



Évaluation des systèmes mobiles et ubiquitaires : proposition de méthodologie et retours d'expérience

Francis JAMBON, LIG/MultiCom, Grenoble




Systèmes mobiles et environnements ubiquitaires

- **Mobilité de l'utilisateur**
 - Avec le système mobile
 - Dans l'environnement ubiquitaire
- **Sensibilité à l'environnement**
 - Localisation
 - Ressources
 - ...



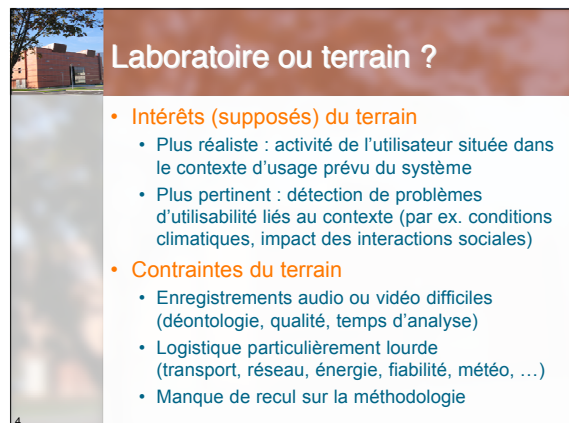

TomTom BlackBerry

→ Évaluer non pas seulement un système interactif mais un **environnement interactif**



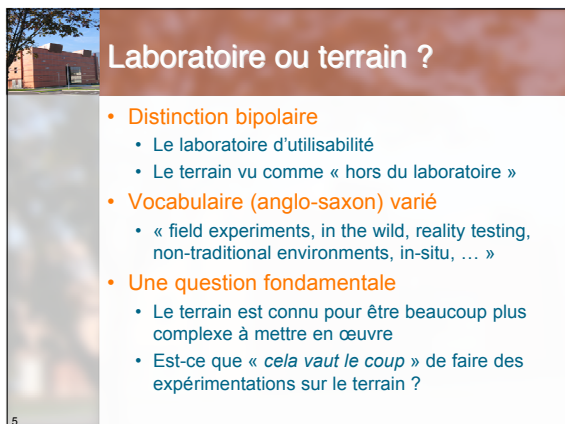
L'environnement interactif

- **Éléments de l'environnement**
 - L'utilisateur (ou les utilisateurs)
 - Le dispositif (ou les dispositifs)
 - Les tâches directement reliées à la mise en œuvre du système évalué
 - Le contexte = le complémentaire des trois autres éléments
- **Comment se situer dans le contexte ?**
 - En **laboratoire** → simulation
 - Sur le **terrain** → contexte réel



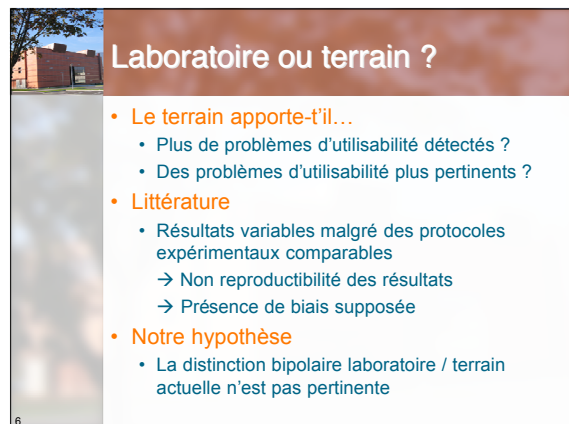
Laboratoire ou terrain ?

- **Intérêts (supposés) du terrain**
 - Plus réaliste : activité de l'utilisateur située dans le contexte d'usage prévu du système
 - Plus pertinent : détection de problèmes d'utilisabilité liés au contexte (par ex. conditions climatiques, impact des interactions sociales)
- **Contraintes du terrain**
 - Enregistrements audio ou vidéo difficiles (déontologie, qualité, temps d'analyse)
 - Logistique particulièrement lourde (transport, réseau, énergie, fiabilité, météo, ...)
 - Manque de recul sur la méthodologie



Laboratoire ou terrain ?

