	<p style="text-align: center;"><b>WebPark</b> <b>IST-2000-31041</b></p> <p style="text-align: center;">Geographically relevant information for mobile users in protected areas</p>
---	--

# Privacy / security assessment report, D2.2.2

Type PU\*

\*Type: CO-consortium, RE-restricted, PU-public

Report Version: 04  
 Report Preparation Date: 17 April 2002  
 Contract Start Date: 17 October 2001  
 Duration: 3 years  
 Project Co-ordinator: Geodan  
 Prepared by: Luc ABADIE, Jonathan RAPER  
 Partners: CU, LNEC, SNP, EADS, GUIZ



**Project funded by the European Community under the "Information Society Technology" Programme (1998-2002)**

Project	Nature	WP	Deliverable	Written by	Delivery date	Contractual delivery
<b>WebPark</b>	<b>R</b>	<b>2</b>	<b>2.2.2</b>	<b>LA, JR</b>	<b>17/04/02</b>	<b>01/05/2002</b>

\*\*Nature: R-Report, P-Prototype, D-Demonstrator



**Abstract:**

Visitors to, and users of, recreation/ protected areas need geographic information to plan their trips safely. Moreover, they want to have contextual information on topics of interest while involved in outdoor activities. To deliver the best information service at home and/or on mobile devices, the WebPark project has to take into account privacy / security issues in association with personalisation. This report focus on the management of personalisation characteristics, and the associated privacy / security issues.

**Partner owning: EADS S&DE, CU**

**Partners contributed: EADS S&DE, CU, Geodan**

## Document information

### Codes

Full document name: WP\_D222\_Privacy\_Security\_assessment\_and\_design\_requirements\_v04.doc

### Revision history

Version number, date	Author(s)	Summary of changes to previous version
01, February 2002	David Gaussens	First draft
02, April 2002	Debbrah Phan	Comments on first draft
03, April 2002	Jonathan Raper	Final draft
03, April 2002	Debbrah Phan	Final draft – document format and logo
04, November 2004	Evert van Kootwijk	Style improvements

### Distribution list

Name	Organisation
Jonathan Raper	City University
David Mountain	City University
Euro Beinat	Geodan
Peter ter Haar	Geodan
Henk Scholten	Geodan
Eduardo Dias	Geodan
Mijntje Spaapen	Geodan
Armanda Rodrigues	LNEC
Antonio Goncalves	LNEC
Christophe Rhin	EADS
Luc Abadie	EADS
Dominique Medal	EADS
Walter Abderhalden	SNP
Reudi Haller	SNP
Katrin Krug	SNP
Rober Weibel	GIUZ
Dirk Bughardt	GIUZ

### Approval

Name	Date and signature	Remarks
Euro Beinat	17 April 2002	Distributed to all partners



<b>1</b>	<b>Introduction.....</b>	<b>5</b>
<b>2</b>	<b>Users requirements assessment.....</b>	<b>5</b>
2.1	<b>Introduction .....</b>	<b>5</b>
2.2	<b>Identification .....</b>	<b>6</b>
2.3	<b>Use cases.....</b>	<b>6</b>
2.4	<b>Topics of interest .....</b>	<b>6</b>
2.5	<b>Location.....</b>	<b>7</b>
2.6	<b>Multi-language capability .....</b>	<b>7</b>
2.7	<b>Conclusion .....</b>	<b>7</b>
<b>3</b>	<b>Privacy/ security issues .....</b>	<b>8</b>
3.1	<b>Introduction .....</b>	<b>8</b>
3.2	<b>Legislation and regulation.....</b>	<b>8</b>
3.2.1	Legal protection of privacy .....	8
3.2.2	National laws .....	8
3.2.3	European Union Data Protection Directive .....	9
3.3	<b>Technical aspects of privacy on networks.....</b>	<b>12</b>
3.3.1	How is confidential data managed? .....	12
3.3.2	P3P recommendation .....	12
3.3.3	Privacy and personalisation.....	13
3.4	<b>Security .....</b>	<b>14</b>
3.4.1	Prevention considerations .....	14
3.4.2	Secure Sockets Layer.....	14
3.5	<b>Profiling and Data mining.....</b>	<b>15</b>
<b>4</b>	<b>User Profile main features.....</b>	<b>16</b>
4.1	<b>Overview .....</b>	<b>16</b>
4.2	<b>Main functional features .....</b>	<b>17</b>
4.2.1	General.....	17
4.2.2	Bookmarks .....	17
4.2.3	Information travel bag .....	17
4.2.4	Tracking users .....	18
4.3	<b>Foreseen Architecture .....</b>	<b>18</b>
4.3.1	Several Components to store User Profile data.....	18
4.3.2	A Distributed architecture .....	18
4.3.3	Multiple databases.....	19
4.3.4	Studied Interfaces .....	19
<b>5</b>	<b>Conclusions.....</b>	<b>19</b>
<b>6</b>	<b>References.....</b>	<b>20</b>
<b>7</b>	<b>Annexes.....</b>	<b>20</b>
7.1	<b>Related white papers .....</b>	<b>20</b>
7.2	<b>Personalization &amp; Privacy Survey .....</b>	<b>21</b>

# 1 Introduction

Visitors to, and users of, recreation and protected areas need geographic information to plan their trips safely and effectively. Moreover, they want to have contextual information on topics of interest whilst they are involved in outdoor activities in these areas. To deliver the best information service before the trip and/or on mobile devices, the WebPark project has to take into account the personalisation characteristics of the user, especially the privacy/ security issues.

The main aims of this report are to identify the user requirements for privacy and security in the WebPark project and to identify the design constraints this will place on the WebPark architecture.

The purpose of this document is to:

- assess and determine the requirements for user profile management based on user needs reports or other existing research related material;
- give an overview of privacy/ security recommendations;
- propose a high level architecture for user profile management and technologies that could be used in the WebPark project.

As a consequence, this document is organised as following:

- the first section (this one) gives an overview of the main objectives of this report;
- the second section analyses the user needs presented in the D2.2.1a Service Template document in order to extract requirements for web-based user profiles;
- the third section gives some recommendations on privacy and security;
- the fourth section deals with user profile components specifications : main features, high-level architecture and foreseen technologies.

