

Interpréter l'intention de conception à partir des observables verbaux : apports du *machine learning* à l'étude de l'activité de conception

Gaëlle BAUDOUX, Dr., LUCID, Université de Liège

Christine MARCELINO, Co-Design Lab, Université de Berkely, Californie

Résumé. Comprendre l'intention de conception est un enjeu central pour l'analyse des processus cognitifs en conception. Les protocoles de verbalisation de type *think-aloud* permettent d'y accéder, mais leur analyse manuelle reste coûteuse. Cette recherche explore le potentiel du *machine learning* pour interpréter automatiquement l'intention de conception à partir de verbalisations issues d'un corpus de dix-neuf sessions de conception, l'intention de conception est classifiée selon trois dimensions (forme, fonction et matérialité) et les résultats d'un modèle de langage contextuel sont comparés à un codage manuel. Les résultats montrent un taux global de concordance de 73,69 %, avec de meilleures performances pour la fonction et la matérialité que pour la forme. L'analyse met en évidence des limites liées à l'ambiguïté des catégories, au recouvrement des intentions et aux limites des observables verbaux. Ces travaux montrent que le *machine learning* constitue un outil complémentaire pour l'analyse de l'activité de conception et ouvrent des perspectives méthodologiques pour des approches multimodales à plus grande échelle.

Mots-clefs. Analyse de l'activité, Protocoles et méthodes, Intention de conception, Observables verbaux, Think-aloud, Machine learning

Introduction

L'analyse des processus cognitifs en conception repose en grande partie sur la capacité à accéder aux intentions qui sous-tendent les décisions prises par les concepteurs. Elle permet de comprendre non seulement ce qui est conçu, mais aussi pourquoi certaines décisions sont prises à un moment donné du processus. L'intention de conception joue ainsi un rôle structurant dans l'interprétation des raisonnements, des choix et des ajustements opérés par les concepteurs au fil de l'activité. Sur le plan théorique, l'intention de conception est étroitement liée aux notions de *rationale* de conception et de réflexion en action. Elle renvoie à la justification des décisions et à la projection d'objectifs, de contraintes et de critères qui orientent l'activité de conception. Contrairement à l'artefact final, l'intention de conception n'est pas directement observable : elle se manifeste à travers les actions, les interactions avec les outils et les verbalisations produites au cours de l'activité. Comprendre l'intention de conception permet ainsi d'accéder à la dynamique cognitive de la conception, dans laquelle les objectifs, les hypothèses et les solutions évoluent en interaction avec la situation.

Pour étudier ces processus, la recherche en conception mobilise largement les protocoles de *think-aloud*, qui consistent à demander aux concepteurs de verbaliser leurs pensées pendant qu'ils réalisent une tâche (Ericsson & Simon, 1993). Des recherches récentes ont cartographié de manière systématique l'utilisation des protocoles de verbalisation en conception, mettant en évidence la variété des processus cognitifs impliqués en conception et documentés par ces méthodes (Hay et al., 2017). Ces protocoles produisent des énoncés verbaux considérés comme des traces valides de l'activité cognitive, permettant d'identifier les intentions, les stratégies et les raisonnements mobilisés. Cependant, l'analyse de ces données repose traditionnellement sur des méthodes de codage manuel, exigeant un important investissement en temps, une formation spécifique des codeurs et une forte implication humaine. Ces contraintes limitent la scalabilité des études et rendent difficile l'analyse de grands corpus de données verbales.

