



**HAL**  
open science

# État de l'art des approches de modélisation et de simulation utilisateur pour la recherche d'information conversationnelle

Pierre Erbacher, Laure Soulier

## ► To cite this version:

Pierre Erbacher, Laure Soulier. État de l'art des approches de modélisation et de simulation utilisateur pour la recherche d'information conversationnelle. CORIA 2021, Apr 2021, Grenoble, France. <hal-03480755>

**HAL Id: hal-03480755**

**<https://hal.science/hal-03480755v1>**

Submitted on 6 Jan 2022

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

---

# État de l'art des approches de modélisation et de simulation utilisateur pour la recherche d'information conversationnelle

**Pierre Erbacher — Laure Soulier**

Sorbonne Université, CNRS, LIP6, F-75005 Paris, France  
{pierre.erbacher,laure.soulier}@lip6.fr

---

*RÉSUMÉ.* La recherche d'information conversationnelle (RIC) est un domaine à la croisée de la RI interactive et des systèmes de dialogue pour des besoins en information sur des domaines ouverts. Afin d'optimiser les interactions entre système et utilisateur et améliorer au mieux l'expérience utilisateur, il est nécessaire d'améliorer les modèles d'interactions en RI par la prise en compte séquentielle des actions hétérogènes. L'apprentissage par renforcement s'est imposé comme un paradigme particulièrement adapté pour optimiser les prises de décision séquentielles dans de nombreux domaines et apparaît récemment en RI. Or, entraîner ces systèmes par renforcement directement sur des utilisateurs n'est pas envisageable. Une solution est d'entraîner les systèmes de RI sur des simulations utilisateurs qui modélisent le comportement de vrais utilisateurs. Ce papier a deux objectifs : 1) proposer un état de l'art des modèles de comportement et de simulations utilisateurs pour l'accès à l'information et 2) discuter des différentes perspectives de recherche dans les simulations utilisateur dans le contexte de la RIC.

*ABSTRACT.* Conversational Information Retrieval (CIR) is an emerging field of Information Retrieval (IR) at the intersection of interactive IR and dialogue systems for open domain information needs. In order to optimize these interactions and enhance the user experience, it is necessary to improve IR models by taking into account sequential heterogeneous user-system interactions. Reinforced learning has emerged as a paradigm particularly suited to optimize sequential decision making in many domains and has recently appeared in IR. However, training these systems by reinforcement learning on users is not feasible. One solution is to train IR systems on user simulations that model the behavior of real users. Our contribution is twofold: 1) reviewing the literature on user modeling and user simulation for information access, and 2) discussing the different research perspectives for user simulations in the context of CIR.

*MOTS-CLÉS :* Recherche d'information conversationnelle, simulation utilisateur.

*KEYWORDS:* Conversational information retrieval, user simulation

---

## 1. Introduction

La recherche d'information conversationnelle (RIC) est désormais un enjeu de recherche majeur pour la communauté de Recherche d'Information (RI) (Culpepper *et al.*, 2018). Elle se positionne à la croisée de nombreux domaines d'application (RI interactive, systèmes conversationnels, systèmes de questions-réponses) et s'en distingue en de nombreux aspects. 1) A l'inverse de la RI interactive (Zhai, 2020) qui s'intéresse également aux *feedbacks* des utilisateurs, son objectif est de trouver de manière plus naturelle les informations pertinentes. 2) Elle est également différente des systèmes de conversation orientés tâches (Bordes *et al.*, 2017) qui sont guidés dans un domaine d'application particulier (i.e. réservation de voyage). Un système de RIC est un cas d'application spécifique des systèmes de conversation orientés tâches, mais sont différents dans la mesure où la recherche est effectuée dans un domaine ouvert. 3) Enfin, très proche des systèmes de questions-réponses (Bordes *et al.*, 2014) qui s'expriment en langage naturel, la RIC se concentre sur des besoins en information exploratoires ou complexes, sur des domaines ouverts, pas spécifiquement centrés sur une entité ou une question précise. L'objectif d'un système de RIC est alors de supporter: 1) répondre à un besoin en information complexe exprimé en langage naturel ou au travers de différents *feedbacks* et 2) anticiper et guider l'utilisateur dans sa séquence d'actions de recherche (Azzopardi *et al.*, 2018 ; Radlinski et Craswell, 2017). Cela nécessite aux systèmes de RIC d'interagir avec l'utilisateur pour mieux comprendre et/ou raffiner son besoin en information. Les interactions entre le système et l'utilisateur sont alors séquentielles (e.g., séquences de reformulation de requêtes, de clics sur des documents, de conversations en langage naturel), nécessitant ainsi la prise de décision du système conversationnel sur le long terme.

