



HAL
open science

Une alternative au lecteur d'écran : l'agent interactionnel dialoguant (AID)

Gérard Uzan, Johana Bodard, Laetitia Gros, Valérie Botherel, Denis Chêne

► To cite this version:

Gérard Uzan, Johana Bodard, Laetitia Gros, Valérie Botherel, Denis Chêne. Une alternative au lecteur d'écran : l'agent interactionnel dialoguant (AID). 2025. <hal-05301860>

HAL Id: hal-05301860

<https://hal.science/hal-05301860v1>

Preprint submitted on 24 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Une alternative au lecteur d'écran : l'agent interactionnel dialoguant (AID)

An Alternative to the Screen Reader: the Interactional Dialog Agent (IDA)

Gérard Uzan
gerard.uzan@univ-paris8.fr
Laboratoire CHArt, Université Paris 8
Saint-Denis, France

Johana Bodard
johana.bodard02@univ-paris8.fr
Laboratoire CHArt, Université Paris 8
Saint-Denis, France

Laetitia Gros
Orange Labs
Lannion, France
laetitia.gros@orange.com

Valérie Botherel
Orange Labs
Lannion, France
valerie.botherel@orange.com

Denis Chêne
Orange Labs
Grenoble, France
denis.chene@orange.com

Résumé

Ce travail renouvelle l'accessibilité numérique pour les personnes aveugles, en s'affranchissant du paradigme du lecteur d'écran. Plutôt que de restituer linéairement, séquentiellement et exhaustivement ce qui est affiché dans un écran ou dans ses fenêtres, et de maintenir l'écran comme interface de commande, nous considérons les documents, sites web et mails comme des espaces d'interaction, porteurs de tâches pouvant être réalisées par un agent virtuel, assistant de l'utilisateur. La problématique n'est plus l'accessibilité (accès à), mais l'utilisabilité (utilisation de). Elle ne pointe plus seulement sur les difficultés d'information, mais sur celles de réalisation de tâches, d'atteinte d'un but ou d'obtention d'un résultat. Cela correspond pour nous à la transition des TIC vers ce que nous appelons les TICAPP (Technologies de l'Information, la Communication, l'Action, le Partage et la Production). Pour cela, nous avons développé un modèle fondé sur un agent interactionnel dialoguant (AID) dont suit ici la description. Il s'agit d'un modèle orienté tâches dans lequel l'AID agit sous le contrôle ou par délégation de l'utilisateur, tenant compte des profils, situations et relations.

Abstract

This work renews digital accessibility for blind people, by moving beyond the screen reader paradigm. Rather than linearly, sequentially, and exhaustively rendering what is displayed on a screen or on its windows, and maintaining the screen as the command interface, we consider documents, websites, and emails as interactive spaces that contain tasks which can be performed by a virtual agent, assisting the user. The problematic is no longer accessibility (access to), but usability (use of). It no longer focuses only on information difficulties, but also on difficulties related to performing tasks, achieving a goal, or obtaining a result. For us, it corresponds to the transition from ICT (Information and Communication Technologies) to what we call ICASPT (Information, Communication, Action, Sharing, and Production Technologies). To this end, we have developed a model based on an interactional dialog agent

(IDA), which is described below. It is a task-oriented model in which the IDA acts under the control of, or by delegation from the user, taking into account profiles, situations, and relations.

CCS Concepts

• **Human-centered computing** → **Accessibility theory, concepts and paradigms**; *Accessibility technologies*; • **Computing methodologies** → **Intelligent agents**.

Mots clés

agent dialoguant, agent interactionnel, rencontre, modèle par tâches, personnes aveugles, TICAPP

Keywords

dialog agent, interactional agent, encounter, task-oriented model, blind people, ICASPT

Reference Format:

Gérard Uzan, Johana Bodard, Laetitia Gros, Valérie Botherel, and Denis Chêne. 2025. Une alternative au lecteur d'écran : l'agent interactionnel dialoguant (AID). In *IHM'25 : Actes étendus de la 36^e conférence Francophone sur l'Interaction Humain-Machine, Novembre 4-7, 2025, Toulouse, France*. 5 pages.