Note that this document makes certain assumptions about the architecture of the WebPark service that may be altered by later design decisions. Also, due to the later than expected start of the project it was not possible to incorporate user testing of security options from WebPark test areas into this report. This perspective will be introduced with the 2002 summer testing using participant observation techniques in the Swiss National Park.

The WebPark project recognizes that the design of appropriate privacy and security measures is of central importance to the success of any mobile information service that involves tracking its users. Users must have confidence in the policies and procedures of the service and they must be assured that it complies with the appropriate legal safeguards. The WebPark project also sees user transparency and control over the stream of tracking information as essential to user acceptance of its location-based services. These issues will have a major influence over the success of the project.

## 2 Users requirements assessment

### 2.1 Introduction

This section deals with the user needs collected and reported in the D2.2.1a Service Template document. A first analysis of this document has shown that user requirements can be classified into several categories:

- ◆ Identification: personal information needed to identify the user;
- ◆ Use cases: different ways of accessing the information;
- ◆ Topics of interest: a classification of the information that can be accessed through the WebPark service,

- ◆ Location: positional information that needs to be associated to a user in order to facilitate location-based services.
  - ◆ Multi-language capability.
- This personalisation and preference information will need to be stored in a central User Profile where it can be accessed by a User Profile Manager component.

## 2.2 Identification

The user needs report has shown that users have diverse interests, suggesting that users will develop individualistic usage behaviour. As a consequence, the users of the WebPark service need to be identified and characterised so that personalisation can be developed. The survey presented in section 7.2 shows that a user will normally be happy to give:

- Name;
- Address;
- Phone number;
- E-mail address

to a service provider. Of these kinds of data, personalisation can best be associated with the phone number, so that a single user can have multiple identities.

## 2.3 Use cases

The user of the WebPark service will most likely access the service for two distinct reasons:

- To plan the trip, requiring the ability to find or compute information on the national park/ protected area such as flora, fauna, walking directions and so on. The user may wish to store the information they have chosen in a kind of 'information travel bag' in order to retrieve it more easily later;
- To acquire location-based information, meaning that the user would retrieve information of interest on fauna, flora, navigation and so on, depending of their position. The user would need to retrieve or update what they had placed in the 'information travel bag'.

In the first use case, the user will retrieve data on request (pulled data), while in the second use case, the user might also consider pushed data, proposed to him by the WebPark service. This information will adapt to their preferences, and will depend on their location. The user will also be able to customise the presentation of information although this will depend on the nature of the information and on the device being used (PDA, mobile phone etc.). The User Profile Manager will need access to the user's personalisation information in order to carry out this customisation.

Since usage patterns and stored preferences are also confidential this kind of data should be securely held and only released under conditions notified to and accepted by the user.

## 2.4 Topics of interest

The information available to national park users can be arranged in a variety of different ways. The user can be asked to specify the most important topics, which can be used to filter any pulled or pushed information. The possible approaches to information arrangement are discussed in the D2.2.1a Service Template document. User preferences of this kind should be secure and confidential.

## 2.5 Location

To deliver location-based services requires the storage of location information at some user-controlled level. This means that User Profile has to be able to store spatial-temporal movement of the user. This record of location and movement must be secure as it may be important for safety reasons and must be kept confidential at all times. It can be suggested from the User Needs Report and some early 'participant observation' exercises that users should be able to control the storage of this information through the provision of a location 'control panel'.

## 2.6 Multi-language capability

The WebPark service will be offered in a number of languages. It is important to store the user preferences and profile data in a language-independent way. With this approach, a transformation of the personalisation data can easily be created for each new language. Consequently, we need tools to translate information into the appropriate language, dependent on user preferences.

## 2.7 Conclusion

In the following table, we try to assess the design requirements of the User Profile in terms of information that needs to be stored.

REQUIREMENT	USER PROFILE FEATURE	FUNCTIONS
<b>GENERAL REQUIREMENT</b>		
Registration in the WebPark service with a new profile	Wizards	Set the different parameters of the WebPark computer (colour, network device, kind of screen, ...)
Personal information	Identity manager	The user profile will be used for the authentication of the end-user in the system
Device	Identity manager	The device owned by the user will help to determine the way that the requested information is presented
Mother language	Identity manager	To be able to display information in the mother language
<b>'VISITOR LOGISTICS' REQUIREMENT</b>		
Lodging/Accommodation	Preferences	The user profile knows the main preference of the user
Store/retrieve Accommodation	Travel bag	The user profile can save interesting information in a personal information bag.
Transport and Access	Travel bag	The user profile can save interesting maps in a personal information bag.
Store/retrieve information of interest	Travel bag	The user profile can save information on topics of interest in a personal information bag in order to retrieve it later quickly
<b>'PARK INFO' REQUIREMENT</b>		
Administration	Preferences	Use as a filter for pulled or pushed data
Research activities	Preferences	Use as a filter for pulled or pushed data
<b>'RECREATIONAL ACTIVITIES' REQUIREMENT</b>		
Walking/Trails	Preferences	Use as a filter for pulled or pushed data
Offered activities	Preferences	Use as a filter for pulled or pushed data
<b>'NATURE' REQUIREMENT</b>		
Flora	Preferences	Use as a filter for pulled or pushed data
Fauna	Preferences	Use as a filter for pulled or pushed data
Geomorphology	Preferences	Use as a filter for pulled or pushed data
Weather and climate	Preferences	Use as a filter for pulled or pushed data
Thematic maps	Preferences	Use as a filter for pulled or pushed data
History	Preferences	Use as a filter for pulled or pushed data
Landscape	Preferences	Use as a filter for pulled or pushed data
Etc.		
<b>'LOCATION' REQUIREMENTS</b>		

REQUIREMENT	USER PROFILE FEATURE	FUNCTIONS
GPS Position	Tracker	Use to retrieve information associated to the position
Elevation	Tracker	Use to retrieve information associated to the position
GPS History	Tracker	Route prediction
Spatial-temporal behaviour	Tracker	Spatial envelope

This list can evolve with the detailed definition of the User Profile.