La plupart des modèles récents se concentrent sur l'aspect RI interactif qui construisent un profil utilisateur ou une modélisation de la session (Sordoni *et al.*, 2015a). Cependant, ils n'embarquent pas de stratégie de décision à long terme, séquentielle et incrémentale. Il est pourtant essentiel et déjà utilisé pour les systèmes de recommandation conversationnel (Sun et Zhang, 2018 ; Jannach *et al.*, 2020). Un intérêt grandissant dans la communauté est de concevoir des systèmes de RI qui embarquent des stratégies pour optimiser au mieux ces séquences de recherche notamment par des interactions en langue naturelle avec les utilisateurs (Zhang *et al.*, 2020). L'apprentissage par renforcement s'est montré particulièrement efficace pour optimiser des prises de décisions séquentielles (Sutton et Barto, 1998) car il apprend à maximiser un signal récompense par interactions successives avec son environnement (Sutton et Barto, 1998). Dans le cadre de la RI, l'idée d'utiliser l'apprentissage par renforcement pour optimiser une politique de RI fait son chemin (Zhang *et al.*, 2020 ; Montazerlghaem *et al.*, 2020). Les politiques apprises optimiseraient sur le long terme les séquences des interactions utilisateur-système de RIC en fonction des différents objectifs de performance ou d'interactions selon différents critères (réduire/augmenter le nombre d'interactions avec l'utilisateur, augmenter la diversité de l'information retrouvée, prédire les futures requêtes, orienter les choix utilisateurs, ...). Idéalement les politiques des systèmes de RIC devraient s'entraîner

sur de vrais utilisateurs. En pratique cela est impossible car les algorithmes d'apprentissage par renforcement sont actuellement très inefficaces : il faudrait des millions d'interactions avec des utilisateurs pour voir émerger une politique correcte. S'entraîner directement sur des données limite l'exploration des scénarios collectés, les systèmes de RIC ne peuvent donc visiter qu'une toute petite partie de l'espace des états. De plus, même si la base de données est très large, il y a peu de chance que les trajectoires optimales soient contenues dedans (Kreyssig *et al.*, 2018).

L'une des solutions est d'entraîner les systèmes de RIC avec des simulations utilisateurs (SU). Les SU génèrent des données synthétiques proches des données générées par de vrais utilisateurs. Cela permet 1) d'estimer les actions les plus probables des utilisateurs notamment dans des nouveaux scénarios et 2) d'optimiser au mieux les politiques de systèmes de RIC en explorant des trajectoires non contenues dans les bases de données. Dans cet article nous nous intéressons à la problématique de modélisation et de simulation utilisateurs dans différents contextes d'accès à l'information (e.g., RI, recommandation, systèmes conversationnels, ...).

1) Nous proposons un état de l'art des approches de modélisation et de simulation utilisateur pour l'accès à l'information (Section 2). Modélisation et simulation utilisateurs sont souvent confondues dans la littérature. Pour un système source, dans notre cas l'utilisateur dont le comportement est inconnu, le modèle associé est un ensemble d'instructions, de règles ou d'équations construites à partir des données observées provenant cet utilisateur. La simulation est un agent qui génère des comportements en suivant les règles d'un modèle. Conjointement, le but est d'émuler le comportement de l'utilisateur (Zeigler, 1984).

2) Nous évoquons les protocoles d'évaluation des SU de l'état de l'art, la difficulté étant de mesurer si le comportement simulé est réaliste (Section 3).

3) Nous discutons enfin les perspectives de recherche pour construire des simulations utilisateurs pour la RIC (Section 4).

## 2. Etat de l'art

### 2.1. Modélisation et interprétation des comportements utilisateurs

Les premières approches de modélisations utilisateurs en RI se sont basées sur des approches cognitives (Belkin, 1984). Ce dernier produit une description des utilisateurs caractérisés par un objectif ou un problème ainsi qu'une connaissance du monde. Il décrit aussi les conditions nécessaires pour qu'un système de RI soit efficace face à des requêtes imprécises. De nombreux dérivés ont été ensuite proposés (e.g., ISP (Kuhlthau, 1991), IS&R (Ingwersen et Järvelin, 2005), Ellis (Ellis, 1989)). D'autres modèles utilisateurs en RI ont été proposés en s'appuyant sur d'autres théories. Par exemple, (Azzopardi, 2014) fait une analogie avec l'économie en appliquant la théorie de production pour modéliser les interactions entre utilisateurs et systèmes.