1 Introduction

Aujourd'hui, les contenus et documents numériques ne sont plus de simples reproductions informatiques de documents consultables sur support papier. Ils sont devenus des espaces d'interaction, permettant de réaliser des tâches plus larges que les seules tâches de lecture et d'écriture. Sous l'angle du processus de concrétisation au sens de Simondon [13], les TIC (Technologies de l'Information et de la Communication) sont désormais des TICAPP : Technologies de l'Information, de la Communication, de l'Action, du Partage et de la Production.

Les personnes aveugles peuvent accéder aux documents numériques et à leurs contenus grâce à des outils informatiques de compensation tels que les lecteurs d'écran, les synthèses vocales et les plages braille. Cependant, malgré l'usage de ces technologies, l'accès aux documents numériques reste difficile et limité pour

ces personnes. Selon l' tude Hom re, qui visait    tudier la vie quotidienne des personnes avec une d ficiency visuelle, la facilit  d'utilisation d'Internet varie en fonction de la t che r alis e [11]. Seule la t che d' change de mails est jug e « tr s facile, m me seul » par plus de la moiti  des r pondants (62 %). Pour toutes les autres t ches (consultation d'information, cours ou r unions, r seaux sociaux, achats en ligne et d marches administratives), un r pondant sur dix d clare renoncer   faire la t che   cause de la d ficiency visuelle, et les difficult s restent nombreuses pour les autres, avec un recours important   l'aide humaine. Par ailleurs, une enq te men e par Access42 et la F d ration des Aveugles et Amblyopes de France a soulign  les difficult s rencontr es par les utilisateurs de lecteurs d' cran [5]. En particulier, 71 % des r pondants d clarent rencontrer des difficult s pour utiliser les composants interactifs des sites web avec leur lecteur d' cran, et seulement un quart d'entre eux disent pouvoir acc der sans probl me aux documents PDF. Par ailleurs, l'usage d'un lecteur d' cran est associ    une augmentation significative du temps n cessaire   la r alisation de t ches sur le web par rapport   des utilisateurs voyants [6, 16]. L'usage d'un lecteur d' cran entra ne  galement un accroissement de la charge mentale, une incertitude interactionnelle, ainsi que de fortes contraintes temporelles [7, 16].

Dans cet article, nous proposons une autre approche, dans laquelle nous cherchons   nous extraire du paradigme du lecteur d' cran. Ce paradigme peut  tre r sum  ainsi :

- La personne aveugle ne peut pas acc der   ce qui est affich    l' cran ;
- Il faudrait donc restituer ce qui est affich    l' cran par une autre modalit  que la vue ;
- Les lecteurs d' cran permettent de restituer, de mani re exhaustive, les contenus affich s   l' cran ;
- D s lors, le probl me d'accessibilit  serait r gl .

Or, cette approche s'av re insuffisante. Il ne suffit pas de restituer par une autre modalit  ce qui est affich    l' cran pour garantir l'accessibilit  des documents num riques. Cependant, la mise en  uvre de l'accessibilit  des documents num riques reste centr e sur cet outil et donc largement orient e sur l'acc s   l'information (le I de TIC), en n gligeant les autres dimensions (les CAPP).

Sans diminuer l'importance du lecteur d' cran, un paradigme orient  sur l'utilisation d'un agent virtuel permettrait de penser autrement les interactions entre les utilisateurs et les contenus auxquels ils sont confront s. Pour une personne aveugle, un contenu n'a pas besoin d' tre affich    l' cran pour  tre utilis , cr e ou modifi . C'est pour le partage avec un environnement qui voit que ce contenu peut  tre affich . Ce qui est partag , ce ne sont donc pas les  crans (barres de menus et d'outils, fen tres, ic nes), mais les contenus (documents, mails, sites et pages web). Pour suppl er   leurs difficult s, les personnes aveugles ont recours   l'assistance humaine. D s lors, pourquoi ne pas envisager l'assistance d'un agent virtuel ? Au lieu, par exemple, de consid rer un site web dans une approche universelle int grale pour l'accessibilit , la prise en compte de la nature de l'activit  du site web (par exemple : site administratif, de commerce ou de vente, de spectacles, etc.), et des sous-activit s induites par cette activit  principale, peut ouvrir vers un mod le de t ches r alisables par l'agent.

2 Travaux connexes

S'appuyant sur les avanc es r centes en intelligence artificielle et en traitement du langage naturel — notamment les agents intelligents, les agents conversationnels et les grands mod les de langue — plusieurs travaux ont propos  des solutions visant   am liorer l'acc s aux documents num riques pour les personnes aveugles ou malvoyantes.