- **Distinction bipolaire**
 - Le laboratoire d'utilisabilité
 - Le terrain vu comme « hors du laboratoire »
- **Vocabulaire (anglo-saxon) varié**
 - « field experiments, in the wild, reality testing, non-traditional environments, in-situ, ... »
- **Une question fondamentale**
 - Le terrain est connu pour être beaucoup plus complexe à mettre en œuvre
 - Est-ce que « *cela vaut le coup* » de faire des expérimentations sur le terrain ?



Laboratoire ou terrain ?

- **Le terrain apporte-t'il...**
 - Plus de problèmes d'utilisabilité détectés ?
 - Des problèmes d'utilisabilité plus pertinents ?
- **Littérature**
 - Résultats variables malgré des protocoles expérimentaux comparables
 - Non reproductibilité des résultats
 - Présence de biais supposée
- **Notre hypothèse**
 - La distinction bipolaire laboratoire / terrain actuelle n'est pas pertinente

Notre hypothèse

- La présence de **personnes** (facilitateur, caméraman, etc.) et **dispositifs** (caméras, etc.) liés à l'observation apporte un biais
Remarque : ce phénomène est un biais bien connu des études en ethnologie
- Il n'existe pas deux, mais plutôt trois approches
 - Le laboratoire
 - Le terrain
 - La situation réelle

7

Approches de l'évaluation

- Le laboratoire d'utilisabilité**
 - Environnement simulé (par exemple via la technique du Magicien d'Oz) et bien contrôlé
 - Mobilité limitée à une salle d'expérimentation
 - Courte durée (quelques heures au max.)
 - Méthodologies et techniques validées



8

Approches de l'évaluation

- Le terrain**
 - Environnement réel et partiellement contrôlé
 - Mobilité limitée à un espace limité
 - Courte durée (quelques heures au max.)
 - Méthodologies et techniques partiellement validées



9

Approches de l'évaluation

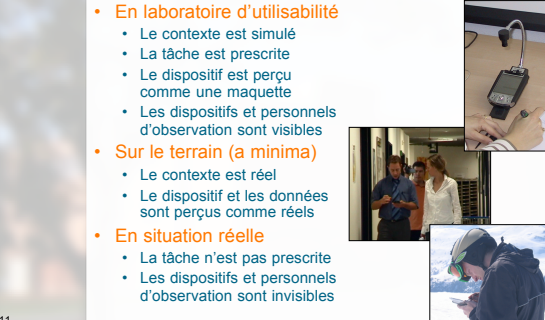
- La situation réelle**
 - Environnement réel et difficilement contrôlable
 - Mobilité totale (par ex. une station de ski)
 - Durée non limitée (par ex. une journée)
 - Méthodologies et techniques non validées



10

Tentative de taxonomie

- En laboratoire d'utilisabilité**
 - Le contexte est simulé
 - La tâche est prescrite
 - Le dispositif est perçu comme une maquette
 - Les dispositifs et personnels d'observation sont visibles
- Sur le terrain (a minima)**
 - Le contexte est réel
 - Le dispositif et les données sont perçus comme réels
- En situation réelle**
 - La tâche n'est pas prescrite
 - Les dispositifs et personnels d'observation sont invisibles



11

Bilan et problématique

- Bilan**
 - Il y a trois approches possibles : laboratoire, terrain et situation réelle
 - Les approches terrain et (par extension) situation réelle sont connues pour être plus complexes à mettre en œuvre
- Problématique**
 - Quelle est la « valeur ajoutée » de ces deux approches ?
 - Niveau quantitatif
 - Niveau qualitatif

12

Principe d'incertitude

- Importance des dispositifs et personnels liés à l'observation
 - Leur **présence** apporte un biais, mais permet aussi d'obtenir des informations riches pour analyser l'interaction
 - Leur **absence** élimine le biais, mais ne permet pas a priori d'obtenir assez d'informations pour analyser l'interaction

Principe d'incertitude : « *Il n'est pas possible à la fois d'observer précisément une situation d'interaction homme-machine, sans, par effets de bord, la perturber* »

13

Démarche

- Résultats de la littérature
 - Les approches laboratoire et terrain ont déjà été largement comparées
- Trois (méta-)expérimentations aux objectifs complémentaires
 - **MapMobile** : déterminer si les données disponibles en situation réelle sont suffisantes pour détecter les principaux problèmes d'utilisabilité
 - **E-skiing et Muséum** : déterminer la faisabilité (méthodologique et technique) de véritables expérimentations en situation réelle