D'autres travaux se sont intéressés à l'analyse statistique des logs utilisateurs pour mettre en évidence des préférences ou des comportements utilisateurs types. (Kelly et Belkin, 2001) ont relié les interactions des utilisateurs à des indicateurs de pertinence

de documents. Plus récemment, (Chuklin *et al.*, 2013) modélisent les comportements des utilisateurs vis-à-vis des pages de résultats (e.g., les utilisateurs qui accèdent à la seconde page de résultat ont tendance à continuer sur les pages de résultat suivantes). Cependant, modéliser le comportement des clics utilisateurs est compliqué, car les clics ne sont pas un indicateur parfait concernant la pertinence des documents retrouvés. En effet, les utilisateurs peuvent être victimes de biais de position (Wang *et al.*, 2018) ou de biais de présentation (Yue *et al.*, 2010) (certains titres sont plus attractifs du fait de la présence d’images ou de mots clefs). La diversité dans les comportements utilisateurs complexifie aussi la modélisation, en effet pour une même requête, des utilisateurs n’ont pas les mêmes attentes donc pas les mêmes comportements de clics (Hu *et al.*, 2011). Un certain nombre de modèles statistiques sont introduits pour interpréter les comportements utilisateurs. L’un des premiers modèles de clics utilisateurs qui permet d’estimer la probabilité de cliquer sur un document est le modèle *cascade* (Craswell *et al.*, 2008). Les auteurs se limitent à un seul clic et supposent que les utilisateurs parcourent la liste de documents retrouvés du haut vers le bas jusqu’à trouver un document pertinent. (Chapelle et Zhang, 2009) proposent un modèle bayésien pour estimer la pertinence des documents en fonction des clics des utilisateurs, malgré le biais de position. (Agichtein *et al.*, 2006) montre des corrélations entre des caractéristiques des comportements utilisateurs et leur satisfaction. (Dupret et Piwowarski, 2008) introduisent le *user browsing model* (UBM) qui permet d’estimer la probabilité de clic sur un document de rang  $r$  en fonction du rang  $r'$  du document précédemment cliqué malgré le biais de position. (Shen *et al.*, 2012) modélisent le comportement utilisateur en fonction des requêtes et des documents tout en prenant en compte le biais induit par la variabilité des comportements utilisateurs. Elle utilise ce modèle pour faire du re-ranking. Récemment, de nombreux travaux se sont intéressés à comprendre le besoin des utilisateurs dans l’objectif de construire sur le long terme des systèmes conversationnels. Les approches se sont principalement intéressées au processus de reformulation de requêtes des utilisateurs (Zamani *et al.*, 2020 ; Sordoni *et al.*, 2015a ; Mustar *et al.*, 2020).

## 2.2. Modèles de simulations utilisateurs pour l’accès à l’information

Dans la partie précédente nous avons réalisé un état de l’art des différentes méthodes pour estimer les comportements utilisateurs à partir de données observées dans un contexte de RI. Dans cette partie on s’intéresse à l’état de l’art de l’utilisation de telles modélisations pour la simulation utilisateur et leurs applications. Nous dénombrons trois grandes catégories de modèles de simulation utilisateurs, très corrélées avec les tendances actuelles de modèles d’apprentissage profond.

La première catégorie de modèles utilisateurs repose principalement sur des modèles utilisateurs formalisés sur la base de règles ou de modèles probabilistes. Historiquement, les simulations utilisateurs ont été très étudiées pour les systèmes conversationnels (Eckert *et al.*, 1997 ; Komatani *et al.*, 2005 ; Pietquin, 2004 ; Schatzmann *et al.*, 2006 ; Scheffler et Young, 2000). Une des premières tentatives est d’utiliser directement des modèles statistiques utilisateur comme simulateur (Eckert *et al.*, 1997).

Celui-ci simule les utilisateurs à partir des probabilités d'apparition d'une phrase en fonction de la phrase précédente. Cette simulation est limitée, car elle ne prend en compte que la réponse précédente du système, les probabilités souffrent du manque de donnée et de plus elle n'est pas guidée par un objectif de dialogue. Pour ajouter de la consistance dans les réponses des utilisateurs, (Scheffler et Young, 2000) introduit dans son simulateur un objectif de dialogue, tiré parmi une liste d'objectif possible. Le modèle suit un objectif de dialogue imposé : les probabilités sont conditionnées par les objectifs des vrais utilisateurs dans le corpus. (Komatani *et al.*, 2005) apporte de la diversité avec un modèle qui conditionne les probabilité sur la capacité de l'utilisateur à utiliser un système de recherche ainsi que son degré d'urgence. À partir de données annotées, un arbre de décision est utilisé pour classifier les phrases du corpus en fonction de l'urgence et de la compétence des utilisateurs. Pour la simulation, un degré d'urgence est tiré et imposé au modèle pour la génération de phrases. Une façon de mettre en oeuvre des règles est d'utiliser des Agendas, comme dans (Schatzmann *et al.*, 2007). Dans ce travail, les auteurs simulent un utilisateur conversationnel sur la base de contraintes et de règles qui permettent d'encoder de manière explicite l'état de l'utilisateur ainsi que son objectif. En plus de générer des phrases suivant un certain modèle utilisateur, le simulateur se voit attribuer un objectif de dialogue ainsi qu'une stratégie de dialogue (Agenda) qui conditionne les phrases générées par le modèle. Ces approches basées sur des Agenda sont très utilisées pour la recommandation conversationnelle car elles permettent de construire facilement des stratégies de dialogue pour des domaines bien spécifiques (Hou *et al.*, 2019 ; Li *et al.*, 2016). (Zhang et Balog, 2020) montrent que des SU Agendas peuvent être particulièrement réalistes : dans le cadre de la recommandation de films, 36% des évaluateurs humains ont confondu une SU agenda avec un vrai utilisateur et 23% des évaluateurs étaient indécis.