Kodandaram et al. (2024) proposent *Savant*, un syst me exploitant un grand mod le de langue pour fournir un cadre d'interactions uniforme, quel que soit le lecteur d' cran et l'application utilis s [10]. L'utilisateur peut manipuler l'interface d'une application gr ce   des commandes en langage naturel qui automatisent des s quences d'actions des lecteurs d' crans sur les  l ments de contr le de l'interface (boutons, menus, champs de saisie, etc.). Par exemple, la commande *Change font size to 12* permet de modifier la taille de la police de caract res sans avoir   naviguer manuellement dans l'interface avec le lecteur d' cran. Cette approche est int ressante, mais elle reste centr e sur l'utilisation d'un lecteur d' cran et la r alisation de t ches avec l'interface de l'application et non avec le document lui-m me.

D'autres travaux proposent des approches alternatives fond es sur le dialogue en langage naturel et/ou la r alisation de t ches avec l'aide d'agents intelligents. Darvishy et al. (2023) pr sentent un syst me permettant la cr ation et l' dition de documents textuels sur un t l phone mobile en dialoguant avec un agent via des commandes vocales simples [4]. Ce syst me propose deux modes d'interaction : un mode « dict e », dans lequel l'utilisateur dicte le texte    crire, et un mode « commande », dans lequel il peut  diter le document en utilisant des instructions vocales telles que *delete last word* pour supprimer le dernier mot saisi.

Certains travaux s'int ressent plus particuli rement   la navigation web ou   la r alisation de t ches sur le web. Baez et al. (2022) introduisent l'approche *Conversational Web Browsing*, dont l'objectif est de permettre   un utilisateur d ficient visuel de naviguer sur un site web en dialoguant avec un agent conversationnel [1]. L'agent fournit une description succincte du site et de sa structure, interpr te les requ tes de l'utilisateur pour lui pr senter les contenus pertinents, et r alise   sa place les actions de navigation. Cette approche se concentre uniquement sur les t ches de lecture et de navigation au sein des contenus web.

Barbosa et al. (2022) proposent une approche centr e sur le but de l'utilisateur lorsqu'il navigue sur un site web et sur les t ches qu'il doit r aliser pour atteindre ce but [2]. L'objectif est de fournir un acc s rapide aux fonctionnalit s les plus couramment retrouv es sur les sites web (par exemple : s'authentifier, chercher un produit, v rifier son prix, etc.). Ces fonctionnalit s sont accessibles soit en saisissant directement la commande associ e (par exemple, *log in* pour s'authentifier), soit en s lectionnant une commande dans la liste des fonctionnalit s automatiquement d tect es par le syst me. De mani re similaire, Cane (2023) propose  galement une approche orient e sur les t ches [3]. Les sites web sont cat goris s en domaines et, pour chaque domaine, un ensemble de commandes correspondant   des t ches pr cises est d fini. L'utilisateur saisit des commandes simples, telles que *search eggs* ou *checkout*, pour r aliser les t ches associ es.

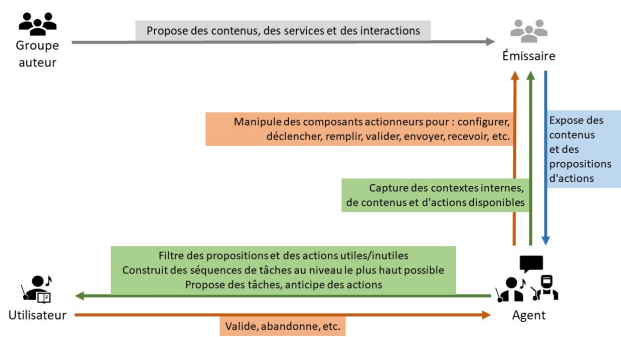


FIGURE 1 : Schéma du modèle

3 Notre modèle

Le modèle que nous présentons dans cet article est un modèle d'interaction entre un utilisateur — potentiellement aveugle — et un document numérique (par exemple : un mail, une page web, un livre, un rapport, etc.). Nous avons adopté une approche centrée sur l'utilisateur et sur les tâches qu'il peut potentiellement réaliser à partir des documents numériques et de leurs contenus.

3.1 Les différents acteurs du modèle

Dans notre modèle, nous distinguons quatre catégories d'acteurs : le groupe auteur, l'émissaire, l'utilisateur et l'agent. La figure 1 présente un schéma général qui positionne ces quatre acteurs dans le modèle.

