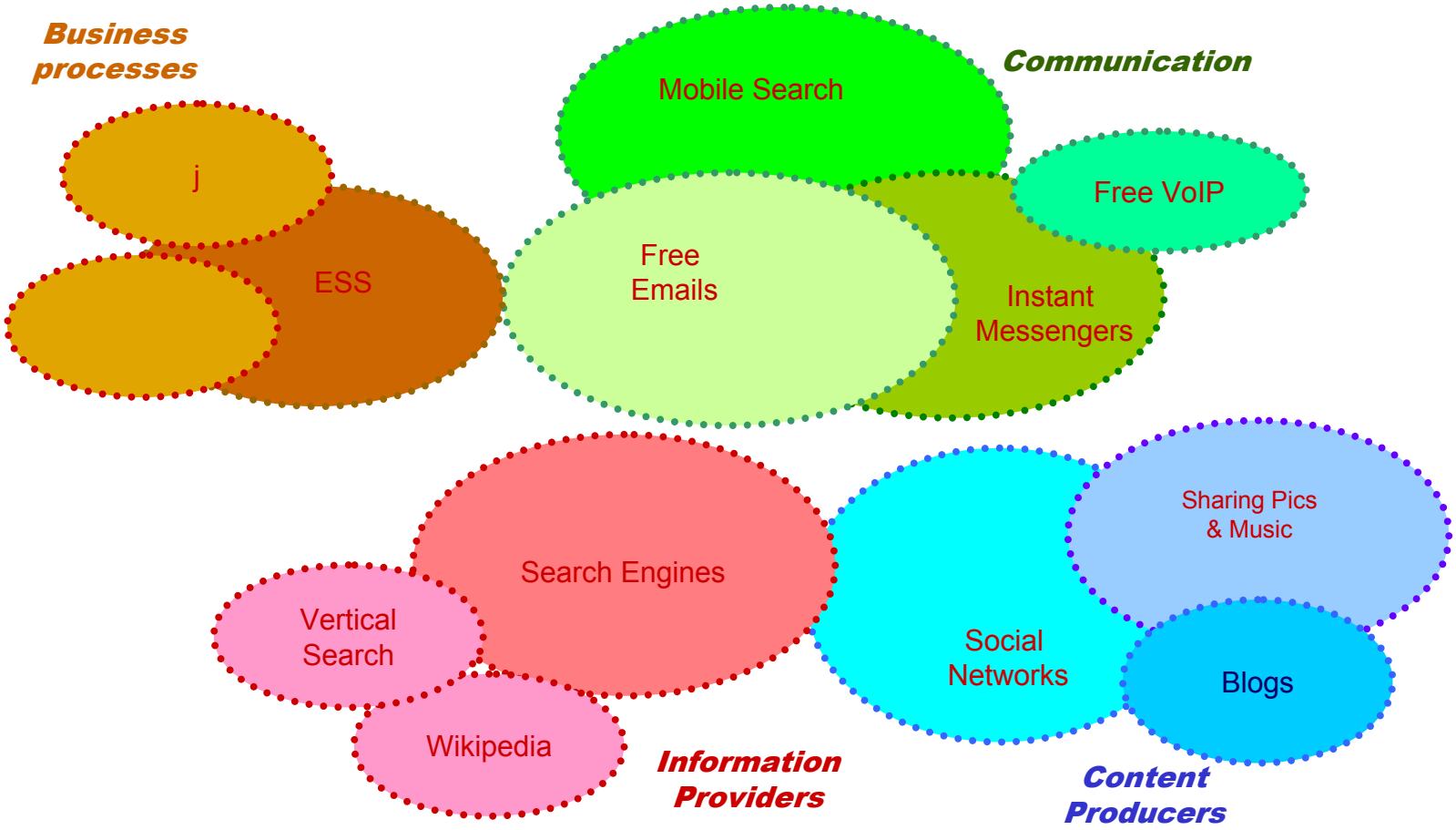
Seville – Spain

http://ipts.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

The encompassing nature of search



The encompassing nature of search



Part 1

Issues related to Content

- Media (manipulation, bias, self-referencing, gateway argument …)
- Individual vs collective creativity
- Ownership of meta-data
- Copyright

Data and People

- Personalization, Customisation,
- Privacy Concerns
- Solutions under discussion

People´s Behaviour and the Decision Maker

- Privacy paradox
- Control paradox
- Responsibility dilemma

Part 2

Mobility and their Disruptive Technologies

SWOT for Europe

Policy Options

Content: Copyright

The **EU Copyright Directive** states in its preamble that *browsing* and *caching* ought to be considered legal exceptions to the reproduction right.

