KW Partner: France Télécom R&D

1 Overview

Challenge
To provide photo album and personal data services via portals.
Solution
Semi-automated aggregation and publication of personal content, digital camera and mobile phones offering today content annotations capabilities (localisation, date..)
Why a Semantic solution
Meta Data extraction and management, content provider mediation are promising areas for semantic portal
Key Business Benefits
B2C enhanced paying services, partnership with the actors of the photo publishing
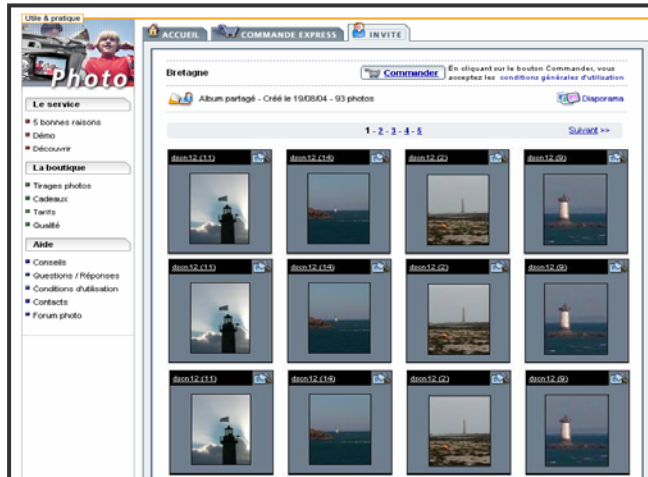**Business Partners**
Orange, Wanandoo, Photo publishing actors

Figure 1 – E-photo album service

<u>**Keys components**</u>

<u>Existing Software</u>
GIS, transformation voice/text, Pattern Recognition in digital image
Domain ontology

<u>Research and development</u>
Semantic meta-data fusion
Semantic data integration
Natural Language Processing

<u>Technology locks</u>
Semantic directory (e.g. UDDI)
Dynamic discovery of Content Providers
Dynamic integration of Content Providers in a directory
Automated annotation

The management of personal data is changing due to new service portals and the appearance of a new generation of customer centred on multimedia (e.g. ILife solution from Apple). These services are limited today to a classic management of data: create/modify/delete photos, photo albums/musical library. They only exploit the superficial potential, which is involved in aggregation and in content inference (Figure 1).

The aim of this use of space is therefore to shed light on this potential, to identify associated problems and to class them into categories: the problems, for which the package solution exists or successful research work and those, which necessitate an investigation by R&D (laboratories, specialist companies, etc…).

Example of the application of new services dedicated to the management of photo albums via a portal:

- Mass manipulation (a digital camera can contain up to 5000 digitalized photographs) and semi-atomisation of the transferral of contents and their organisation,
- Semantic aggregation of contents (photos with other photos, photos with relevant content, etc.)

- Dynamic access to suppliers of external contents via the portal,
- Process of organisation and enhancement of contents (e.g. transformation voice/text, Recognition of form, Geographical location, etc.)

These services could equally be applied to other content types, for example the availability directory of a music library or all other content, since outside of specific treatment of a target category of the multimedia content it is intrinsically applied to more general problems:
- Due to the explosion of production of content, their exploitation needs very efficient, personalized, intuitive and automated tools and services,
- Rapid access and structure of the published contents on the web via inference solutions.

2 Current Practices and Technologies

2.1 Typical business practices

A tourist with a digital camera, under the spell of Brittany, takes photos of landscapes, of his/her friends, of his family, of all what he wants…
Back home, he loads his photos on his personal computer and stores them inside directories. Afterwards, directories constitute very often the only mean of accessing photos. Generally if the names attributed to directories seem to be temporarily suggestive, this feeling vanishes quickly with time when you are seeking in your past photos.
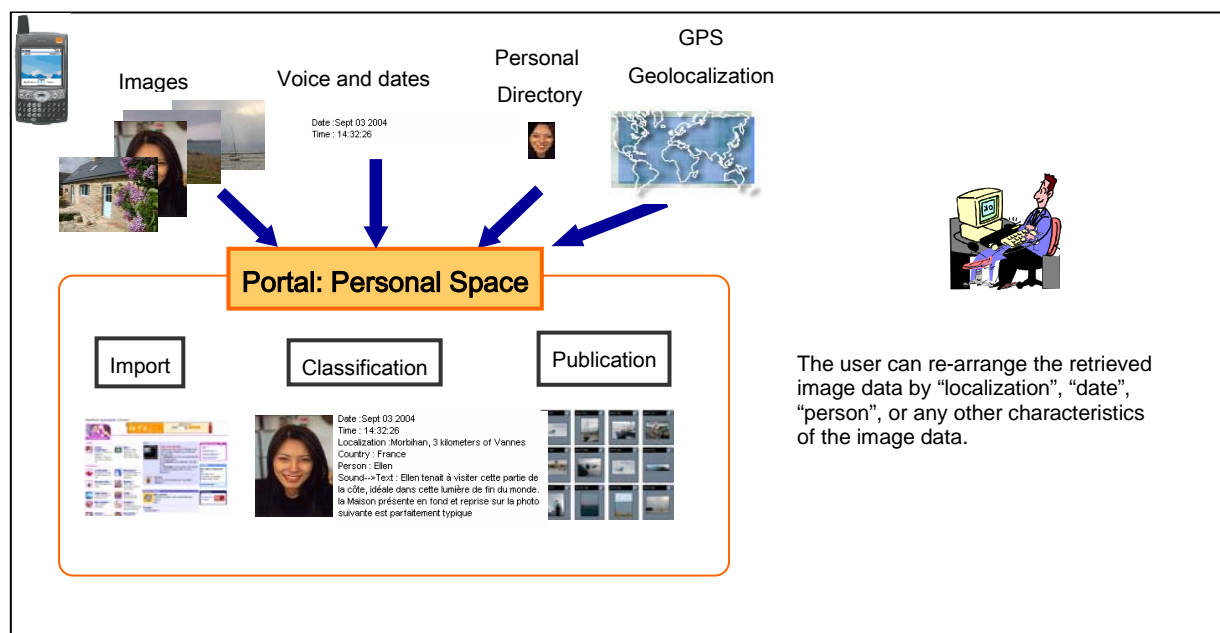