Le **groupe auteur** désigne l'ensemble des concepteurs, réalisateurs et diffuseurs des documents et de leurs contenus. Il peut s'agir, par exemple, des signataires et expéditeurs d'un mail, de la maîtrise d'œuvre d'un site web, des auteurs d'un rapport, ou encore de la chaîne d'édition et/ou de diffusion d'un document.

L'**émissaire** désigne tout document numérique avec ses contenus. Nous le nommons émissaire car il est une forme déléguée du groupe auteur auprès de l'utilisateur avec une ou plusieurs missions (par exemple : informer, commercer ou divertir). Il représente le groupe auteur de manière altérée, asynchrone et ubiquitaire.

L'**utilisateur** est la personne qui est confrontée au contenu. Il a un but à atteindre ou un résultat à obtenir.

L'**agent** agit pour le compte de l'utilisateur afin de satisfaire le but qu'il veut atteindre ou le résultat qu'il veut obtenir. Il peut être :

- collaboratif : il partage le même objectif que l'utilisateur ;
- coopératif : il fait avec l'utilisateur ;
- déléгатif : il fait à la place de l'utilisateur.

Il comporte une part conversationnelle :

- il recueille la demande de l'utilisateur et la fait préciser si nécessaire ;
- il accompagne l'utilisateur en décryptant l'émissaire.

Il a une mémoire de (et dans) la situation immédiate, des rencontres antérieures entre l'utilisateur et les émissaires, ainsi que des profils personnels de l'utilisateur et des émissaires.

Sans agent, l'utilisateur est confronté à un émissaire passif. Il doit découvrir et comprendre l'émissaire, détecter les tâches disponibles, puis sélectionner et réaliser celles qui lui permettront

d'atteindre son but ou d'obtenir le résultat souhaité. L'interaction est à l'initiative de l'utilisateur.

Parmi les TICAPP, les agents conversationnels actuels tentent de réaliser au mieux le « I » (Information), avec un « C » (Communication) marqué par de l'empathie et de la bienveillance — souvent au prix d'hallucinations. Notre modèle d'agent, que nous désignons par **Agent Interactionnel Dialoguant (AID)**, tente d'explorer la problématique de la partie APP (Action, Partage, Production) des TICAPP. L'AID a des capacités de prévention, d'initiative et de délégation, dans le respect de l'autodétermination de l'utilisateur. Il agit comme un assistant disposant de mémoires des situations, des relations, des buts, des résultats, des caractéristiques et des comportements de l'utilisateur et des émissaires.

3.2 Le modèle de tâches

Nous distinguons quatre niveaux de tâches : la macro-tâche, les mezzotâches, les mini-tâches et les micro-tâches.

La **macro-tâche** est la tâche globale d'intérêt, c'est-à-dire l'activité englobante répondant à une finalité permise et/ou suggérée par l'émissaire. Par exemple : acheter un ensemble de produits sur un site d'e-commerce, se connecter à une réunion en visio-conférence, écouter une émission de radio sur Internet, etc. La macro-tâche est composée d'une ou plusieurs mezzotâches.

Les **mezzotâches** sont l'ensemble des tâches de niveau intermédiaire, organisées en séquences à l'intérieur de la macro-tâche. Elles produisent des résultats intermédiaires utiles aux mezzotâches qui leur succèdent. Un but atteint ou un résultat obtenu clôt la tâche. La nature des mezzotâches dépend du type de contenu (mails, sites et pages web, documents numériques sous forme de fichier) et du domaine (site d'e-commerce, site culturel, document administratif, etc.). Par exemple, dans le cadre d'un site de théâtre, un utilisateur qui souhaite obtenir des places pour une séance précise d'un spectacle (macro-tâche) devra réaliser les mezzotâches suivantes : s'informer, réserver, payer, valider. En explorant différents sites web, correspondant à des propositions de macro-tâches différentes, on peut dégager un ensemble commun de mezzotâches, par exemple : s'informer, s'authentifier, contacter, réserver, etc.