Search engine's spidering process requires at least one initial reproduction of the content in order to be able to index the information;

There are some similitudes between copying for browsing and for indexing;

The cache copy provision was originally foreseen for Internet Service Providers to speed up the process, but cache copies of Search Engines resemble more an archive

Search Engines (SE) are already very important AV-SE will be even more so given

- the explosion of AV content;
- the need for the organisation and categorisation of all sorts of information, particularly in audio-visual (AV) form

SE are in the process of becoming fully-fledged information portals, rivalling traditional media players

- Due to the shift from pure information retrievers to categorisers,
- Due to the shift from an information ´pull´ to a ´push´ to the user?

Application of Law

- Courts in different countries appear to have drawn different conclusions on the same copyright issues.
- is a slight refocusing of copyright law may be necessary to offer more certainty?

How to incentivise the creation of quality content?

Given the proliferation of digital content, it becomes more difficult to locate specific content.

How to deal with metadata?

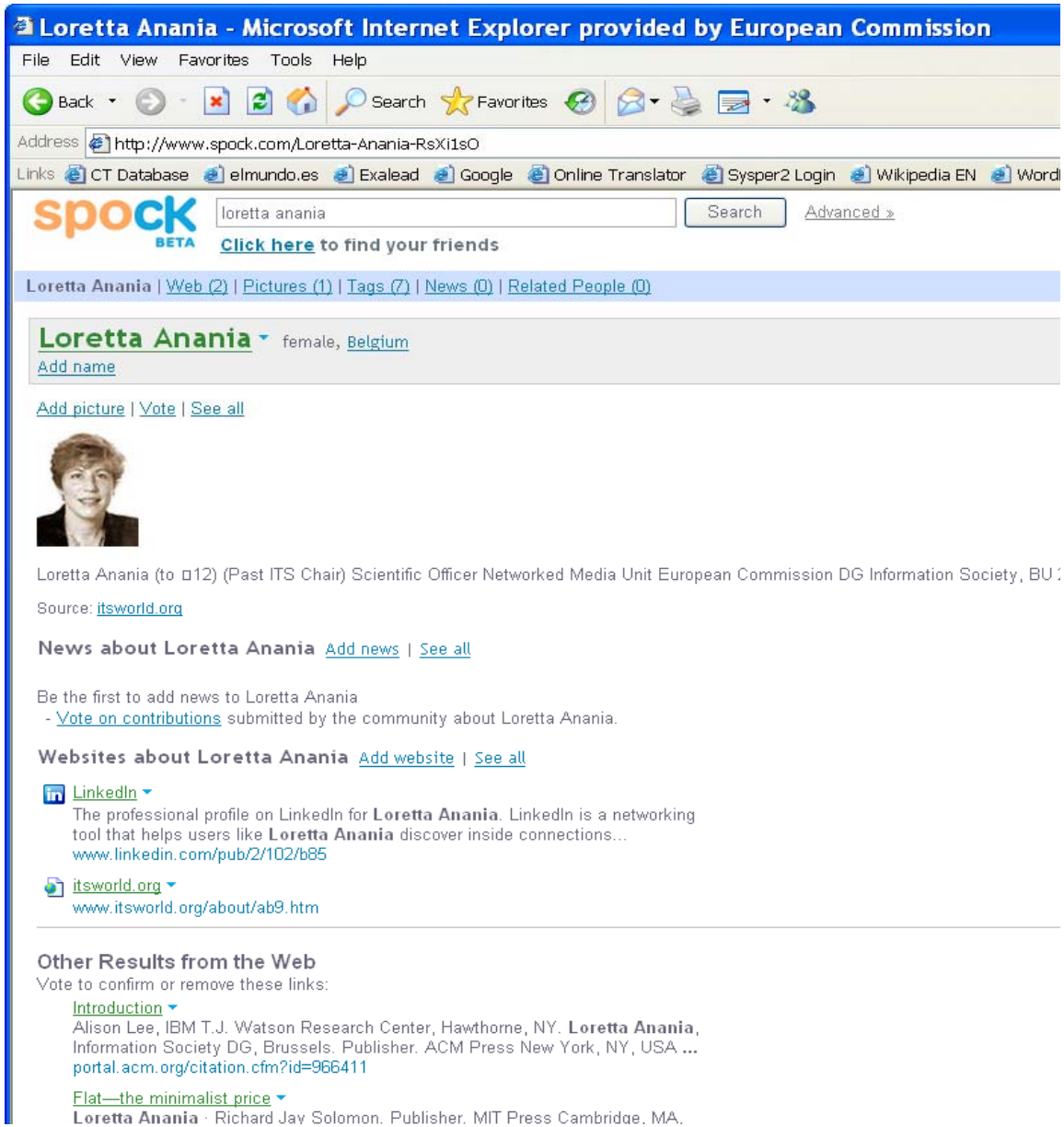
It becomes comparatively more important to promote the development of methods for accurate labelling of AV content than to incentivise creation.