Certains travaux plus récents conservent l'Agenda pour gérer la stratégie de dialogue utilisateur (Li *et al.*, 2016). Les auteurs proposent un modèle hybride plus réaliste qui s'appuie sur l'agenda pour conserver l'état de l'utilisateur mais qui profite de la puissance des modèles de langue *seq2seq* pour améliorer les modules de génération et de compréhension de langue naturelle. A l'inverse, (Asri *et al.*, 2016) et (Kreyssig *et al.*, 2018) proposent un modèle neuronal qui s'affranchit de l'agenda en utilisant des modèles de langue *seq2seq* pour générer et interpréter le langage naturel. En revanche ils conservent la génération d'objectifs (contraintes/requêtes) au niveau sémantique. Pour simuler le comportement dynamique de l'utilisateur, les contraintes sont tirées indépendamment selon une probabilité observée dans le dataset.

La deuxième catégorie de travaux s'appuie sur l'apprentissage profond pour simuler des utilisateurs à partir des données (Chandramohan *et al.*, 2011 ; Chen *et al.*, 2019 ; Zhao *et al.*, 2019). Par exemple, (Zhao *et al.*, 2019) propose d'apprendre une simulation utilisateur qui renseigne sur la pertinence des recommandations faites par le système. Les auteurs utilisent une approche générative adversaire (Goodfellow *et al.*, 2014) pour converger vers une simulation utilisateur. Le discriminateur doit faire la différence entre des feedbacks simulés et réels en fonction de l'historique de recherche utilisateur. Le générateur apprend à créer des feedbacks en fonction d'un historique d'interactions. En s'appuyant sur les approches par apprentissage par ren-

forcement, (Chen *et al.*, 2019) proposent de modéliser conjointement la politique des clics utilisateurs et la fonction de récompense grâce à une approche générative. En effet, la politique utilisateur est optimisée pour générer les actions qui maximisent le signal de récompense. En parallèle, la fonction de récompense a le rôle du discriminateur, elle est optimisée pour minimiser la différence de récompense entre l'action prise par la politique utilisateur et l'action prise par un utilisateur réel observé dans la base de données. Une approche d'apprentissage à partir des données repose sur l'apprentissage par renforcement inverse dont l'objectif est d'apprendre la fonction de récompense à partir des observations de comportement d'un autre agent (souvent humain) (Russell, 1998). Cette approche s'oppose aux modèles précédents qui reposaient sur des fonctions de récompense définies explicitement. Dans ce contexte, (Chandramohan *et al.*, 2011) proposent d'utiliser l'apprentissage par renforcement inverse pour retrouver la fonction de récompense décrivant les politiques utilisateurs observé dans des données. Dans le papier, une expérience est réalisée à partir de données conversationnelles dans le domaine du renseignement touristique.

La troisième catégorie de travaux s'appuie sur les limites des modèles d'apprentissage profond, requérant un nombre important de données, pour développer des modèles moins coûteux. Pour cela, (Zhang *et al.*, 2019) propose d'inclure l'humain dans la boucle d'apprentissage de manière parcimonieuse. Le modèle utilisateur apprend en parallèle du système de recommandation. Le modèle utilisateur hybride propose au choix une réponse générée par le modèle utilisateur ou par un humain. Le but est d'optimiser l'entraînement pour limiter l'intervention humaine. Pour cela, le modèle utilisateur se voit attribuer un budget limité. A chaque tour, le modèle peut choisir comment allouer ses crédits pour répondre au système de recommandation : 0 crédit pour répondre en utilisant la politique apprise, 1 crédit pour avoir accès à la réponse d'un humain, 2 crédit pour collecter une discussion humain-humain. Le modèle utilisateur apprend des nouvelles données collectées. Au début de chaque conversation, un générateur d'objectif tire des objectifs non réussis ou non explorés. Malgré une utilisation parcimonieuse des interventions humaines, inclure l'humain dans la boucle d'apprentissage reste coûteux.

### 3. Evaluation des simulations utilisateurs

Comme expliqué dans l'introduction, l'un des objectifs des SU est d'évaluer les systèmes d'accès à l'information (Eckert *et al.*, 1997). Il est donc nécessaire de s'assurer de la qualité des SU utilisés pour l'évaluation. D'après (Pietquin et Hastie, 2008), une bonne métrique pour l'évaluation de SU doit être capable de mesurer la qualité de la SU indépendamment de la tâche et doit être corrélée avec les évaluations de vrais utilisateurs. On s'intéresse ici aux différentes méthodes d'évaluation des SU. (Pietquin et Hastie, 2008) distinguent deux types d'évaluation.