Figure 2 – Proposed semantic-enhanced photo album

The tourist loads his photos on the portal and more accurately on his personal storage space (Figure 2). A Personal storage space is dedicated to one customer and can contain any personal data that this customer wants to entrust to the portal and especially we can imagine that our tourist entrusted his personal contacts (like in mailboxes) to the portal. A contact is made up of a first name, last name, and a photo. This is the relevant information relative to a contact that the portal will need later as part of this scenario.

Anyway, the portal recovers photos and associated meta-data (Date and time, localization, vocal message describing photos). The portal first applies treatments to the provided meta-data:

- It turns a physical localization (x, y coordinates) into a geographical one (a town…) by using a distant geographical information system.
- It turns vocal messages into text by applying natural language processing. Afterwards, the portal indexes the text.
- Next, the portal makes those meta-data complete:

Based on indexed words, the portal is able to retrieve external relevant information coming from content providers in order to enrich photos. For instance, we can imagine that our tourist takes a photo of a lighthouse (Figure 3) and enters the following information: "I took a photo of the lighthouse of the Cordouan". Afterwards, the portal indexes words "lighthouse" and "Cordouan". Next, the portal identifies the appropriate content providers and through a mediation services, retrieves the relevant information relative to this lighthouse.

Figure 3 – Integration of relevant external data from keywords

If our tourist says in a vocal message that he took a photo of his wife, the portal is able to learn and to store this information by associating indexed words and personal contacts. Otherwise, the portal applies algorithms of pattern recognition based on photos of contacts in order to detect the presence of his wife.

2.2 System requirements Analysis

To make the search of photos easier for customers, we need metadata. It is relatively obvious to define useful meta-data for photos (localization, date and time, persons present on the photo…) It is less easy to define when and how to associate those meta-data with photos. Concerning the moment, we thought that is preferable to associate meta-data with photos at the same time that the tourist takes photos. Concerning how to associate them, a digital camera is already able to associate automatically meta-data with photos. Indeed, a digital camera determines the date and time of the snapshot. Some digital cameras linked with a GPS can supply the exact localization, but we can also imagine, that tomorrow some cameras will allow our tourist to associate manually meta-data like vocal messages describing the photos he took.

These Use Cases make the following problematic stand out:

- With some existing solutions:

1. Geo-localization
   The needs in terms of geo-localization require the integration of a GIS (Geographical Information System). These days, this type of problem gives rise to solutions that are already industrialized.
2. Transformation voice/data and pattern recognition
   The problems of transformation from voice to data and text indexation obtained in this way such as based on recognition of patterns are resolved and we can use existing solutions for a rapid service deployment.

- With some potential semantic type solutions available in few months time:

3. Semi-automated integration of photos to the portal whatever the structure or syntax of the sources,
4. Aggregation of external and heterogeneous contents,
5. The availability of a user-friendly interface with efficient and personalized research capabilities

- With the following open research topics:

6. The dynamic discovery of new content providers,
7. Dynamic integration of content providers inside a semantic directory (a la UDDI),
8. The fusion of data requires prioritisation amongst the content providers for a choice of equivalent semantic information. Example: The same lighthouse gives information to 3 different content providers. A possible solution that is conceivable for this prioritisation is supported by trusted indications attributed to each content provider.

2.3 Review of the current systems

There exist current solutions like PBase http://www.pbase.com or e.photo http://photos.wanadoo.fr/ but they are limited. They provide only simple loading publication capabilities.