The way forward?

Regulation (e.g. copyright law) is only one modalities, technological standardisation (e.g. robot exclusion protocols) and commercial deals between market players are others. All should be used.

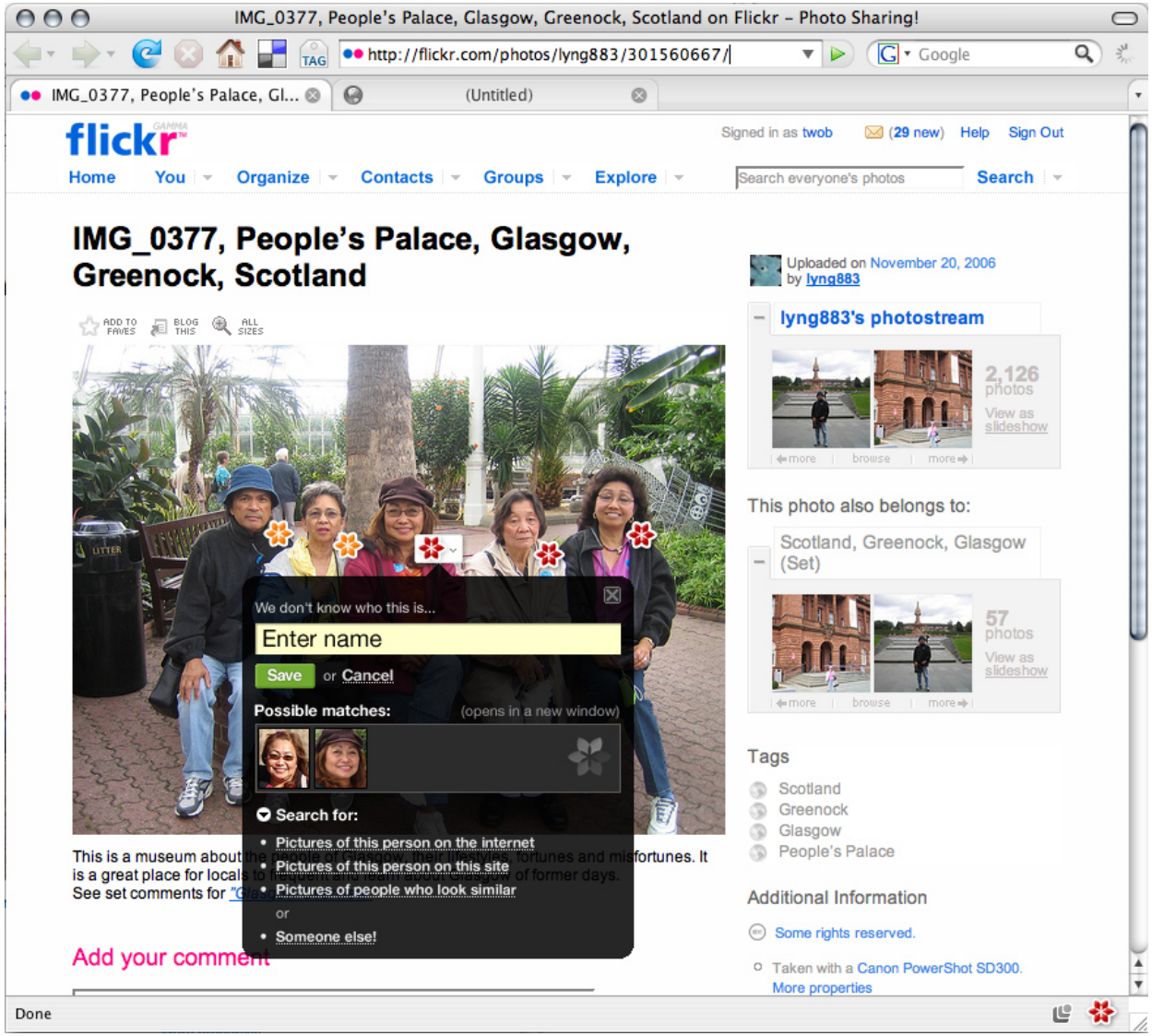
People Search Engines

- Gathering and filter existing information about people;
- Tag and create meta-data on people;
- Structure and deliver personal data;
- Invite users to confirm automatically generated meta-data and create new meta-data;



Info Syndication
+ Similarity Search
+ local geo-information

Sources
increasingly more from social networks sites

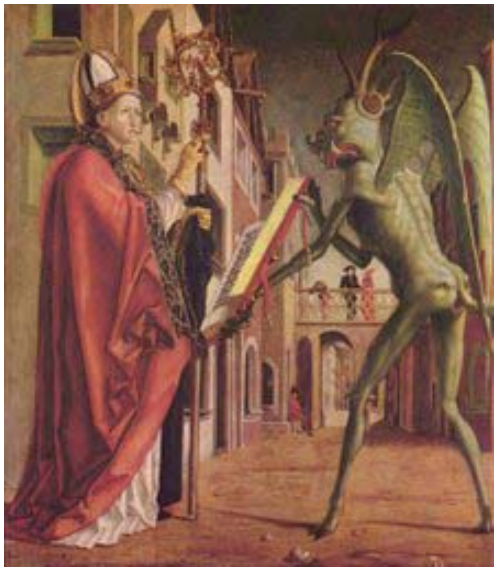


Privacy: the concerns

- **Risk of Loss of Privacy**
  - Privacy ≠ Security
  - Privacy is context dependent
  - Privacy has a cultural element and changes with time
- **Risk of User Surveillance**
  - Using search engines to find other users' personal information
    - Big Brother watching you (e.g. government)
    - Little Brother watching you (e.g. neighbours, companies, etc)
  - Search engines logging search data of their users
  - Data logged forever
  - Searchable forever
- **Risk of Misuse**
  - False associations,
  - Defamation;
  - Racism…

The Context is changing

- Bluring ´offline´ and ´online´ privacy;
- New understanding of what is private and what is public (e.g. ´that I was in public does not mean that image should be public);