L'évaluation directe mesure directement les SU au travers de métriques. Par exemple, une approche d'évaluation consiste à mesurer si la politique a réalisé son objectif (e.g., qualité des phrases générées dans un système de dialogue grâce à des mesures BLEU/ROUGE/METEOR (Papineni *et al.*, 2002 ; Lin, 2004 ; Banerjee et

Lavie, 2005 ; Sordoni *et al.*, 2015b), comportement de clic vraisemblable mesuré par la perplexité (Dupret et Piwowarski, 2008)). Cependant, dans le contexte des systèmes de dialogue, (Liu *et al.*, 2016) démontre que les métriques automatiques qui évaluent le chevauchement des mots avec des phrases de références comme BLEU, ROUGE et METEOR sont très peu corrélées avec le jugement humain, notamment car il existe une grande diversité de réponses valides pour chaque tour de conversation (Liu *et al.*, 2016 ; Zhang et Balog, 2020). L'évaluation humaine est aussi utilisée pour évaluer la qualité des réponses d'une simulation utilisateur. Par exemple, (Zhang et Balog, 2020) observe le nombre de fois que les évaluateurs humains ne peuvent pas faire la différence entre des vraies réponses utilisateurs et des dialogues générés par SU pour comparer la qualité entre des SU.

L'évaluation indirecte mesure l'impacte d'une SU sur les performances du système de dialogue (Schatzmann *et al.*, 2006). (Shi *et al.*, 2019) évalue le réalisme de son modèle utilisateurs grâce à l'annotation humaine directe et indirecte. Pour l'évaluation directe, les évaluateurs jugent directement la qualité des phrases de différents SU. Les SU sont évalués sur la qualité de la langue (comme la grammaire), la cohérence des phrases entre les différents tour ainsi que diversité des phrases. Les évaluateurs donnent aussi une note générale des simulations. L'évaluation indirecte des SU est mesurée par le biais des performances des systèmes de recommandation qui ont interagit avec les SU. Des évaluateurs humains analysent a posteriori la qualité de la recommandation. Ils notent leur satisfaction, le naturel des conversations et l'efficacité des recommandations.

#### 4. Perspectives de recherche

Dans ce qui suit, nous discutons les perspectives de recherche que nous envisageons pour proposer des modèles de simulation utilisateur pour la RIC.

**Vers des modèles de simulation utilisateurs séquentiels et multi-actions.** Actuellement, les SU conversationnelles n'embarquent pas de modèles de clics utilisateurs. Générer des comportements utilisateurs complet sous la forme de séquences de clics et requêtes (dialogue) semble essentiel pour avoir des cadres d'évaluations réalistes pour les systèmes de RIC. Pour pallier au manque de donnée, une idée serait d'incorporer des modèles de clics déjà existants (Dupret et Piwowarski, 2008) à des modèles de langues. Une autre idée est d'apprendre conjointement le modèle de langue et le modèle de clics utilisateurs. Simuler à la fois la conversation et les clics semble essentielle pour apprendre des systèmes de RIC, notamment dans des tâches de clarification de requêtes (Zamani *et al.*, 2020).

**Vers des modèles de simulation utilisateurs guidé par la diversité.** L'une des difficultés apporté par les modèles neuronaux est de générer de la diversité dans les comportements. En effet, de tel modèles apprennent à satisfaire les comportements moyens observés dans les données. Or il existe une forte variabilité dans les comportements utilisateurs (White et Drucker, 2007). Certains sont plus persévérants ou curieux dans leur recherche tandis que d'autres abandonnent rapidement. Apprendre un seul modèle utilisateur sur des données ne reflète pas cette diversité des comporte-

ments. Pour une tâche de conversation (Engelbrecht *et al.*, 2009) introduit le MeMo workbench qui permet de modéliser les utilisateurs avec leur différentes caractéristique comme l'âge, le niveau d'expertise, l'anxiété etc. En revanche les modèles récents dans la SU modélise le comportement d'un seul utilisateur a partir des données générés par un ensemble de vrais utilisateurs. Une des idées serait de générer des politiques utilisateurs réalistes mais avec des styles de recherche différents. Une méthode serait de regrouper et modéliser à la main les différents utilisateurs suivant un certain nombre de critères (curiosité, persévérance...) comme (Engelbrecht *et al.*, 2009). Une autre solution serait d'utiliser des GAN pour générer des styles différents tout en s'assurant du réalisme grâce au discriminateur.