Les **mini-tâches** sont l'ensemble des tâches permises par la présence d'éléments de structure ou de contenus lexico-sémantiques récupérables et porteurs de tâches. Elles incluent les tâches explicitement proposées par le groupe auteur, que nous appelons **micro-tâches** (exploitables sous une autre forme que celle proposée par le groupe auteur). Par exemple, un numéro de téléphone écrit sous forme textuelle brute peut être identifié et récupéré par son contexte et sa structure, et est alors est porteur de tâches propoables à l'utilisateur : appeler le numéro, insérer le numéro dans un carnet de contacts, etc.

Les **micro-tâches** sont l'ensemble des tâches explicitement proposées par le groupe auteur à travers des interacteurs passifs (afficheurs et capteurs, par exemple : texte, image, compteur, jauge) ou activables (actionneurs, par exemple : champ de saisie, bouton, lecteur).

Sans agent, l'utilisateur doit créer, ordonnancer et maintenir dans sa propre mémoire de travail les mezzotâches à l'intérieur de la macro-tâche. Il doit également ordonnancer les micro-tâches dans les mezzotâches. En dehors des micro-tâches, les mini-tâches

TABLE 1 : Exemple d'exploitation des traces

Contexte	Il est 16 h. (<i>trace de situation</i>) Depuis 60 jours, l'utilisateur �coute quotidiennement le Journal de 18 h de France Culture, sauf deux fois o� il a �coute celui de France Inter. (<i>trace de relation</i>)
Sc�nario	U : Mets le journal de 18h. A : C'est trop t�t. Je te pr�viens vers 18h? U : Oui. A : Tr�s bien. � tout � l'heure. ...
� 17 :59	A : Je mets le journal? U : Oui.
� 18 :00	L'agent lance la diffusion du Journal de 18h de France Culture.

U : utilisateur, A : agent interactionnel dialoguant

ne sont disponibles qu'avec la pr sence de l'agent. Accompagnant l'atteinte d'un but ou l'obtention d'un r sultat, l'agent  labore, avec et pour l'utilisateur, les s quences de t ches et les interactions n cessaires   leur r alisation. Celles-ci sont compos es de tours de paroles, d'actions et de s quences. Ainsi, le bin me agent-utilisateur entre dans une s rie de **rencontres** successives avec chacun des  missaires.

3.3 La rencontre entre un utilisateur et un  missaire

La visite de l'utilisateur sur un mail, une page web ou un document est d finie comme une **rencontre** entre un utilisateur et un  missaire. Cette rencontre se compose d'interactions et/ou de dialogues. La s quence d'interactions peut  tre enrichie par la prise en compte du **contexte** de la rencontre. Ce contexte comprend   la fois le contexte li    la situation imm diate [15, 17] et celui li    la relation entre l' missaire et l'utilisateur [12, 14].

Le contexte de **situation** porte sur :

- le(s) mat riel(s) et le(s) logiciel(s) utilis s pour acc der   l' missaire ;
- l'environnement temporel de la rencontre : date et heure ;
- l'environnement g o-spatial et organisationnel de la rencontre : l'utilisateur est chez lui,   son travail, dans un lieu public, etc.

Le contexte de **relation** comprend :

- les donn es concernant les rencontres ant rieures, qui peuvent  tre associ es   des informations permettant des calculs de fr quence, de r cence, de dur e, ou de saisonnalit  (par exemple : produits r guli rement achet s ou listes de parcours SNCF fr quemment consult s) ;
- les donn es relatives aux transformations de l' missaire ou de l'utilisateur survenues entre la rencontre pr c dente et la rencontre actuelle (activit s supprim es ou ajout es, obsolescence des donn es ant rieures) ;
- les donn es partag es (par exemple, bon de commande ou suivi d'un colis),
- les donn es li es   la qualit  de la relation entre le groupe auteur et l'utilisateur (relation professionnelle, relation familiale, etc.),
- les donn es de profil de l'utilisateur et de l' missaire.

De rencontre en rencontre, ces donn es constituent des **traces** (de situation, de relation et de profil). La m morisation et l'exploitation de ces traces visent plusieurs objectifs :

- ne pas repartir *ex nihilo*   chaque nouvelle rencontre ;
- optimiser la rencontre conversationnelle selon les maximes de Grice [8] : sobri t  et minimisation verbales des empanns et du nombre de tours de parole ;
- optimiser la rencontre interactionnelle :
 - r duire le nombre d'enchaînements en substituant des s quences de mini/micro-t ches par une mezzo-t che effectu e,
 - r duire le nombre d'enchaînements en r duisant les  tapes de choix, d'insertion de donn es et de d clenchement en les d l guant   l'agent gr ce   la m moire partag e,
 - maintenir le contr le de l'utilisateur en lui gardant la supervision et la capacit  d'interception dans les t ches d l gu es,
 - optimiser l'imbrication entre dialogue et actions en r duisant ou supprimant des tours de paroles et accroissant la d l gation ;
 - profiler des mezzo-t ches communes instanciables.