- From individual rights to collective rights;
- …



Personalsation and Customisation

- ***User-side***: more relevant better retrieval results; adequate offers,
- ***Technical side***: user-driven ranking; speed vs processing trade-off
- ***Economic side***: targeted advertising – selling "personalised eyeballs"; preventing click-fraud

Tensions

- Balancing between promoting innovation and assuring privacy;
- Personalised services require more personal data (Faustian Bargain?)

1. Social Norms
2. Market-led / Industry-Self-Regulation
3. Governmental Regulation
4. Technology-based solutions

Underlying idea: The power of masses!

- **Reputation-based systems?**
  - e.g. similar to eBay
- **Self-regulating control systems?**
  - e.g similar to Wikipedia
- **Negotiated market for personal data?**
  - e.g Laundon Model

Policy Options: Market-led Options

- **Self-Regulation by Industry**
  - Self-regulation is a social norm.
  - Exist (partially) in Germany
- **Co-regulation by Industry and Government**
  - Both code regulation and social norm
- **Perceived Potential Risks triggered by SE**
  - Access to harmful content;
  - Access to illegal content;
  - Discrimination of content;
  - Misleading consumers;
  - Exploiting protected works and data
  - Influence on opinion making;
  - Distortion of competition, including the transfer of market power to other markets, e.g. advertising