## 3 Privacy/ security issues

### 3.1 Introduction

In section 2 it was seen that location-based services and customisation of the service will imply the use of personal data. Prevention of the interception/ surveillance of this information or the unlawful access to it and the legal protection of individuals with regard to their personal data are the major concerns of service providers and users in the information society. To prevent unlawful access to computer facilities and online interception by hackers or incidental access and the possible manipulation of contents is a matter of security through prevention measures (see section 3.4 below).

Protection of privacy is also subject to legal regulation in the area of access to personal data stored in digital form.

### 3.2 Legislation and regulation

#### 3.2.1 Legal protection of privacy

Article 12 of the Universal Declaration of the Rights of the Man of 10th December 1948, established that no-one should be the object of arbitrary interference into their private life, into that of their family, into their home or into their correspondence, and, also that all persons are entitled to the protection of the law against such interference. These principles were reaffirmed in Article 8 of the European Convention on Human Rights, which have been incorporated in the statutes of most European states.

#### 3.2.2 National laws

Currently each European state has its own privacy laws, although the EU Directive on the 'Protection of individuals with regard to the processing of personal data' (95/46/EC) has standardized the protection of personal data in computer form within the EU. This has led to the setting up of Data Protection offices in each EU state:  
([http://europa.eu.int/comm/internal\\_market/en/index.htm](http://europa.eu.int/comm/internal_market/en/index.htm)).

In the UK, in addition to the Data Protection office there is a law of confidence extended by the European Convention of Human Rights. This states that any information given in confidence from one party to another must be respected, and if this provision is violated then the injured party can sue for damages. There must be a guarantee in WebPark that, for example, current location must not be disclosed to a 3rd party without permission to avoid a breach of confidence. Recent court cases have clarified the right of individuals in the UK to sue for breach of confidence when there is no national security or public interest defence by the discloser.

In France, the Commission Nationale de l'Informatique et des Libertés (CNIL) manages the rules for privacy and security. In both Spain (Article 18-4) and the Netherlands (Article 10-2) the constitutions contain written guarantees that the use of personal data should not infringe privacy.

In the United States of America an alternative path to the protection of privacy has developed since the end of the 19th century (Brandeis and Warren 1890). There have been several common law judgements and three statute laws, including the Freedom of Information Act (FIA) of 1966, the Privacy Act (PA) of 1974 and the Electronic Communication Privacy Act (ECPA) of 1986 in this area. The ECPA contains provisions on provisions of using wireless electronic devices.

Canada recently enacted the 'Personal Protection and Electronic Documents Act ' (April 2000), (<http://law.about.com/library/briefs/ucanadaprivacy.htm>). This legislation covers personal data collected during the transfer of commercial information or during government handling. No consent of the user is necessary when that information is for personal/ domestic use or legal, journalistic, artistic or literary use. It permits the dissemination of: names, titles, business addresses but prohibits organisations from using personal data without the previous consent of the user concerned, except for situations where life is under threat, during legal investigations and in authorised scientific research.

The current status of privacy provisions the world can be found at:  
<http://www.privacyinternational.org/survey/phr2000/countrieshp.html#Heading7>

### 3.2.3 European Union Data Protection Directive

#### 3.2.3.1 Protection privacy in the European Union

Processing of personal data means any operation or set of operations which is performed upon personal data, for example as indicated in article 2b of Directive 95/46: 'the collection, recording, organisation, storage, adaptation, or alteration, retrieval, consultation, disclosure by transmission, dissemination, alignment or combination, blocking, erasure or destruction', whether or not carried out automatically.

The implementation of the EU Data Protection Directive is different in each state but it is based on a number of basic principles:

- ◆ fair obtaining and use of data;
- ◆ data up to date;
- ◆ accurate and not excessive;
- ◆ held securely.

The user is entitled to object to data processing if the personal information is passed to any third party, or if damage/distress is caused, or if the information is sensitive and is used without consent. Currently 'sensitive' means racial, political, health, religious, moral or trade union affiliations and does not include location specifically.

The EU data protection directive also prohibits secondary uses of data without informed consent:

- ◆ Creating personally-identifiable online profiles will have to have full informed consent from the user;
- ◆ Upfront notice must be given when data is collected – no techniques like web bugs will be allowed;
- ◆ No transfer of data to non-EU countries unless there is adequate privacy protection there.

#### Guarantee of the right to privacy

As a principle of data protection, it is specified expressly that the holders (Internet service providers, mobile operators, etc) must fulfil a series of requirements to guarantee the right to privacy of concerned individuals about their stored data.

These regulations include the requirement to:

- Inform the Data Protection Registrar about the creation of the file in the territory where the file is located. This information will identify the holder of the file, the processor of the data, its physical location, the aim of its creation, the types of data stored, the procedure for collection, the structure of the file, the computer applications using the data, the uses to which the file will be put and the measures of security taken.
- Establish the mechanisms of accuracy control for the data stored and to foresee the means of their upgrade, modification or deletion, when it is no longer needed for the ends for which they were gathered.
- Guarantee the security of the data held in their power and be responsible for the activities of the staff in charge of maintaining the file, and the security measures adopted in connection with the physical and logical storage of data and the control of unauthorised access to the information.

#### **Processing of personal files**

The processing of the personal files can be carried out if their collection is compliant with the data protection principles. Processing is regulated such that:

- Personal data must only be processed insofar as is pertinent and necessary to the specific purpose of the processing and the consent given by the user;
- The data obtained by the processing of the file will be used exclusively for the foreseen purpose, with no use or storage for any other purpose other than that previously consented to;
- Individuals who are the subject of processing must be informed in an unequivocal way about the existence and purpose of the file for which their data are held; they must consent to the processing that is requested; and they must be afforded rights of access, rectification, deletion and opposition to the processing of their data.