**Vers des modèles de simulation utilisateurs pour des domaines ouverts (RIC).**

En plus de la propriété de variété dans le comportement, la RIC impose de pouvoir évoluer sur une variété de domaines (domaines ouverts). La grande majorité des simulations utilisateurs pour les systèmes de recommandation conversationnelle (SRC), par exemple, sont spécifiques à un domaine limitant par conséquent les SRC à un nombre limité de domaines (e.g., suggestion de restaurants). Les assistants virtuels pour la RI ont besoin de proposer des services couvrant un nombre croissant de domaines. Une première étape vers ce domaine ouvert serait de bénéficier des pistes de recherche des SRC multi-domaines (Rastogi *et al.*, 2020), propulsés avec la publication de quelques jeux de données multi-domaines (Rastogi *et al.*, 2020 ; Budzianowski *et al.*, 2018).

**Vers la mise en place de benchmark pour les systèmes de RIC.** Les perspectives de recherche précédentes permettraient à long terme de proposer des cadres formels d'évaluation pour les modèles de RIC afin de couvrir des sessions de recherches plus réalistes du comportement utilisateur et de son besoin en information. A l'opposé, l'évaluation humaine est trop coûteuse pour tester la robustesse des modèles pour divers besoins en information et divers comportements. L'une des solutions est d'établir des modèles utilisateurs 'benchmark', qui seront utilisés pour mesurer les performances des différents modèles proposés. A noter que ce cadre a déjà été mis en place pour les systèmes de RC (Zhang et Balog, 2020).

## 5. Conclusion

Cet article s'intéresse à la RIC et surtout à la mise en place de cadre d'entraînement et d'évaluation pour les approches centrées utilisateurs. Bien que l'évaluation humaine soit la plus adaptée, elle est cependant trop coûteuse pour les modèles basés sur l'apprentissage profond ou par renforcement. L'objectif de ce papier est de faire une synthèse de l'état de l'art des approches de modélisation et de simulation utilisateurs dans diverses tâches d'accès à l'information pour ouvrir des perspectives de mises en oeuvre des systèmes de RIC.

## Remerciements

Ce travail est effectué dans le cadre du projet ANR JCJC SESAMS (ANR-18-CE23-0001).

## 6. Bibliographie

- Agichtein E., Brill E., Dumais S., “Improving Web Search Ranking by Incorporating User Behavior Information”, *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, Association for Computing Machinery, New York, NY, USA, p. 19–26, 2006.
- Asri L. E., He J., Suleman K., “A Sequence-to-Sequence Model for User Simulation in Spoken Dialogue Systems”, in N. Morgan (ed.), *Interspeech 2016, 17th Annual Conference of the International Speech Communication Association, San Francisco, CA, USA, September 8-12, 2016*, ISCA, p. 1151-1155, 2016.
- Azzopardi L., “Modelling Interaction with Economic Models of Search”, *Proceedings of the 37th International ACM SIGIR Conference on Research Development in Information Retrieval*, SIGIR '14, Association for Computing Machinery, New York, NY, USA, p. 3–12, 2014.
- Azzopardi L., Dubiel M., Halvey M., Dalton J., “Conceptualizing agent-human interactions during the conversational search process”, 2018.
- Banerjee S., Lavie A., “METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments”, *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, Association for Computational Linguistics, Ann Arbor, Michigan, p. 65-72, June, 2005.
- Belkin N. J., “Cognitive models and information transfer”, *Social Science Information Studies*, vol. 4, n° 2-3, p. 111-129, 4, 1984.
- Bordes A., Boureau Y.-L., Weston J., “Learning End-to-End Goal-Oriented Dialog”, *5th International Conference on Learning Representations, {ICLR} 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.
- Bordes A., Chopra S., Weston J., “Question Answering with Subgraph Embeddings”, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, Doha, Qatar, p. 615-620, 2014.
- Budzianowski P., Wen T.-H., Tseng B.-H., Casanueva I., Ultes S., Ramadan O., Gašić M., “MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling”, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA, p. 5016-5026, 2018.
- Chandramohan S., Geist M., Lefèvre F., Pietquin O., User Simulation in Dialogue Systems using Inverse Reinforcement Learning, Technical report, 2011.
- Chapelle O., Zhang Y., “A Dynamic Bayesian Network click model for web search ranking”, *WWW'09 - Proceedings of the 18th International World Wide Web Conference*, ACM Press, New York, New York, USA, p. 1-10, 2009.
- Chen X., Li S., Li H., Jiang S., Qi Y., Song L., “Generative Adversarial User Model for Reinforcement Learning Based Recommendation System”, in K. Chaudhuri, R. Salakhutdinov (eds), *Proceedings of the 36th International Conference on Machine Learning*, vol. 97 of *Proceedings of Machine Learning Research*, PMLR, p. 1052-1061, 09–15 Jun, 2019.
- Chuklin A., Serdyukov P., De Rijke M., “Modeling clicks beyond the first result page”, *International Conference on Information and Knowledge Management, Proceedings*, Association for Computing Machinery, p. 1217-1220, 2013.