Policy Options: Regulation

- **Re-conceptalization of privacy rights**
- **EU Data Protection Directive**
  - Control over data collection, data circulation, data usage, adequacy of data
  - Rights of accessing data, right to delete data
- **Recommendations on the Art. 29 Data Protection Working Party on Search Engines**

Opinion on data protection issues related to search engines (4 April 2008):

Applicability of EC Directives:

1. The **Data Protection Directive** (95/46/EC) generally **applies** to the processing of personal data by search engines, **even when their headquarters are outside of the EEA**.
2. Non-EU based search engine providers should inform their users about the conditions in which they must comply with the Data Protection Directive, whether by establishment or by the use of equipment.
3. The **Data Retention Directive (2006/24/EC) does not apply to internet search engines**.

"Obligations on search engine providers

6. Retention periods should be minimised and be proportionate to each purpose put forward by search engine providers. (…) In case search engine providers retain personal data longer than **6 months**, they must demonstrate comprehensively that it is strictly necessary for the service. (…)

9. Search engine providers must give users clear and intelligible information about their identity and **location and about the data** they intend to collect, store or transmit, as well as the purpose for which they are collected.

10. Enrichment of user profiles with data not provided by the users themselves is to be based on the consent of the users.

14. Search engine providers that specialise in the creation of value added operations, such as profiles of natural persons (so called '**people search** engines') and **facial recognition software on images** must have a legitimate ground for processing, such as **consent**, and meet all other requirements of the Data Protection Directive, such as the obligation to guarantee the quality of data and fairness of processing.

Rights of users

15. Users of search engine services have the right to access, inspect and correct if necessary, according to Article 12 of the Data Protection Directive (95/46/EC), all their personal data, including their profiles and search history.

16. **Cross-correlation of data** originating from different services belonging to the search engine provider may only be performed if **consent has been granted by the user for that specific service**.

"

Policy Options: Summary

- **Issue**
  - If properly handled, Search Engine Providers, Advertisers and Users can greatly benefit from personalization;
  - To provide more added-valued services for the user, more data is required about the user. Is this a Faustian Bargain?

- **Options**
  - Viable bottom-up socio-based approaches not yet in sight;
  - Industry-led self-regulation may be viable only in those domains where there is a common interest
  - Legislation in place, but insuffient or non-viable to solve upcoming problems (e.g. difficult to enforce)
  - Privacy Enhancing Technologies too little discussed in this context

Dependent.

35% of the searchers use a search engine daily and
47% of searchers use it once a week.
32% consider themselves "addicts"
44% of searchers say they regularly use one single search engine, 48% will use just two of three search engines

Happy

87% of the internet users say they have successful search experiences most of the time

Naïve

68% of users say that search engines are a fair and unbiased source of information
33% of users are aware of the difference between sponsored and 'organic' results. Only 16% of the searchers state that they can consistently distinguish between paid and unpaid results.

[Source: Internet Pew Project]

Internet confidence

Overall: low

Very low for individual-related aspects: safeguards, self-protection

Privacy risks

Very concerned (70 – 80%) about negative consequences of spreading personal data – stealth use, financial fraud, dossiering