#### **Liability for treatment**

The person responsible for the files will assume liability for non-fulfilment of any legal obligations. These will include the need to:

- Notify the creation of the file to the Data Protection Registrar;
- Treat the data according to the regulations;
- Inform the data subjects of the existence and purpose of the file and of their rights of access, rectification, deletion and opposition to the automated treatment of their data;
- Obtain the consent of data subjects as is necessary about any automated treatment of their data and obtain consent in writing when the processing is of sensitive data;
- Ensure that the data are only used for the purpose for which they were gathered;
- Maintain the data such that incorrect files are rectified, incomplete ones are corrected and unfounded ones are deleted;
- Ensure the confidentiality of the data and that the duty of confidentiality is passed on to other processors of the data;

- Adopt reasonable technical and organisational measures to guarantee the security of the data, and to adopt especially secure measures in connection with sensitive data;
- Dissociate the data obtained so that the data can not be linked to other data sets in a damaging way; and
- Control transfers of data, especially those made to third countries.

In the case of non-fulfilment of any of the above provisions ones, the data subject has a right to receive compensation for any damage suffered.

### 3.2.3.2 Positioning and privacy

The use of wireless navigation and positioning systems, in particular Global Position System (GPS), has raised a number of legal questions of great importance during the last two decades. See for example Raper (2002), and Sovocool (2000) at:  
[http://thelenreid.com/articles/article/art\\_37.htm](http://thelenreid.com/articles/article/art_37.htm).

A record of the individual's position in space through time constitutes a very sensitive dataset and can reveal many details of their private life. However, the position relates to a mobile device location and not necessarily the person concerned. For full details of the geolocation technology see the D2.1.1b Geo-accuracy document.

For location-based services, whichever geolocation technology that is used, two approaches can be used to guarantee the right to the privacy of the user:

- a. To use technologies that allow the separation of the user identification from the user positioning, or
- b. To get the user's express consent.

In Article 6 of the European Directive 97/66/EC, the regime for the treatment of personal data by the telecommunications operators is set out in relation to traffic data and billing. The precepts about processing are contained in the annex of the Directive. The data used to establish location for the purpose of the telecommunication should be stored only as long as is needed to complete billing of the service to the user, unless specific consent has been given. However, recent increases in security concerns have led some governments to require operators to keep this data for much longer periods for the purpose of national security.

In similar ways the US Wireless Communications and Public Safety Act 1999 prohibits the unauthorised disclosure of positioning information by telecommunication operators except in strictly regulated circumstances.

In the overview of their Mobile Positioning System (MPS) (<http://www.ericsson.com/mps/>), three levels of subscriber protection of privacy are contemplated:

- a. A request and authorisation to log on to the MPS (for service providers);
- b. Retrieval of information on a user's location must be authorised; and
- c. Subscribers can choose not to be located.

### 3.2.3.3 Conclusion

The implications of this regulation and legislation are that WebPark must develop compatible business models from the outset. It is also wise to assume that location is 'sensitive' even though according to the law no special extra consent for the service provider needed to use this information.

### 3.3 Technical aspects of privacy on networks

#### 3.3.1 How is confidential data managed?

On the World Wide Web, applications use profile and personal data to check transactions or to customise services according to the user's preferences. In this context, privacy is important to ensure consumer trust. But privacy policies are often difficult to understand, hard to find, and take a long time to read. Many policies are changed frequently without notice. WebPark must aim to simplify this process yet still allow the user a high level of involvement in maintaining the data.

How is this confidential data passed from client to server? Servers can get data from the browser, using several methods to personalize end-users web page. Two kinds of data collection exist:

- ◆ Getting data in real-time:
  - By the standard HTTP Protocol: last URL visited, IP, Kind of browser, domain name, organization and so on. Hence, amazon.com compute their first page depending on the URL where you come from!
  - By browser advertisements. They may hide a small application that gets specific data;
  - By cookies: it's possible to get access permissions;
  - By web bugs : it's possible to get some passwords.
- ◆ Getting data with some time delay:
  - Organizations provide services based on 'double click' tracking. When a user clicks on banners, before accessing the desired web site, another web site logs the user context, the site from which referred and the next link. These organizations collect this data and sell databases to data mining operators.

The WebPark service needs to find solutions for each problem, although some of them are complicated :

- ◆ **Seal Programs** ensure a label for affiliated web sites : TRUSTe, BBBOnline, CPA WebTrust, Japanese Privacy Mark ...
- ◆ **Laws and regulations** protecting consumers: the European Data Protection Directive requires all European Union countries to adopt similar comprehensive privacy laws.
- ◆ **Privacy tools** allow the user to bypass some tracking.

The key lesson is that the global architecture of the WebPark service must design the management of confidential data into its basic architecture.

#### 3.3.2 P3P recommendation

On 28th January 2002, the W3C consortium released its recommendation for privacy preferences. The Platform for Privacy Preferences Project (P3P) enables a web site to express their privacy practices in a standard form so that they can be retrieved automatically and interpreted easily by user agents. WebPark could adopt this approach to privacy management and communication.

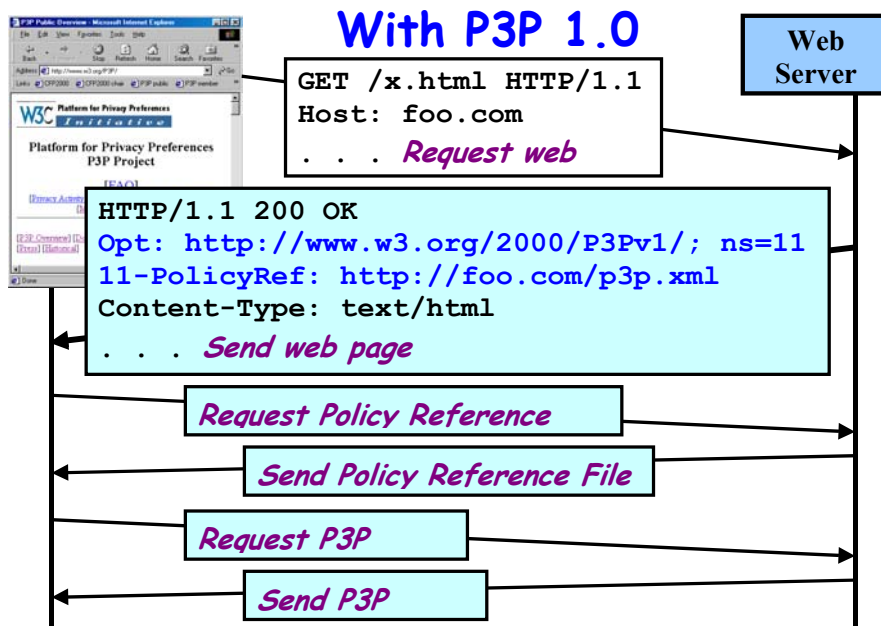


Figure 1: P3P communication

Advantages:

- ◆ The system **informs users** before the release of any personal information;
- ◆ Web sites disclose their privacy practices in standard machine-readable formats: **User agents can test** whether a **site's practices** are compliant with a law or code;
- ◆ Implementation is possible under the existing HTTP protocol.