14

MapMobile

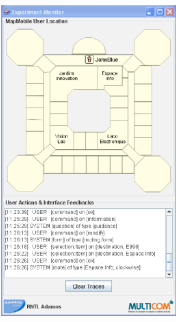
- Dispositif mobile, proactif et sensible au contexte
 - Géo-localisation
 - Guidage proactif
 - Messagerie géo-localisée
- Conditions expérimentales
 - Scénario : visite professionnelle dans une entreprise inconnue
 - Bâtiment de France Télécom R&D Meylan pendant les heures de bureau
 - 12 sujets répartis en âge, genre, expérience et profil sociologique
- Objectifs (qualitatifs)
 - Ergonomie du dispositif
 - Acceptabilité de la proactivité



15

MapMobile – Observations

- Méta-évaluation
 - Une expérimentation
 - Deux ensembles d'observations
 - **Terrain** : commentaires du sujet, vidéo du contexte, film d'écran du dispositif
 - **Situation réelle simulée** : actions utilisateur, retours d'informations du système, position
 - Technique d'observation
 - Temps-réel et à distance via un bus de données spécialisé (sauf pour la caméra contexte)



16



RNTL Adamos
Mapmobile Experiments - June 2005

MapMobile – Analyses

- Analyse Terrain
 - Entretiens semi-structurés
- Analyse Situation réelle simulée
 - Détection de variations par rapport à une référence (variations temporelle et de séquences d'activités)
 - Détection de patterns d'usage
- Méta-analyse
 - Comparaison des résultats obtenus en aveugle (la référence est l'analyse terrain)

18

MapMobile – Résultats

- **Détections conjointes**
 - Principaux problèmes ergonomiques
 - Acceptation/refus de la proactivité
- **Faux négatifs**
 - Biais induits par la présence du facilitateur
 - Confusion exploration/incompréhension
 - Problèmes indétectables : ceux ne conduisant pas à une modification de l'activité perceptible (par le dispositif)
- **Faux positifs**
 - Événements externes (par ex. les amis)

19

Bilan intermédiaire

- **Ce que nous apprend l'expérimentation MapMobile**
 - Les données « terrain » peuvent suffire pour la détection des principaux problèmes ergonomiques
 - *Mais...*
 - Ce n'est pas exhaustif
 - Certains problèmes sont mal interprétés
 - D'autres outils d'observations, notamment la vidéo seraient très utiles...

20

Technique du Cheval de Troie


- **Objectif**
 - Contourner les limites de l'observation liées au principe d'incertitude
- **Principe : le double usage**
 - Dissimuler un dispositif d'observation sous un autre usage, visible, mais crédible
- **Exemple**
 - E-Skiing : mini-caméra fixées sur le casque des skieurs
 - Présentées comme une fonctionnalité permettant de revivre la descente
 - Permettent d'observer l'interaction des skieurs avec le smartphone



21

Technique du Cheval de Troie

- **Généralisation : deux cas possibles**
 - Si il existe dans l'environnement un dispositif ayant potentiellement les capacités d'acquisition : le compléter (*par exemple avec un système d'enregistrement*)
 - Si il n'en existe pas : ajouter le dispositif d'acquisition souhaité et lui trouver un scénario d'usage crédible (*cela ne sera pas possible si le dispositif d'acquisition semble incongru ou inadmissible dans le contexte*)



22

Technique du Cheval de Troie


- **Aspect déontologique**
 - La technique du Cheval de Troie impose de masquer (temporairement) à l'utilisateur l'usage réel des données enregistrées
 - L'usage des données personnelles doit être rigoureusement encadré
 - Déclaration CNIL
 - Anonymisation ou possibilité de destruction des données personnelles
 - Les sujets doivent être informés de l'usage réel des données (au moins a posteriori)



23

E-Skiing

- **Système proactif permettant « d'enrichir » la pratique du ski**
 - Connaître ses performances (vitesse maximum, distance, temps, etc.)
 - Revivre ses descentes (vidéo, chemin parcouru)
- **Conditions expérimentales**
 - Un groupe d'utilisateurs : 4 personnes ayant l'habitude de skier ensemble
 - Une journée de ski + une soirée informelle de type « focus-group »
- **Objectifs (qualitatifs)**
 - Utilisabilité du dispositif
 - Acceptabilité du système (notamment l'aspect proactif)