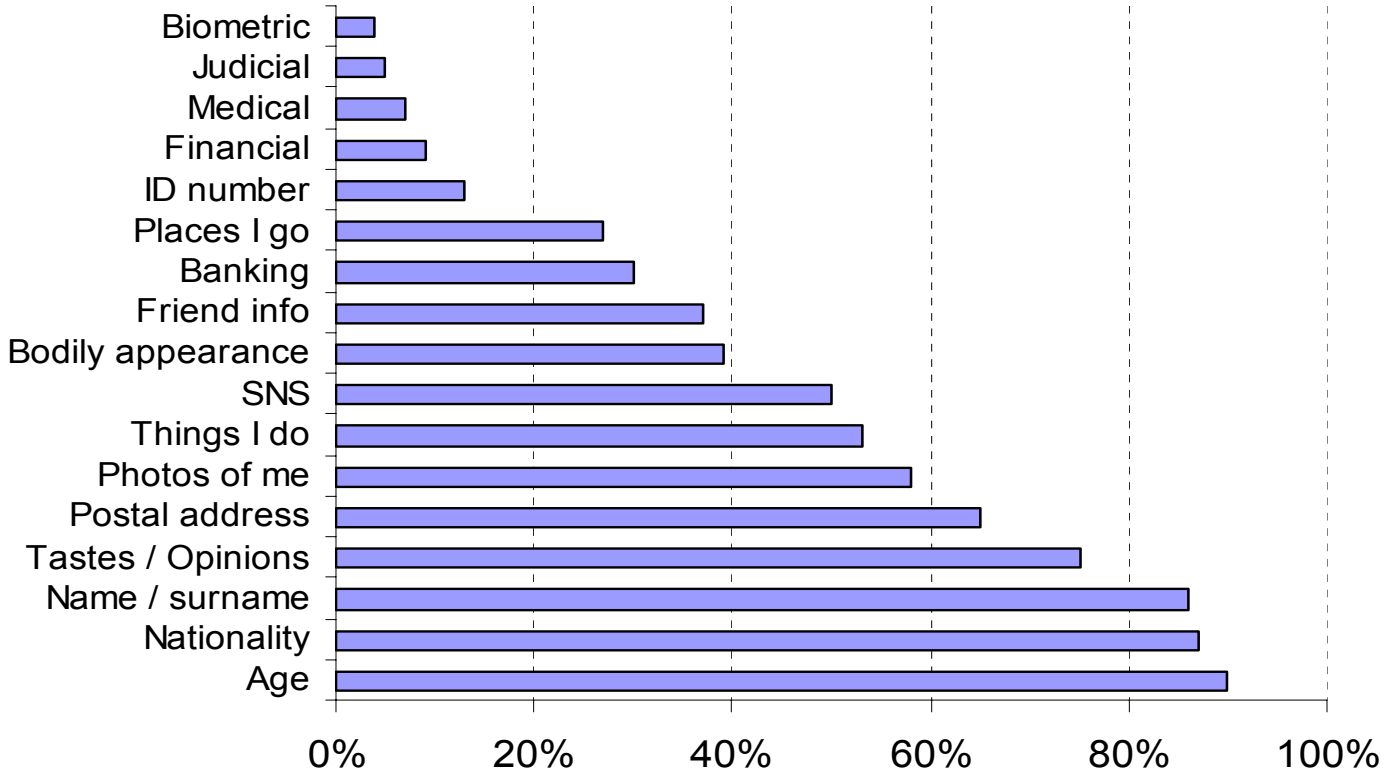
Concerned (60 – 70%) about reputation and identity damage

"Privacy Paradox"

Users state that they are concerned about privacy, but do not act accordingly

[Source: "Young People and Emerging digital Services" W. Lusoli, C. Miltgen, IPTS 2009

Young people happily give out biographic information (~80%), social data (50-60%), much less so sensitive information (5%).



[Source: "Young People and Emerging digital Services" W. Lusoli, C. Miltgen, IPTS 2009

Young people give personal data online to…

- receive better services
- connect with others
- personalised commercial offers are least appreciated

To shield their identity, young people…

- use active strategies – dummy email accounts, tinkering with own personal details
- do not use tools provided externally – trust badges and privacy-enhancing technologies

Young people believe that it is their own responsibility to protect their data online and of the companies they transact with

	% Agree
It is my responsibility to protect my personal data	32
It is the responsibility of the company I transact with to protect my personal data online	27
It is everybody's responsibility to make sure personal data are safe online	26
It is the government responsibility to protect my personal data	8
It is the responsibility of the police and courts to ensure that personal data are protected online	7

[Source: "Young People and Emerging digital Services" W. Lusoli, C. Miltgen, IPTS 2009

Internet confidence

	% Agree
The internet provides a trusted environment in which to make transactions for leisure, work and business	38
In general, the internet is now a robust and safe environment in which to transact	30
The internet is safe enough to preserve my privacy as I carry out leisure, business and personal activities	29
I am confident that I can protect my privacy online	27
The internet has enough safeguards to make me feel comfortable giving my personal details online	27

[Source: "Young People and Emerging digital Services" W. Lusoli, C. Miltgen, IPTS 2009

Remedies and responsibilities

They think that there are solutions (~80%). Technical solutions to preserve personal data are favoured over awareness raising

They ask for guarantees of procedural fairness (~65%) and for increased user control (~50%)

Control paradox. Respondents are confident they can protect their own privacy online, but few claim that it is efficient to "give users more direct control on their own identity data"

Trust

Young people trust what they know

	Mean
A company I am familiar with	3,27
The local council	3,02
A well-known company	2,91
The European Union	2,89
The national government	2,81
A non-profit association	2,61
An unknown company	1,82
Note The scale spans from 'very much trust' (5) to 'not trust at all' (1); (3) is the average point.	