Dans ce contexte, les approches basées sur le *machine learning* offrent des perspectives nouvelles pour l'analyse de l'intention de conception. Cette perspective est au cœur des discussions contemporaines sur l'avenir de la cognition en conception, qui soulignent la nécessité d'approches méthodologiques mixtes

et innovantes (Hay, Cash & McKilligan, 2020). En permettant l'identification automatique de patterns dans de larges ensembles d'énoncés verbaux, ces méthodes ouvrent la voie à une analyse plus systématique et reproductible des données issues des protocoles *think-aloud*. Cette étude explore la capacité du *machine learning* à inférer l'intention de conception à partir de données verbales recueillies lors d'une étude portant sur un logiciel génératif multimodal basé sur le dessin pour l'architecture. Les énoncés verbaux des participants sont utilisés à la fois comme support d'analyse et comme données d'entraînement pour un modèle de langage contextuel (*LLM*) chargé de classifier automatiquement l'intention de conception. Dans ce contexte, cette recherche vise à répondre à la question suivante : "Dans quelle mesure des approches de *machine learning* peuvent-elles permettre d'identifier et de classifier l'intention de conception à partir d'énoncés verbaux issus de protocoles *think-aloud*, dans un contexte de conception architecturale ?"

La contribution principale de ce travail réside dans l'introduction de techniques de traitement automatique du langage (*NLP*) pour soutenir l'analyse de l'intention de conception, en complément des méthodes qualitatives traditionnelles. Nous présentons le cadre théorique et expérimental de l'étude, la méthodologie de *machine learning* mise en œuvre, ainsi que les résultats obtenus et leurs implications pour la recherche en conception et en analyse de l'activité.

Cadre théorique et méthodologique

Analyse de l'activité et cognition de conception

Cette recherche s'inscrit dans le champ de l'analyse de l'activité de conception, qui vise à comprendre comment les concepteurs mobilisent leurs connaissances, leurs outils et leur environnement pour produire des solutions. Les travaux en cognition de conception montrent que l'activité de conception n'est ni linéaire ni strictement planifiée : elle se caractérise par des allers-retours constants entre formulation du problème, génération de solutions, évaluation et reformulation (Cross et al., 1996). Les intentions évoluent ainsi en fonction des contraintes, des opportunités et des retours produits par la situation de conception.

Dans cette perspective, l'activité de conception est fondamentalement située. Schön (1992) décrit la conception comme une réflexion en action, dans laquelle le concepteur ajuste en permanence son raisonnement en réponse aux effets de ses actions et aux résistances ou suggestions de la situation. L'intention de conception se construit et se transforme dans l'interaction entre le concepteur, les outils mobilisés et l'environnement de conception. L'introduction de nouvelles technologies, et en particulier d'outils génératifs basés sur l'intelligence artificielle, modifie cette dynamique et reconfigure les modes de raisonnement et d'exploration des solutions.

Définition et modélisation de l'intention de conception

L'intention de conception est un concept central mais intrinsèquement polysémique. Dans la recherche en conception assistée par ordinateur (CAO), elle désigne la logique sous-jacente qui guide la structure d'un modèle et anticipe son comportement lors de futures modifications (Henderson, 1993 ; Otey et al., 2014). Elle ne se limite pas à la description de la forme finale, mais renvoie aux objectifs, contraintes et justifications qui motivent les choix de conception.

Henderson définit l'intention de conception comme la finalité ou la justification sous-jacente d'un artefact, tandis que les travaux sur le *design rationale* mettent l'accent sur l'explicitation des raisons pour lesquelles une décision a été prise plutôt qu'une autre (Moran & Carroll, 1996). Dans cette perspective, l'intention de conception est considérée comme une représentation explicite et communicable du raisonnement de conception, distincte de l'artefact final. Ganeshan et al. (1991) proposent ainsi de représenter l'intention de conception à travers des modèles distinguant objectifs, actions et justifications.

Dans le contexte de cette étude, l'intention de conception est abordée comme une manifestation cognitive observable, exprimée à travers les énoncés verbaux produits par les concepteurs au cours de l'activité. En architecture, cette intention peut être appréhendée à travers des dimensions fondamentales telles que la forme, la fonction et la matérialité, qui structurent traditionnellement le raisonnement architectural. Ces catégories ne prétendent pas couvrir l'ensemble de la cognition de conception, mais offrent un cadre opératoire permettant de classifier et d'analyser les intentions exprimées.