Le tableau 1 pr sente un exemple d'exploitation des traces.

4 Discussion

Par rapport aux travaux pr sent s dans la section 2, notre mod le, int gratif, relie le contexte (situation, relation, profil) et l'interaction. Ce lien est chapeaut  par l'objectif de l'utilisateur. Comme le fait Cane (2023) concernant les pages web [3], le profil de l' missaire d termine des t ches disponibles et des r sultats potentiels atteignables. Au-del  de la page web, d'autres cat gories d' missaires sont envisageables : mail, rapport, devis, facture, etc.

Dans notre mod le, les interactions s'articulent autour des t ches disponibles. Nous distinguons sept cat gories d'actions primaires dans l'interaction, implicitement ou explicitement r alis es, par commande ou d l gation, isol es ou incorpor es dans une s quence :

- le pointage ;
- la commande ;
- la dict e ;
- le choix (ou dialogue) ;

- la requête informationnelle ;
- la substitution (faire autrement) ;
- la bifurcation (changer de but).

Chacun des travaux cités dans la section 2 se concentre sur une partie de ces actions. Kodandaram et al. (2024) [10] et Darvishy et al. (2023) [4] proposent une approche de substitution modale du tactilo-haptique vers la reconnaissance vocale, les interfaces restant égales par ailleurs. Kodandaram et al. explorent les commandes, plus précisément les commandes d'application, en conservant le lecteur d'écran. Darvishy et al. (2023), quant à eux, explorent également le pointage et la dictée. Baez et al. (2022) s'écartent du lecteur d'écran et s'intéressent à l'action de choix [1]. Barbosa et al. (2022) se focalisent sur les commandes incorporées dans les pages web, en proposant une re-présentation de la page qui permet de proposer des micro-tâches et des mezzo-tâches. Cependant, ils n'osent pas franchir le pas de s'affranchir du lecteur d'écran [2]. Enfin, Cane (2023) présente une approche qui couvre plusieurs catégories d'actions (pointage, commande, choix, substitution) et qui s'affranchit potentiellement du lecteur d'écran. C'est l'approche qui a la plus grande proximité avec la nôtre.

5 Perspectives et travaux futurs

Nous envisageons d'appliquer le modèle à trois catégories d'émissaire : le mail, la page web et le document rédigé, partageable, affichable et/ou imprimable (par exemple : livre, rapport, revue, article, etc.).

La première étape consiste à caractériser chacune de ces catégories. À partir de contenus, nous définirons des propriétés utiles pour l'interaction, permettant de décrire la situation, la relation entre le groupe auteur et l'utilisateur, ainsi que les profils des émissaires.

Prenons l'exemple d'un mail de demande envoyé à l'utilisateur par Pierre Dupont. Des descripteurs possibles seraient :

- pour le profil : adresse mail de l'expéditeur et des personnes en copie, signature(s), données de contenus (objet et corps du mail), etc.
- pour la situation : ton du mail (professionnel, amical, familial, etc.), date et heure d'envoi, date et heure de consultation, flux de mails faible ou important, etc.
- pour la relation :
 - premier contact de l'expéditeur/contacts fréquents/...,
 - organisme/activité/produit/actualité/...,
 - demande unique/demande faite par d'autres personnes/...

Certaines propriétés peuvent être communes à deux ou trois catégories de descripteurs (profil, situation, relation).

Par la suite, nous définirons un ensemble de buts à atteindre et de résultats à obtenir, et, de manière associée, les macro-tâches et les mezzo-tâches correspondantes. Ce travail nécessite une analyse approfondie des contenus afin d'identifier et de classer les macro-tâches et mezzo-tâches communes. Nous recherchons une classification la plus exhaustive possible (pour autant restant ouverte).

Puis, nous développerons un prototype implémentant notre modèle sur la base d'un prototype existant fondé sur un modèle d'agent proposant [9, 16]. Enfin, une évaluation du prototype sera menée auprès de personnes aveugles, afin de vérifier la pertinence du modèle proposé.