Disadvantages:

- ◆ **Doesn't** provide **secure way** to transfer data;
- ◆ More **data to transfer** between mobile client and server!
- ◆ The browser needs **P3P built in**. Then, web browsers automatically retrieve P3P privacy policies and compare them to users' privacy preferences.
- ◆ In any case, the browser notifies policies to the end-users. It may involve a communication **with pop-up messages**. This kind of interface may annoy end-users.

With this protocol, end-users may set their preferences on each used browser used, although no browser yet supports this protocol. The most practical solution for the WebPark project would be to define a P3P approach that is built in and linked to our user profile component.

### 3.3.3 Privacy and personalisation

The legislation and regulation of privacy needs to be extended to personalisation information. On the WebPark service this means that the user profile must be accessible only by the user, that any tracking or data mining of behaviour is proportionate and effective, and that the service must not pass on private information to 3rd parties without agreement. There will be a number of new aspects to privacy management for location-based services relating to the tracking of location that will have to be respected. However, WebPark aims to engage users with its services since the applications on offer will explicitly require tracking, e.g. travel advice in a potentially dangerous area or a desire to ask location-based questions.

Attitudes to privacy of information can be ascertained through surveys such as the one carried out by Personalisation.org. These results are given in detail in Annex 7.2. This study identified important features in the privacy domain:

About the level of agreement:

- ◆ It helpful and convenient when a web site remembers basic information about user (73%). But it seems that users who like to store personal information (preferences etc.) are much fewer (only 50%).
- ◆ Users don't like online solicitations for shopping based on previous experience (69%).
- ◆ An online shopping is less easy because of control. The offline mode is better (67%).
- ◆ Before registering on web sites, most users read the privacy statement (51%). They say that privacy statements aren't easy to read (62%) but they think that a privacy statement is necessary to share personal information (58%).

About personal information:

- ◆ People are happy to leave email, name and address without worrying too much about the consequences. But storage of hobbies or personal information gives rise to greater level of concern;
- ◆ People are reluctant to leave a Credit Card Number or a Social Security Numbers.

## 3.4 Security

### 3.4.1 Prevention considerations

A first step in information transfer security is the development of confidence between the parties in communication. Authentication is a manner of providing such confidence. A definition of authentication in the online environment can be 'with the intent to sign a record to execute or adopt an electronic symbol, sound, message, or process referring to, attached to, included in, or logically associated or linked with that record' (section 102 of the US Uniform Computer Information Transaction Act 1999).

Concerning electronic and digital signatures, several legal approaches by States around the world have been implemented explaining how to use a form of electronic signature:

[http://www.ilpf.org/groups/analysis\\_IEDSII.htm](http://www.ilpf.org/groups/analysis_IEDSII.htm)

These include the Electronic Signatures in Global and National Commerce (E-SIGN) Act (USA), General Law of e-signature of Japan Personal Information Protection and Electronic Documents Act (Canada), EU Electronic Signature Directive, 1999:

<http://www.ict.etsi.fr/eessi/Documents/e-sign-directive.pdf>

Except in the US and Canada these provisions do not have a mandatory effect and they are confined to intra-national agreements. International proposals include: the UNCITRAL Model Rules on Electronic Signatures, Internet Law and Policy Forum (ILPF), International Telecommunication Union (ITU) and others.

Electronic signature making is generally based on encryption or cryptographic techniques. Legal recognition needs the participation of a certification body. The EU Directive on a Community Framework for electronic signature (1999/93/EC) is based on the 'no denial' principle implemented in the Member States regulation. A standard form of qualified electronic signature (EESSI) is being developed by European standardisation bodies (CEN, CENELEC and ETSI).

### 3.4.2 Secure Sockets Layer

Secure Sockets Layer (SSL) is a protocol for transmitting private information via the Internet. First, the SSL protocol is designed to provide privacy between communicating applications (a client and a server). Secondly, the protocol is designed to authenticate the server, and

optionally the client. SSL works by using a private key to encrypt data that's transferred over the SSL connection.

In fact, the SSL protocol is a *de facto* industry standard security protocol, providing:

- **Authentication:** the identities of the parties communicating are verified, ensuring that the source and the destination of data are trusted.
- **Confidentiality:** all data being transferred is encrypted to a known level of security, guarding against third party snooping.
- **Message Integrity:** the data received is guaranteed to be the data that was sent, preventing a third party from modifying the data during transmission.

An interesting characteristic of the SSL protocol is that it is 'application protocol' independent. A 'higher level' application protocol (e.g. HTTP, FTP, IIOP) can layer on top of the SSL protocol transparently.

As a consequence, SSL versions of useful software for WebPark already exist, for example, Apache HTTP server, or Orbit (from Iona), and FTP servers. They are used to ensure secure transfers of data through the Internet as well as private information concerning the end-users (like credit card numbers). Moreover, in the case of WebPark, SSL ensures that the customer is sure of the origin of the data they buy.

Generally, the electronic commerce service consists of three main elements: the Shopping Basket Manager, the Invoicing, and the Payment Terminal.

- ◆ The **Shopping Basket Manager** is a clearinghouse for shopping basket objects, i.e. it offers methods for creating and releasing basket objects. A shopping basket object provides methods that enable customers to choose or remove datasets;
- ◆ The **Invoicing object** provides a method that computes an invoice from a given shopping basket object; and
- ◆ The **Payment Terminal** object effects a bank transaction from an invoice computed by the Invoicing object.