- Craswell N., Zoeter O., Taylor M., Ramsey B., “An experimental comparison of click position-bias models”, *WSDM'08 - Proceedings of the 2008 International Conference on Web Search and Data Mining*, ACM Press, New York, New York, USA, p. 87-94, 2008.
- Culpepper J. S., Diaz F., Smucker M. D., “Research Frontiers in Information Retrieval: Report from the Third Strategic Workshop on Information Retrieval in Lorne (SWIRL 2018)”, *SIGIR Forum*, vol. 52, n<sup>o</sup> 1, p. 34–90, 2018.
- Dupret G., Piwowarski B., “A user browsing model to predict search engine click data from past observations”, *ACM SIGIR 2008 - 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Proceedings*, p. 331-338, 2008.
- Eckert W., Levin E., Pieraccini R., “User modeling for spoken dialogue system evaluation”, *IEEE Workshop on Automatic Speech Recognition and Understanding Proceedings*, IEEE, p. 80-87, 1997.
- Ellis D. G., “A Behavioral Approach to Information Retrieval System Design”, *Journal of Documentation*, vol. 45, n<sup>o</sup> 3, p. 171-212, 1989.
- Engelbrecht K.-P., Quade M., Möller S., “Analysis of a new simulation approach to dialog system evaluation”, *Speech Communication*, vol. 51, n<sup>o</sup> 12, p. 1234-1252, 2009.
- Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., “Generative Adversarial Nets”, *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS'14*, MIT Press, Cambridge, MA, USA, p. 2672–2680, 2014.
- Hou Y., Fang M., Che W., Liu T., A Corpus-free State2Seq User Simulator for Task-oriented Dialogue, Technical report, 2019.
- Hu B., Zhang Y., Chen W., Wang G., Yang Q., “Characterizing search intent diversity into click models”, *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*, ACM Press, New York, New York, USA, p. 17-26, 2011.
- Ingwersen P., Järvelin K., *The Turn: Integration of Information Seeking and Retrieval in Context (The Information Retrieval Series)*, Springer-Verlag New York, Inc., 2005.
- Jannach D., Manzoor A., Cai W., Chen L., “A Survey on Conversational Recommender Systems”, *arXiv*, 4, 2020.
- Kelly D., Belkin N. J., “Reading time, scrolling and interaction”, *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '01*, Association for Computing Machinery (ACM), New York, New York, USA, p. 408-409, 2001.
- Komatani K., Ueno S., Kawahara T., Okuno H. G., “User modeling in spoken dialogue systems to generate flexible guidance”, *User Modelling and User-Adapted Interaction*, vol. 15, Springer, p. 169-183, 3, 2005.
- Kreyssig F. L., Nigo Casanueva I. [U+FFFD], Budzianowski P., Gašić M. G., Neural User Simulation for Corpus-based Policy Optimisation for Spoken Dialogue Systems, Technical report, 2018.
- Kuhlthau C. C., “Inside the search process: Information seeking from the user’s perspective”, *Journal of the Association for Information Science and Technology (JASIST)*, vol. 42, n<sup>o</sup> 5, p. 361-371, 1991.
- Li X., Lipton Z. C., Dhingra B., Li L., Gao J., Chen Y.-N., A User Simulator for Task-Completion Dialogues \*, Technical report, 2016.