Protocoles think-aloud comme méthode d'accès à l'activité cognitive

Les protocoles de *think-aloud* constituent une méthode de référence pour étudier les processus cognitifs en situation de conception (Ericsson & Simon, 1993). Ils reposent sur la verbalisation en temps réel des pensées du concepteur pendant l'exécution d'une tâche. Les énoncés verbaux produits sont considérés comme des traces empiriques valides de l'activité cognitive, permettant de rendre explicites des processus autrement inobservables.

En recherche en conception, ces protocoles sont largement utilisés pour analyser les stratégies, les intentions et l'évolution des états de connaissance des concepteurs. Les énoncés verbaux sont généralement segmentés, codés et catégorisés afin d'identifier des types d'activités ou d'intentions. Toutefois, cette analyse repose sur un codage manuel exigeant, nécessitant des règles précises, une formation des analystes et un important investissement en temps. De plus, la subjectivité inhérente au codage peut conduire à des divergences d'interprétation, notamment lorsque les énoncés relèvent de plusieurs dimensions simultanément. Pour renforcer la validité interne, il est recommandé de recourir à une validation croisée par plusieurs codeurs en aveugle et d'évaluer le consensus inter-codeurs à l'aide d'indicateurs standardisés (comme par exemple les coefficients de fiabilité). Cette étape méthodologique permet de réduire le risque de codage arbitraire et d'augmenter la fiabilité des catégories analytiques utilisées mais elle est coûteuse en temps et en ressources.

Ces limites méthodologiques constituent un frein à l'analyse de corpus volumineux, en particulier dans des contextes de conception qui génèrent de grandes quantités de données verbales et interactionnelles.

Apports du NLP et du machine learning pour l'analyse de l'activité

Le traitement automatique du langage (*NLP*) offre des perspectives pour dépasser certaines limites des méthodes manuelles d'analyse des données verbales. Depuis les premiers travaux de Winograd, la recherche en interaction humain-machine souligne que la compréhension du langage ne peut se réduire à une analyse syntaxique, mais nécessite une prise en compte du contexte sémantique et cognitif dans lequel les énoncés sont produits. Des chercheurs ont ainsi exploré des cadres conceptuels tels que le behavioural design space, reliant motifs cognitifs et exploration de l'espace de conception (Nielsen et al., 2021).

Les approches contemporaines de *NLP* reposent sur des représentations vectorielles du langage permettant de capturer des relations sémantiques complexes (Siddharth, Blessing & Luo, 2022). Ces méthodes ont été utilisées pour analyser des documents de conception, extraire des raisonnements implicites et identifier des patterns cognitifs dans des corpus textuels (Liang et al., 2018). Plus récemment, les modèles de langage de grande taille (*LLM*) ont montré leur capacité à traiter des volumes importants de données verbales tout en intégrant une compréhension contextuelle fine.

Dans le cadre de l'analyse de l'activité de conception, le *machine learning* permet d'envisager une classification automatique des énoncés verbaux en fonction de l'intention exprimée. Contrairement aux approches à base de règles prédéfinies, ces modèles peuvent s'adapter à la variabilité du langage et aux évolutions de l'intention au cours de l'activité. Cette capacité est particulièrement pertinente dans des contextes de conception assistée par des outils génératifs, où les interactions peuvent produire des formes inédites de raisonnement et de verbalisation.

Cette étude s'inscrit dans cette dynamique en examinant dans quelle mesure des modèles de *machine learning* peuvent inférer l'intention de conception à partir d'énoncés verbaux produits lors d'une activité de conception architecturale médiée par un outil génératif multimodal.

Méthode

Situation d'étude et protocole expérimental

Cette étude vise à analyser l'intention de conception telle qu'elle s'exprime au cours de l'activité de conception, en s'appuyant sur des protocoles de *think-aloud* recueillis dans une situation de conception architecturale médiée par un outil génératif multimodal. Vingt participants, étudiants en architecture à l'Université de Liège, ont pris part à l'étude. Tous étaient francophones et disposaient d'une formation préalable aux méthodes traditionnelles de conception architecturale.