Références

- [1] Marcos Baez, Claudia Maria Cutrupi, Maristella Matera, Isabella Possaghi, Emanuela Pucci, Gianluca Spadone, Cinzia Cappiello, and Antonella Pasquale. 2022. Exploring challenges for Conversational Web Browsing with Blind and Visually Impaired Users. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3491101.3519832
- [2] Natã M. Barbosa, Jordan Hayes, Smirity Kaushik, and Yang Wang. 2022. "Every Website Is a Puzzle!" : Facilitating Access to Common Website Features for People with Visual Impairments. *ACM Trans. Access. Comput.* 15, 3 (July 2022), 19 :1–19 :35. doi:10.1145/3519032
- [3] David Cane. 2023. Task Aware Browsing : AI Assisted Web Access for Blind Users. *Journal on Technology and Persons with Disabilities* (2023), 12–27. <http://hdl.handle.net/10211.3/225162>
- [4] Alireza Darvishy, Hans-Peter Hutter, Edin Beljulji, and Zeno Heeb. 2023. A new conversational interaction concept for document creation and editing on mobile devices for visually impaired users. doi:10.48550/arXiv.2301.02546 arXiv :2301.02546 [cs].
- [5] Sylvie Duchateau and Boulay Denis. 2018. *Étude sur l'usage des lecteurs d'écran et des outils et logiciels "basse vision" en France et en Francophonie*. Technical Report. 53 pages. https://www.firah.org/upload/notices/3/2018/rapport_etude_le_version-finale_21mars2018.pdf
- [6] Stéphanie Giraud. 2014. *L'accessibilité des interfaces informatiques riches pour les déficients visuels*. Thèse de doctorat. Université Nice Sophia Antipolis. <https://theses.hal.science/tel-01154444>
- [7] Stéphanie Giraud, Gérard Uzan, and Pierre Thérouanne. 2011. L'accessibilité des interfaces informatiques pour les déficients visuels. In *L'ergonomie des objets et environnements physiques et numériques*, Jérôme Dinet et Christian Bastien (Ed.). Hermes - Sciences Lavoisier, 279–304. <https://hal.science/hal-01731086>
- [8] H. P. Grice. 1975. Logic and Conversation. In *Speech Acts*. Brill, 41–58. doi:10.1163/9789004368811_003 Section : Speech Acts.
- [9] Mohammed Hanine, Susana Herrera, Denis Chêne, and Gérard Uzan. 2021. Agent virtuel pour l'utilisation effective du web par des aveugles. In *Handicap, Vieillesse, Autonomie, Insertion, Technologies*. IFRATH, Université Paris-Est Créteil, 57–63.
- [10] Satwik Ram Kodandaram, Utku Uckun, Xiaojun Bi, IV Ramakrishnan, and Vikas Ashok. 2024. Enabling Uniform Computer Interaction Experience for Blind Users through Large Language Models. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3663548.3675605
- [11] Caroline Pigeon, Anna Rita Galiano, Myriam Evennou, Gérard Uzan, and Nicolas Baltenneck. 2024. Usage d'internet des personnes atteintes d'une déficience visuelle et ses déterminants. Exploitation des données de l'étude Homère. IFRATH, Paris. <https://hal.science/hal-04604666> hal-04604666.
- [12] John R. Searle. 1983. *Intentionality : An Essay in the Philosophy of Mind*. Cambridge University Press.
- [13] Gilbert Simondon. 2012. *Du mode d'existence des objets techniques*. Aubier, Paris.
- [14] Dan Sperber and Deirdre Wilson. 1989. *La Pertinence : Communication et cognition*. Les Éditions de Minuit.
- [15] Lucille Alice Suchman. 1987. *Plans and Situated Actions : The Problem of Human-Machine Communication* (cambridge university press ed.). Cambridge.
- [16] Gérard Uzan. 2005. *Ergonomie cognitive du handicap visuel : une contribution à la conception d'aides informatiques*. Thèse de doctorat. Université René Descartes - Paris 5.
- [17] Julia Velkovska and Marc Relieu. 2021. Pour une conception « située » de l'intelligence artificielle : Des interactions hybrides aux configurations socio-techniques. *Réseaux* 229, 5 (Nov. 2021), 215–229. doi:10.3917/res.229.0215 Publisher : La Découverte Section : Sociologie.