Today, it's too early to know if the WebPark service will use these objects. However, if we decide to employ this technique, the user profile would be the logical solution for the storage of data, because information about the e-commerce is important to developing the user profile and to categorize the type of service offered.

### 3.5 Profiling and Data mining

The WebPark service will require two kinds of application. The first one will focus on contents, which will attract users by the value of the information they provide. The second kind of application will focus on electronic commerce which will allow the user to customize the service they get and this will require the service to anticipate their needs. In these two cases, we have to ensure that the user profile can provide these new tools:

- ◆ **Save all the queries:** if we save the different queries of the user, we can analyse them using several different methods (Clustering, Classification, Regression, Cross-selling) to improve the quality of our WebPark services;
- ◆ Improve the quality of the system with **customisation:** according to the ability of the end-users, we think that it may be interesting to provide a dynamic interface like Personal Yahoo. In this case, the user profile can offer several ways to do it.
- ◆ Detect the kind of profile to **personalise** services and interfaces: before customisation, the system can offer its interface according the detected profile.

- ◆ **Keep in contact** with end-users: the user profile has to store all the information to contact end-users at any time. It's important for several reasons like emergency, push services.

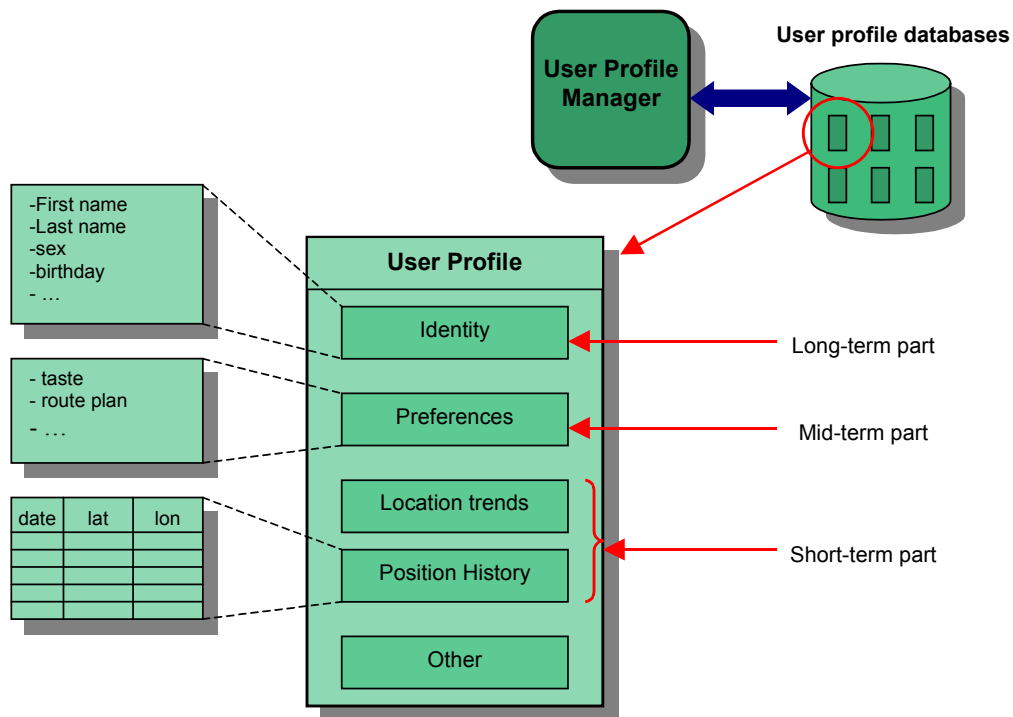
There is also a substantial opportunity for WebPark to develop spatial-temporal data mining to allow the user's movement to be summarized given appropriate spatial-temporal data storage in the profile. This can be conducted at two levels: basic geographical on-line analytical processing (G-OLAP), and full spatial-temporal trends detection. Such techniques will also permit the reduction in detailed location data being stored (e.g. by storing movement 'envelopes') and therefore can reduce the privacy concerns about a detailed profile. Hence this approach could be used to determine whether the user was familiar with the area from summarised data to avoid giving information that the user might find obvious.

Each of these functions will require careful grounding against our privacy and security policy.

## 4 User Profile main features

### 4.1 Overview

The user profile component has to store all the relevant information on a user. According to the users requirements (see section 2), the data stored in the user profile can be divided in several parts as shown in the following figure:



**Figure 2: What a user profile component stores**

- A long-term part which deals with the user identity: first name, last name, sex, birthday, address, job title, device characteristics and the phone number;

- A mid-term part that contains user preferences: parameters for certain services, tastes, etc. These preferences can be customised by the end-user.
- A short-term part that holds the location-based information for a user: his last geographical positions and elevation received, and also his location trends computed by the system.

## 4.2 Main functional features

### 4.2.1 General

The main functional features of the user profile component are described below:

Functional Features	
Identity	<ul style="list-style-type: none"> <li>➤ Create a profile</li> <li>➤ Modify personal information (password etc.)</li> <li>➤ Delete profile</li> </ul>
Preferences	<ul style="list-style-type: none"> <li>➤ Add preference</li> <li>➤ Modify preference</li> </ul>
Position History	<ul style="list-style-type: none"> <li>➤ Add a set of positions and elevations</li> <li>➤ Read a set of positions and elevations</li> <li>➤ Delete a set of positions and elevations</li> </ul>
Spatial Temporal Behavior	<ul style="list-style-type: none"> <li>➤ Add current activities computed with positions</li> <li>➤ Read current activities</li> <li>➤ Delete current activities</li> </ul>
Configuration	<ul style="list-style-type: none"> <li>➤ Add a point (a map, accommodation) to personal 'information travel bag'</li> <li>➤ Modify a point from 'information travel bag'</li> <li>➤ Delete a point from 'information travel bag'</li> <li>➤ Add an URL to personal bookmarks</li> <li>➤ Modify an URL from the bookmarks</li> <li>➤ Delete an URL from the bookmarks</li> </ul>

### 4.2.2 Bookmarks

Mobile users have new kinds of user interface needs. Given the difficulty of managing bookmarks, everybody knows that it's difficult to hold a separate set of bookmarks on each computer used. The main idea of this feature is to provide a ubiquitous set of bookmarks stored by the user profile component.