La tâche consistait à concevoir un logement annexe de petite échelle, destiné à être implanté dans un espace résidentiel existant (par exemple un jardin privé), à l'aide d'un outil génératif intégrant plusieurs modalités d'interaction : texte, croquis et références visuelles. Ce type de programme architectural a été choisi pour sa complexité modérée, permettant d'observer des raisonnements liés à la forme, aux usages et à la matérialité sans imposer de contraintes excessives.

Avant le début de la session expérimentale, les participants ont bénéficié d'une présentation de l'outil, suivie d'une phase de prise en main de dix minutes. La session de conception, d'une durée totale de cinquante minutes, était structurée en trois phases indicatives :

- une phase orientée vers la fonction (20 minutes);
- une phase orientée vers la forme (20 minutes);
- une phase de conception libre (10 minutes).

Ces consignes visaient à favoriser l'expression verbale de différentes dimensions de l'intention de conception, sans contraindre strictement le déroulement de l'activité.

Données recueillies et préparation du corpus

Les données principales de l'étude sont constituées des énoncés verbaux produits par les participants durant les protocoles *think-aloud*. Les sessions ont été enregistrées en vidéo, puis transcrites sous forme de fichiers horodatés. Dix-neuf participants ont été retenus pour l'analyse, un participant ayant été exclu en raison des données de transcriptions insuffisantes.

Les énoncés verbaux ont été segmentés en unités correspondant à des intentions ou actions de conception explicites. Le tableau 1 illustre un exemple de cette segmentation. Ont été conservés uniquement les segments faisant référence à l'activité de conception ou à l'intention de conception, incluant les objectifs, les caractéristiques du projet et les justifications associées. Les segments hors activité (discours du facilitateur, commentaires métacognitifs, ou propos hors sujet) ont été exclus du corpus.

Tableau 1 – Extrait illustratif de la segmentation des énoncés verbaux.

start_time	end_time	Énoncé verbal
00:41:41,151	00:41:43,312	Ok, là je commence à trop réfléchir comme chez moi.
00:44:44,152	00:44:56,886	Ça c'est nouveau. Mais ça ressemble... je vois ce que tu veux dire. Déjà, avec l'histoire du verre autour. Puis une façade, etc. Ça commence déjà à ressembler à quelque chose.
00:44:56,886	00:45:02,687	En fait, ce type de rendu est vraiment lié à la façade réaliste. Comparé à l'autre, effectivement.
00:45:27,289	00:45:40,788	Je commence à voir ce qu'il a en tête. Non, en fait non. Après, il veut mettre une sorte de mur en bois au bout. Peut-être parce que je n'ai pas bien dessiné.

Les segments non pertinents pour l'analyse de l'activité (discours du facilitateur, commentaires métacognitifs, propos hors sujet ou énoncés ambigus) ont été exclus.

Dans la mesure où les modèles de *machine learning* les plus performants utilisés dans cette étude reposent sur des corpus d'entraînement majoritairement anglophones (Devlin et al., 2019), les transcriptions ont été traduites en anglais avant traitement. Les transcriptions ont été traduites en anglais selon deux modalités distinctes :

- pour les participants P1 à P12, une traduction automatique (Google Translate et ChatGPT) suivie d'une relecture par un locuteur natif francophone;
- pour les participants P13 à P20, une traduction réalisée directement par un locuteur natif francophone.

Cette double approche permettait d'évaluer la robustesse d'une chaîne de traitement computationnelle relativement indépendante de la langue source. L'ensemble des transcriptions a été vérifié dans le contexte du dialogue complet afin d'assurer la fidélité des traductions.