[Source: "Young People and Emerging digital Services" W. Lusoli, C. Miltgen, IPTS 2009

Take away messages

There is no single option to meet the privacy requirements

A problem caused by technology might probably be also solved by technology: Consider more privacy-enhancing-technologies

Regard the Privacy discussion as an opportunity

Policy makers have to struggle with behavioural paradoxes and balance between dilemmas

Privacy paradox, responsibility paradox, trust dilemma, …

Web Search
Mobile Search

SWOT for Europe
Policy Options

Structure

- Many search engines (>70 world-wide)
- Few players with significant share
- Revenues (mainly) through advertising

Trends

- Integration: horizontally & vertically
  - With other applications
  - With content provision
- Increased centrality of search
- Need for user data: personalisation & stickiness
- Proactive engines: from ´pull´ to ´push

- Market concentration possibly to continue in the audio-visual search era !

Operational Cost of providing web search

- Content crawling and indexing (cost proportional to content volume)
- Query answering (cost and revenue proportional to traffic and user population). Ideally, once traffic increases, the investments are compensated by revenue.

Audio-Visual search far more demanding than text-based. Operation costs will further increase.

Challenges

- Reduce hardware related costs
- Make data centres more efficient
- Reduce power/electricity comsumption
- ….

Structure

- actors include telecom operators, device manufacturers, search technology providers, content providers and advertisers.
- market in its infancy and fragmented.
- Current walled-garden business model may suffer structural changes.

Trends

- Huge Growth rates (Mobile broadband connection: 3m (March 2007), 32m (March 2008) [Source GSM Association]
- Main attraction, include Weather, local information, maps & directions,

	2006	2007	2008	2009	2010	2011
Mobile Internet Users	*337,3*	*405,5*	*489,6*	*596,4*	*757,1*	*982,4*
Mobile Search Users	*266,0*	*327,2*	*410,7*	*561,8*	*672,3*	*901,1*
Mobile Search Ad-Revenues [m$]	*6,8*	*63,1*	*221,3*	*580,3*	*1148,9*	*2361,5*

[Source: eMarketer July 2007]

Challenges

- Mobility imposes very specific search, interaction, retrieve, display requirements,
- Mobility requires coupling of content search with other technologies, e.g. location
- Heterogeneous mobile-fixed environments bring thin different types of constraints
- Personalised search and content adaptation

Mobile Search: SWOT

Strengths

- Technological puzzle pieces in place
- Good research standards
- Content of higher quality for mobile use (geo, cadastre, …)
- Strong Industrial landscape
- Public funded broadcasting

Weaknesses

- Fragmentation (roaming, regulation, cultural)
- Need for better / understandable / more secure pricing models
- Roaming charges
- Strategic decisions on innovation and investments are outside EU
- Venture capital / Entrepreneurship
- Lack of interoperability and (open) standards