During the preparation of the trip, users browse their profile and insert their most important URLs in the area they plan to visit. The URL of the bookmark appears directly on the web page of the system (fixed or mobile computer). It's easy to delete or modify them.

### 4.2.3 Information travel bag

The 'Information travel bag' is a new concept specially designed for end users that would like to prepare their trip. Before travelling, if they need to see maps, to select hotels, to search monuments and so on then they will need a specific folder to store their useful results. Otherwise, during their trip, they will need to search again.

The Information Travel bag' is a set of web pages that manages the data stored in the profile. It's possible to store all kinds of data like URLs, addresses and mapping.

#### 4.2.4 Tracking users

Most of the mobile services need a position as a foundation of geographic context. The User Profile may allow the storage location data from GPS related data such recent envelope and conjectured activities.

### 4.3 Foreseen Architecture

#### 4.3.1 Several Components to store User Profile data

For privacy/ security purposes, the user profile must be split into at least three main components:

- Identity manager, which will be in charge of long-term data. It will store the files with user's overall spatial-temporal behaviour;
- Preferences manager, which will be in charge of mid-term data. It will store the topics of interest of a user.
- Position manager, which will be in charge of short-term data. It will store the last position of the user, the history of position, elevations associated with these positions. This component is a spatial-temporal repository, because all positions are associated with a date.

Another component will be required to manage 'information travel bag' and personal bookmarks.

All these components store the data in a different database. The link between all databases will be an identification key given by the Identity manager every time a new profile is created.

This separation of datasets answers some of the protection of privacy/ security concerns: 'The person responsible for the files will assume the implicit liability for the non fulfilment of their obligations'. Among those that it is necessary to highlight include those of:

- The need to dissociate individual users and the datasets collected;
- The need to record location so that location and identity are only matched for specific purposes.

#### 4.3.2 A Distributed architecture

We recommend that the user profile components are based on a distributed architecture. The different components will be accessed by all the other components. Each component of the user profile manages the pool of connections to its database. It ensures the sharing and the security of the personal data.

This distributed architecture has the following main characteristics:

- ◆ **Scalability:** they are distributed components. They may be distributed among several hosts in the same way;
- ◆ **Few development constraints:** the components already exist. Partners can use them in heterogeneous framework. It's easy to integrate them;
- ◆ **Performance:** the user profile components manage all the connections to the profile. They open several connections to their databases, and they share the resources to quickly ensure all the queries succeed;
- ◆ **Security:** Users and components access data via the user profile components. These components are on the same network as the databases: it's impossible to access the databases directly without it;

Moreover, the databases can be installed on different servers: Identity information, preferences information, positional information can be physically separated. This increase privacy protection and security of personal data.

### 4.3.3 Multiple databases

Each user profile component ensures access to a database. The database is not directly accessible; all the queries are submitted to the component, which is the only one, allowed to access it.

Data are stored in relational databases and the access is ensured by an API. Therefore all the applications that handle user profiles are built on top of this API, which will be kept confidential.

By using different databases, we split up information concerning the end-user: identity, preferences, positions information are separated. No direct link is available.

### 4.3.4 Studied Interfaces

#### 4.3.4.1 IIOP protocols

All the interfaces with the user profile components may be via Java RMI or via CORBA. Therefore, the components constituting the user profile don't need to stay on the same machine as other components.

A RMI service enables remote accesses to user profiles. An RMI interface implies that the components will be written in Java. In addition, a CORBA service can be developed. The interest of CORBA is that it lets the service be independent of the implementation language and the type of platform. Moreover CORBA facilitates the integration job since a CORBA component is easier to reuse and/or to integrate to a heterogeneous system than another type of component.

#### 4.3.4.2 LDAP

The issue of the user profile (and in particular the Identity manager) is linked with the issue of directories (or address lists) because they both store information about the identity of users. The LDAP protocol is a standard way to access directories. It provides an interface to allow access all kind of data. Consequently, several directory services could communicate with our user profile component if we decide to implement this protocol. But do we need it?

It's early to answer this question because it all depends on the architecture adopted and our end-users needs. But if we integrate LDAP, we will provide:

- ◆ **An Interchange format (LDIF):** the LDAP protocol allows the LDIF import/export of data;
- ◆ A way to **quickly access** the data: LDAP uses a directory information tree to improve the performance of access;
- ◆ A **secure form of access:** LDAP uses ACLS and ensures the agreement of data with the SSL protocol. Current projects are working on a way to replace NIS and NIS+ by the LDAP protocol;
- ◆ **JNDI services:** JNDI is a standard API to communicate between LDAP and client applications. JNDI uses the same concept as JDBC for the database.

## 5 Conclusions

This section gives a brief summary of some of the key lessons of previous experience, from the literature review and from an assessment of regulation and the law. Together they provide a number of guidelines for the further design of WebPark.

These points can be considered to be the foundations of a proposed WebPark privacy and security policy:



- ◆ The WebPark service should only gather the data necessary to deliver specified services and it should always seek to justify carefully any data collected and stored in the user profile, (especially on location) by constant communication with the user;
- ◆ For the WebPark service the profile must only be accessible by and for the user, and the service must not pass on private information to third parties without agreement;
- ◆ The user should have access to the user profile data to check its accuracy, edit its contents or to update preferences as it is suggested that this will engage the user with the database and build confidence in the service;
- ◆ The default operation of the service should be to provide information in 'pull' mode on user demand, with the availability of 'push' services for user-specified services;
- ◆ Location data should be not be collected by default, but if authorised should be collected independently of the device identity and matched when necessary;
- ◆ Data mining and profiling will require careful grounding against our privacy and security policy, especially with respect to location, which should be summarised wherever possible to reduce the physical and privacy overheads of the location information;
- ◆ The personalisation information in the user profile must be configured to take advantage of protocols such as P3P to enable transparent access to the WebPark privacy policy;
- ◆ The architecture of the WebPark service should be designed to strictly guard against unauthorised access to the user profile; and
- ◆ The regulation and legislation of privacy must be built into the business models that the WebPark service develops from the outset.