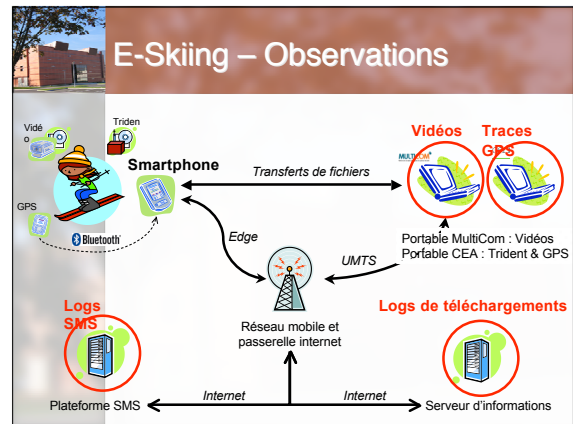



24

E-Skiing – Observations

- Architecture technique**
 - Skiers : accéléromètre (« Trident »), GPS, mini-caméra et smartphone
 - Chalet : deux ordinateurs portables pour le traitement des données
 - Serveurs : serveur de données et plateforme SMS de l'opérateur téléphonique
 - Réseau : réseaux mobile et passerelle internet de l'opérateur téléphonique (et transfert direct via carte mémoire pour les vidéos)
- Acquisition des données (en différé)**
 - Données enregistrées sur les ordinateurs portables : vidéos et traces GPS
 - Logs des serveurs : envois de SMS et lecture des performances

25



E-Skiing – Observations

- Technique du Cheval de Troie**
 - La mini-caméra est présentée comme une des fonctionnalités du service, mais sert principalement à l'observation des interactions du skieur avec le smartphone
- Limites techniques et logistiques**
 - Malgré une forte anticipation (deux tests partiels et un pré-test en situation), de nombreux problèmes techniques et logistiques se sont produits
 - Arrêts inopinés de certains enregistreurs
 - Mauvais cadrage des mini-caméras
 - Saturation lors du traitement des données
 - ...

27



E-Skiing – Analyse

- Reconstruction de l'activité**
 - L'activité des skieurs a été reconstruite sans aide de la vidéo, grâce aux traces GPS et aux logs des serveurs de données et SMS

Date	Commande	Paramètre	Value
9:32	SMS	Mobile 1	
10:46	get	Vitesse_1	45.411
10:46	set	Temps_1	590.294
10:46	set	Distance_1	944
10:46	set	IndiceLat_1	
10:46	page	Vitesse_1	
10:46	get	Temps_1	
10:46	get	Distance_1	
10:46	get	IndiceLat_1	
10:56	page	Vitesse_1	
10:57	get	Temps_1	
10:57	get	Distance_1	
10:57	page	Vitesse_1	
11:20	get	Temps_1	
11:20	get	Distance_1	
11:20	get	IndiceLat_1	
11:20	SMS	Mobile 1	
12:40	SMS	Mobile 3	
12:40	page	Vitesse_1	
12:56	get	Temps_1	
12:56	get	Distance_1	
12:56	get	IndiceLat_1	
12:57	SMS	Mobile 1	

30

E-Skiing – Analyse

- **Résultats**
 - Peu de résultats concernant **l'utilisabilité** (notamment en raison des problèmes d'enregistrement vidéo)
 - Un **usage** complètement inattendu du système : une utilisation exclusive lors des pauses, au bar !
- **Retours d'expérience**
 - Les problèmes techniques et logistiques sont critiques et peuvent difficilement être anticipés...