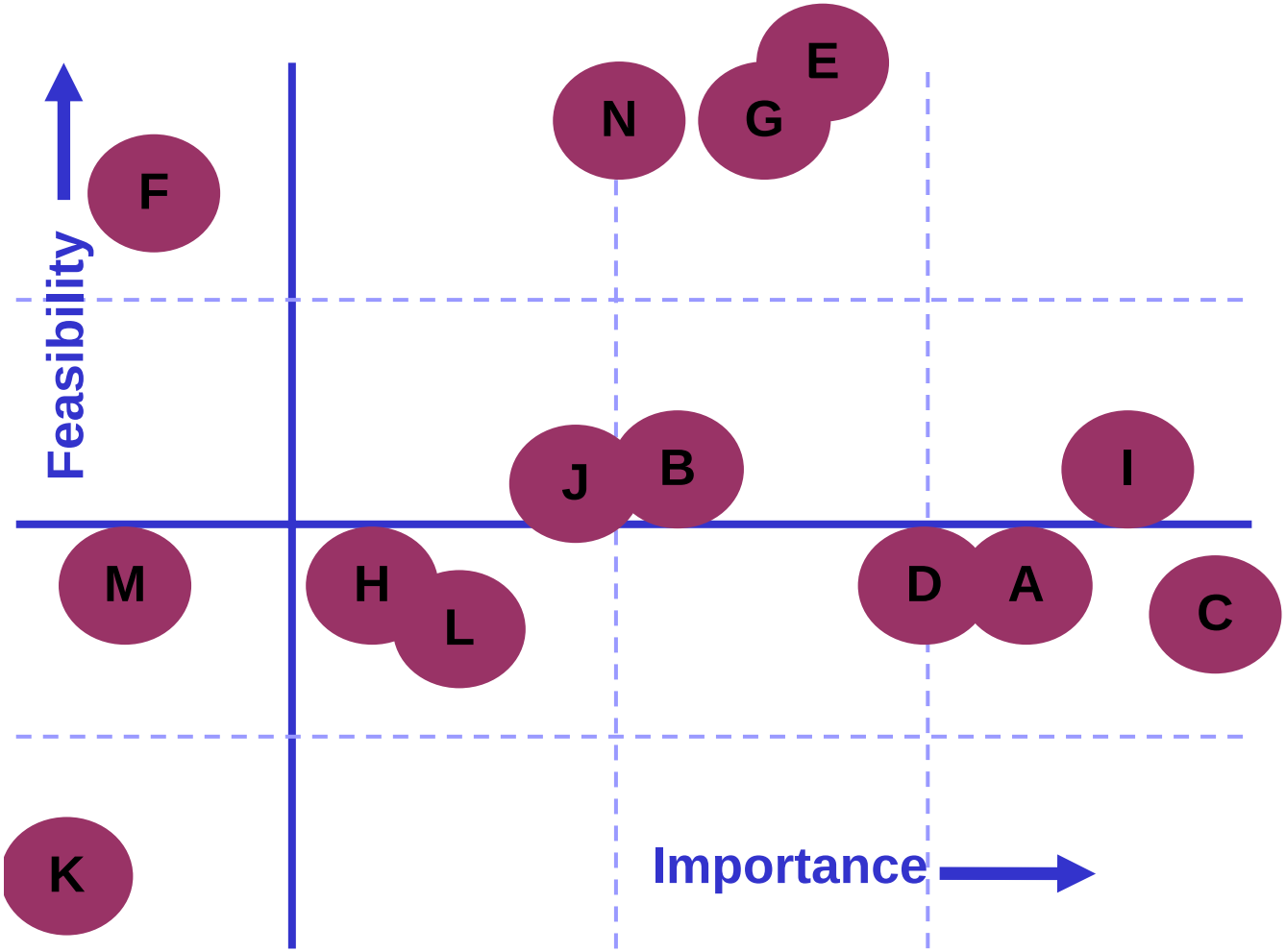
Opportunities

- Improving integration between web/mobile/pc platform for a richer user experience
- Niche markets/services
- Local content (multicultural)
- New regulatory framework needed for API's, privacy, …
- Data portability
- Liberation of public data
- Disruptions (cloud computing)

Threats

- Lack of technology development
- Fragmented market (silos, platforms)
- Privacy issues, data protection
- Companies outside EU will control the developments in mobile search
- Asymmetry of regulation
- Regulatory lag (spectrum management)

Policy Options



A. Raise user-awareness
B. Create tools for user privacy/manage ID
C. Support innovators/entrepreneurs
D. Reform regulatory framework
E. Promote local user-centric env. (liv labs)
F. Develop public services (emergency, …)
G. Research programmes
H. Promote self-regulation
I. Promote standards/interoperability
J. Develop accessibility of content
K. Support EU champion
L. Public procurement
M. Promote EU internal market\
N. Establishing a multi-stakeholder discussion platform

Thank you for your attention

ramon.compano@ec.europa.eu
Institute for Prospective Technological Studies
JRC IPTS, Seville