## 6 References

[CNIL]	<a href="http://www.cnil.fr/thematic/them01.htm#necessaire">http://www.cnil.fr/thematic/them01.htm#necessaire</a> , <a href="http://www.cnil.fr/thematic/them01.htm#protection">http://www.cnil.fr/thematic/them01.htm#protection</a> , <a href="http://www.cnil.fr/thematic/them01.htm#100sites">http://www.cnil.fr/thematic/them01.htm#100sites</a>
[Brandeis L.D. and Warren, A.]	The right of privacy. Harvard Law Review, 1890
[Sovocool, D.R.]	<b>GPS/Wireless: Legal Issues Related to Wireless Navigation and Location Systems, 2000</b> <a href="http://thelenreid.com/articles/article/art_37.htm">http://thelenreid.com/articles/article/art_37.htm</a>
[Raper, J.F.]	<b>Location privacy: a new challenge for GI science. AGILE 2002, Palma, Mallorca., 2002.</b>
[P3P]	<a href="http://www.research.att.com/projects/p3p/">http://www.research.att.com/projects/p3p/</a> , <a href="http://www.w3.org/P3P/">http://www.w3.org/P3P/</a>
[Seal Programs]	TRUSTe, <a href="http://www.truste.org">http://www.truste.org</a> , BBBOnline, <a href="http://www.bbbonline.org">http://www.bbbonline.org</a> , CPA WebTrust, <a href="http://www.cpawebtrust.org/">http://www.cpawebtrust.org/</a> , Japanese Privacy Mark, <a href="http://www.iipdec.or.jp/security/privacy/">http://www.iipdec.or.jp/security/privacy/</a> .
[Personalisation Consortium]	<a href="http://www.personalisation.org">http://www.personalisation.org</a>

## 7 Annexes

### 7.1 Related white papers

[ E. BEINAT ]	Privacy and location based services
---------------	-------------------------------------

## 7.2 Personalization & Privacy Survey

A survey from the web site <http://www.personalisation.org> :

1.	How many hours a week, on average, do you surf the web (not including e-mail)?	Numbers	Percentage
	0-1	122	3%
	2-7	1548	34%
	8+	2850	63%
	Total	4520	100%

2.	How much have you spent online within the last 6 months?	Numbers	Percentage
	\$0	662	15%
	\$1 - \$100	1657	37%
	\$100+	2194	49%
	Total	4513	100%

3.	Are you currently a registered user with any web sites (providing name and email address)?	Numbers	Percentage
	Yes	4231	94%
	No	274	6%
	Total	4505	100%

4.	Have you provided personal information to any web sites (address, phone number, hobbies/interests)?	Numbers	Percentage
	Yes	4231	97%
	No	274	3%
	Total	4511	100%



5.	Please indicate your level of agreement with the following statements:	1 Strongly disagree	2 Disagree	3 Neither agree or disagree	4 Agree	5 Strongly Agree
1.	Before registering on web sites I always read the privacy statement.	9% (386)	18% (829)	22% (1008)	33% (1473)	18% (820)
2.	Most privacy statements on web sites are simple and easy to understand.	12% (542)	24% (1070)	26% (1189)	30% (1364)	8% (343)
3.	A privacy statement is necessary for me to share personal information.	5% (227)	11% (517)	24% (1096)	31% (1420)	27% (1231)
4.	I find it helpful and convenient when a web site remembers basic information about me (e.g., my name and address).	4% (179)	5% (224)	17% (753)	42% (1891)	31% (1415)
5.	I find it helpful and convenient when a web site remembers more personal information about me (e.g., my preferred colours, music or delivery options).	7% (324)	13% (588)	30% (1363)	31% (1389)	19% (838)
6.	Banner ads and "pop ups" are an invasion of my privacy.	8% (367)	21% (950)	35% (1560)	16% (736)	19% (867)
7.	Online solicitations (offers) from the web site hinder my shopping experience.	7% (318)	22% (1014)	40% (1790)	19% (850)	11% (514)
8.	I am willing to give information about myself in order to receive an online experience truly personalized for me.	4% (198)	11% (507)	33% (1510)	41% (1837)	10% (435)
9.	I have more control shopping online than I do offline.	9% (424)	20% (902)	38% (1718)	22% (993)	10% (441)
10.	It bothers me when web site requests personal information I've already provided (e.g., my mailing address).	5% (206)	9% (429)	23% (1042)	38% (1724)	24% (1102)

6.	What pieces of information would you provide a web-shopping site that DOES NOT provide any features personalized for you? (Please check all that apply.)		
	Name	3751	85%
	Address	2642	60%
	Credit Card Number	845	19%
	Income	855	19%
	Job Title	1416	32%
	Phone Number	1262	29%
	Hobbies/Interests	2222	51%



<b>6.</b>	What pieces of information would you provide a web-shopping site that DOES NOT provide any features personalized for you? (Please check all that apply.)		
	Social Security Number	295	7%
	Mother's Maiden Name	607	14%
	E-mail Address	3856	88%

<b>7.</b>	What pieces of information would you provide a web site that used the information that you gave them to personalize/customize your experience? (Please check all that apply.)		
	Name	4266	96%
	Address	3600	81%
	Credit Card Number	973	22%
	Income	1508	34%
	Job Title	2235	50%
	Phone Number	1988	45%
	Hobbies/Interests	3426	76%
	Social Security Number	270	6%
	Mother's Maiden Name	1001	22%
	E-mail Address	4232	95%

<b>8.</b>	Do you know what a web browser cookie is?	Numbers	Percentage
	Yes	3463	77%
	No	1035	23%
	Total	4498	100%

<b>9.</b>	If yes, please indicate which of the following statements apply to you. (Please check all that apply.)	Numbers	Percentage
	I know how to reject cookies.	1532	43%
	I generally accept cookies.	2204	62%
	A warning pop up comes on before I am given cookies.	888	25%
	I erase my cookies from my hard drive periodically.	1881	53%
	Cookies are an invasion of my privacy.	669	19%