31

Bilan intermédiaire

- **Ce que nous apprend l'expérimentation E-Skiing**
 - La mise en œuvre d'expérimentations en situation réelle est possible
 - Mais...*
 - Les problèmes techniques et logistiques sont critiques et peuvent difficilement être anticipés
 - Il est possible de reconstruire l'activité des utilisateurs avec des enregistrements vidéo partiels
 - Mais...*
 - L'analyse de **l'utilisabilité** est difficile
 - On obtient cependant de très intéressantes informations concernant **l'usage**

32

Muséum de Lyon

- **Visite « augmentée »**
 - Le visiteur vit l'expérience d'être traqué par des paparazzi virtuels
- **Conditions expérimentales**
 - Le parcours des visiteurs (environ 500) est enregistré, ainsi que leur utilisation des bornes interactives
- **Objectifs (quantitatifs)**
 - Analyse des parcours
 - Calcul des temps de visite



33

Muséum – Observations

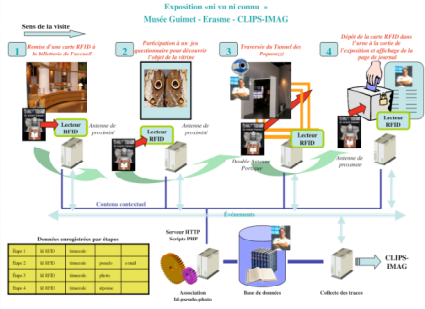
- **Technique du Cheval de Troie**
 - Utilisation d'étiquettes RFID cachées dans une carte distribuée à chaque visiteur avec son ticket d'entrée
 - Détection et identification unique des visiteurs aux points de passage et lors de l'utilisation des bornes interactives
- **Acquisition des données (en différé)**
 - Centralisation des données sur un serveur local avec téléchargement journalier au laboratoire



34

Muséum – Observations

Exposition-salon vu et revu - Musée Galimé - Erasme - CLIPS-IMAG



Principales caractéristiques par étage			
Étage 1	RFID	Interactif	
Étage 2	RFID	Interactif	Interactif
Étage 3	RFID	Interactif	Interactif
Étage 4	RFID	Interactif	Interactif

35

Muséum – Observations

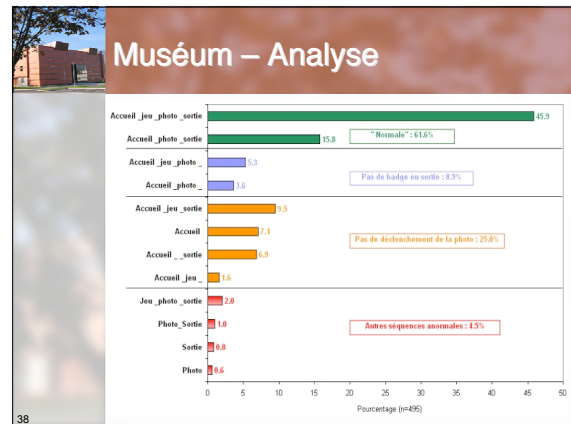
- **Limites techniques et logistiques**
 - Important retard à la mise en place (trois semaines d'expérimentation au lieu des trois mois prévus)
 - Et ce fut quasiment le seul problème...

36

Muséum – Analyse

- **Reconstruction de l'activité**
 - Filtrage des données
 - Non respect des consignes lors de la distribution et la récupération des cartes (présence de doublons)
 - Interprétation des données
 - Comportement aberrants difficilement compréhensibles sans une observation directe (utilisation des témoignages des personnels du musée)
 - Analyse des données
 - Types de parcours
 - Temps de visite

37



Bilan intermédiaire

- **Ce que nous apprend l'expérimentation au Muséum de Lyon**
 - La mise en œuvre d'expérimentations en situation réelle est possible
 - Mais...*
 - Les problèmes techniques et logistiques sont toujours aussi critiques
 - La reconstruction de l'activité demande toujours des informations complémentaires
 - La technique du Cheval de Troie est généralisable

39

Synthèse

- **Pas trois, mais 2 (+1) configurations**
 - Le **laboratoire** et le **terrain** sont finalement peu différents du fait que le protocole est semblable : l'utilisateur se trouve dans une « bulle » l'isolant partiellement du monde extérieur
 - Le **terrain** n'apporte en plus que les éléments d'utilisabilité lié au contexte, ou une plus grande facilité pour les environnements sont peu facilement simulables mais aussi de nouvelles variables indépendantes...
 - La **situation réelle** se distingue principalement par l'absence de tâche prescrite ce qui change de manière importante le protocole expérimental
- **Utilisabilité versus Usage**
 - Le **laboratoire** est adapté à l'utilisabilité et la **situation réelle** à l'usage