- Lin C.-Y., “ROUGE: A Package for Automatic Evaluation of Summaries”, *Text Summarization Branches Out*, Association for Computational Linguistics, Barcelona, Spain, p. 74-81, July, 2004.
- Liu C.-W., Lowe R., Serban I., Noseworthy M., Charlin L., Pineau J., “How {NOT} To Evaluate Your Dialogue System: An Empirical Study of Unsupervised Evaluation Metrics for Dialogue Response Generation”, *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Austin, Texas, p. 2122-2132, 2016.
- MontazerAlghaem A., Zamani H., Allan J., “A Reinforcement Learning Framework for Relevance Feedback”, *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '20, Association for Computing Machinery, New York, NY, USA, p. 59-68, 2020.
- Mustar A., Lamprier S., Piwowarski B., “Using {BERT} and {BART} for Query Suggestion”, in I. Cantador, M. Chevalier, M. Melucci, J. Mothe (eds), *Proceedings of the Joint Conference of the Information Retrieval Communities in Europe {(CIRCLE} 2020)*, Samatan, Gers, France, July 6-9, 2020, vol. 2621 of {CEUR} Workshop Proceedings, CEUR-WS.org, 2020.
- Papineni K., Roukos S., Ward T., Zhu W.-J., “{B}leu: a Method for Automatic Evaluation of Machine Translation”, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, p. 311-318, 2002.
- Pietquin O., A Framework for Unsupervised Learning of Dialogue Strategies, PhD thesis, Facult[U+FFFD] Polytechnique de Mons, TCTS Lab (Belgique), apr, 2004.
- Pietquin O., Hastie H., “A survey on metrics for the evaluation of user simulations”, *The Knowledge Engineering Review*, vol. 00, p. 0-1, 2008.
- Radlinski F., Craswell N., “A Theoretical Framework for Conversational Search”, *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, CHIIR '17, Association for Computing Machinery, New York, NY, USA, p. 117-126, 2017.
- Rastogi A., Zang X., Sunkara S., Gupta R., Khaitan P., “Towards Scalable Multi-Domain Conversational Agents: The Schema-Guided Dialogue Dataset”, *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, n<sup>o</sup> 05, p. 8689-8696, Apr., 2020.
- Russell S., “Learning agents for uncertain environments (extended abstract)”, *Proceedings of the eleventh annual conference on Computational learning theory - COLT' 98*, ACM Press, New York, New York, USA, p. 101-103, 1998.
- Schatzmann J., Thomson B., Weilhammer K., Ye H., Young S., Agenda-Based User Simulation for Bootstrapping a POMDP Dialogue System, Technical report, 2007.
- Schatzmann J., Weilhammer K., Stuttle M., Young S., “A Survey of Statistical User Simulation Techniques for Reinforcement-Learning of Dialogue Management Strategies”, *The Knowledge Engineering Review*, vol. 00, p. 0-1, 2006.
- Scheffler K., Young S., “Probabilistic simulation of human-machine dialogues”, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2, Institute of Electrical and Electronics Engineers Inc., p. 1217-1220, 2000.
- Shen S., Hu B., Chen W., Yang Q., “Personalized click model through collaborative filtering”, *Proceedings of the fifth ACM international conference on Web search and data mining - WSDM '12*, ACM Press, New York, New York, USA, p. 323, 2012.

- Shi W., Qian K., Wang X., Yu Z., “How to Build User Simulators to Train {RL}-based Dialog Systems”, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, Hong Kong, China, p. 1990-2000, 2019.
- Sordoni A., Bengio Y., Vahabi H., Lioma C., Simonsen J. G., Nie J. Y., “A hierarchical recurrent encoder-decoder for generative context-aware query suggestion”, *International Conference on Information and Knowledge Management, Proceedings*, vol. 19-23-Oct-2015, Association for Computing Machinery, New York, NY, USA, p. 553-562, 10, 2015a.
- Sordoni A., Galley M., Auli M., Brockett C., Ji Y., Mitchell M., Nie J.-Y., Gao J., Dolan B., “A Neural Network Approach to Context-Sensitive Generation of Conversational Responses”, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Association for Computational Linguistics, Stroudsburg, PA, USA, p. 196-205, 2015b.
- Sun Y., Zhang Y., “Conversational Recommender System”, *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018* p. 235-244, 6, 2018.
- Sutton R. S., Barto A. G., *Reinforcement Learning: An Introduction*, MIT Press, 1998.
- Wang X., Golbandi N., Bendersky M., Metzler D., Najork M., “Position Bias Estimation for Unbiased Learning to Rank in Personal Search”, *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, WSDM '18*, Association for Computing Machinery, New York, NY, USA, p. 610–618, 2018.
- White R. W., Drucker S. M., “Investigating behavioral variability in web search”, *16th International World Wide Web Conference, WWW2007*, p. 21-30, 2007.
- Yue Y., Patel R., Roehrig H., *Beyond Position Bias: Examining Result Attractiveness as a Source of Presentation Bias in Clickthrough Data*, 2010.
- Zamani H., Dumais S., Craswell N., Bennett P., Lueck G., “Generating Clarifying Questions for Information Retrieval”, *Proceedings of The Web Conference 2020*, ACM, New York, NY, USA, p. 418-428, 4, 2020.
- Zeigler B. P., *Theory of Modelling and Simulation | Guide books*, Krieger Publishing Co., Inc, 1984.
- Zhai C., “Interactive Information Retrieval: Models, Algorithms, and Evaluation”, *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '20*, Association for Computing Machinery, New York, NY, USA, p. 2444–2447, 2020.
- Zhang S., Balog K., “Evaluating Conversational Recommender Systems via User Simulation”, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, p. 1512-1520, 8, 2020.
- Zhang W., Zhao X., Zhao L., Yin D., Yang G. H., Beutel A., “Deep Reinforcement Learning for Information Retrieval: Fundamentals and Advances”, *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, Inc, p. 2468-2471, 7, 2020.
- Zhang Z., Li X., Gao J., Chen E., “Budgeted Policy Learning for Task-Oriented Dialogue Systems”, *Proceedings of the 57th Annual Meeting of the Association for Computational Lin-*

*guistics*, Association for Computational Linguistics, Florence, Italy, p. 3742-3751, July, 2019.

Zhao X., Xia L., Zou L., Yin D., Tang J., Simulating User Feedback for Reinforcement Learning Based Recommendations, Technical report, 2019.