40

Synthèse

- **Des approches complémentaires**
 - **Laboratoire** ou **terrain** : utilisabilité
 - **Situation réelle** : usage + utilisabilité filtrée par le contexte d'usage

41

Méthodologie proposée

- **Des approches complémentaires**
 - Le **laboratoire** permet de s'intéresser à l'utilisabilité dans un cadre méthodologique strict
 - Le **terrain** permet de valider le système dans un contexte un peu plus réaliste ou d'éviter une simulation délicate
- La **situation réelle** permet de valider (ou découvrir) les usages et de détecter des problèmes d'utilisabilité liés au contexte d'usage

42


Conclusion

- **Distinction**
 - Laboratoire
 - (Terrain)
 - Situation réelle
- **Principe et technique**
 - Principe d'incertitude
 - Technique du Cheval de Troie
- **Méthodologie proposée**
 - Utiliser de manière complémentaire le laboratoire et la situation réelle
- **Quelques retours d'expérience pour les expérimentations en situation réelle**
 - Ne pas hésiter à prendre du retard
 - Tester, tester, tester, tester, tester, tester...

43

Perspectives

- **Analyse de l'activité humaine**
 - Reconstruction des activités à partir de données hétérogènes et/ou incomplètes
 - Automatiser la reconstruction en se basant sur une modélisation de l'activité humaine
- **Domaines d'application**
 - **Projet TRACES** (MSTIC Université J. Fourier) : Analyse des modifications des déplacements des individus en cas de risques de crues rapides



44

Perspectives

- **Domaines d'application**
 - **Projet TELEOS** (ANR Blanc) : Analyse de l'activité de chirurgiens orthopédiques experts et novices (situations réelles et simulées)



45

Financement et remerciements

- **Projets supports aux expérimentations**
 - **RNTL Franco-Finlandais Adamos** (Adaptive Mobile Services)
Partenaires : University of Oulu, VTT Electronics, CLIPS-IMAG, CEA-Leti, France Telecom, MSH-Alpes, ST Microelectronics
 - **Région Rhône-Alpes IMERA** (Interfaces Mobiles pour l' Environnement Réel Augmenté)
Partenaires : ICTT Lyon, CLIPS-IMAG, MSH-Alpes, Assetium, Tagproduct, France Telecom
- **Remerciements**
 - Les expérimentations ont fait appel aux multiples compétences de toute l'équipe MultiCom !

46

Bibliographie

Jambon, F. **User Evaluation of Mobile Devices: In-Situ versus Laboratory Experiments.** *International Journal of Mobile Computer-Human Interaction*, 1, 2 (2009), pp. 56-71.

Jambon, F., Mandran, N., Meillon, B. et Perrot, C. **Évaluation des systèmes mobiles et ubiquitaires : proposition de méthodologie et retours d'expérience.** *Ergo-IA'08 "L'humain au coeur des systèmes et de leur développement"* (15-16-17 Octobre, Biarritz, France), 2008, pp. 107-116

Jambon, F., Golanski, C. et Pommier, P.-J. **Meta-Evaluation of a Context-Aware Mobile Device Usability.** *International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM'07)* (November 4-9, IEEE Computer Society Press, 2007, pp. 21-26.

Jambon, F., Mandran, N. et Perrot, C. **La RFID au service de l'analyse du parcours muséal des visiteurs.** *La lettre de l'OCIM (Office de Coopération et d'Information Muséographiques)*, 113 (2007), pp. 11-17.

Jambon, F. **Retours d'expérience lors d'évaluation de dispositifs mobiles en situation réelle.** *3e Journées Francophones Mobilité et Ubiquité - Ubimob'06* (5-8 septembre, Paris, France), ACM Press, 2006, pp. 135-138.

Jambon, F., Golanski, C. et Pommier, P.-J. **Évaluation des dispositifs mobiles : sur le terrain ou en laboratoire ?** *18e Conférence Francophone sur l'Interaction Homme-Machine - IHM'06* (18-21 Avril, Montréal, Canada), ACM Press, 2006, pp. 35-42.

47

Évaluation des systèmes mobiles et ubiquitaires : proposition de méthodologie et retours d'expérience

Francis JAMBON, LIG/MultiCom, Grenoble



48