Catégorisation de l'intention de conception

Chaque énoncé verbal a été codé manuellement selon une grille de classification reposant sur trois dimensions fondamentales de la conception architecturale : forme, fonction et matérialité. Cette catégorisation a servi de référence pour l'entraînement et l'évaluation des modèles de *machine learning*. La forme renvoie aux aspects visuels et géométriques du projet (ex. : “*J’imaginai quelque chose de plus cubique.*”). La fonction concerne les usages, la distribution spatiale et les objectifs d’usage (ex. : “*Je vais juste rajouter le mobilier.*”). La matérialité recouvre les considérations liées aux matériaux, textures et qualités sensorielles (ex. : “*Pour la texture sur la façade, je verrais bien un bardage bois vertical.*”). Cette classification constitue une formalisation opératoire de l'intention de conception, permettant de relier les énoncés verbaux à des dimensions analytiques de l'activité.

Données issues des interactions multimodales

En complément des énoncés verbaux, le corpus inclut des données textuelles issues des interactions avec l'outil génératif, extraites de fichiers JSON (illustrées à la figure 1). Ces fichiers contiennent des métadonnées relatives aux prompts (horodatage, entrées utilisateur, sorties générées par l'IA), permettant de reconstruire la séquence des interactions.

Les informations issues des différentes modalités (texte, croquis, références visuelles) fournissent un contexte supplémentaire pour interpréter l'intention de conception, en articulant verbalisation, action et réponse de l'outil.

```
303 {
304   "Action": "Text to Image Generation ",
305   "Sketch": "None",
306   "Description_of_the_sketch": "Generate a professional rendering of the architectural design of a office shed in the backyard from the interior. It should be a RAW photo, architecture photography, and hyper realistic. Incorporate this description: can you draw a facade of the project of a mid-century californian pool house with a bedroom and bathroom and a small office. add the pool in the front",
307   "Generated_image": "0_36",
308   "Generated_time": "2025-02-21 10:41:35"
309 },
```

Fig. 1 – Extrait illustratif des informations extraites des fichiers JSON.

Approche de machine learning

Évaluation exploratoire sans apprentissage (*zero-shot*). Une première expérimentation a consisté à évaluer la faisabilité d'une classification automatique à l'aide d'un modèle *Sentence-BERT*, sans ajustement préalable sur les données de l'étude. BERT (*Bidirectional Encoder Representations from Transformers*) est un modèle de langage pré-entraîné pour la compréhension contextuelle du langage naturel (Devlin et al., 2019).

Les définitions des catégories Forme, Fonction et Matérialité ont été fournies au modèle afin de conditionner la sortie. Le modèle calcule la similarité sémantique entre un énoncé verbal et la description des catégories, en produisant des représentations vectorielles au niveau de la phrase. Cette approche *zero-shot* a été évaluée sur les données des cinq premiers participants et a permis d'atteindre une précision initiale de 65 %, suggérant la pertinence du *machine learning* pour cette tâche.

Affinage du modèle supervisé. Sur la base de ces résultats, un modèle *TFBertForSequenceClassification* a ensuite été affiné afin d'améliorer la classification des énoncés verbaux. Ce modèle repose sur l'architecture *bert-base-uncased*, pré-entraînée sur Wikipedia et BookCorpus (représentant environ 2,5 milliards de mots) (Devlin et al., 2019). Les mécanismes d'auto-attention de cette architecture permettent de capturer les dépendances contextuelles entre les mots d'un énoncé.

Ce modèle a été entraîné sur les énoncés verbaux issus des participants P1 à P6 dans un cadre d'apprentissage supervisé. L'entraînement a été réalisé sur quatre passes complètes du jeu de données (*epochs*), un choix cohérent avec les pratiques courantes de fine-tuning de modèles BERT sur des corpus de petite taille, et compatible avec les capacités de calcul disponibles.

Au cours de l'entraînement, la précision du modèle a évolué progressivement à chaque itération. La première passe correspond à une phase d'ajustement initial, au cours de laquelle le modèle commence à associer les énoncés verbaux aux catégories d'intention de conception. Les passes suivantes permettent un raffinement progressif de ces associations. À l'issue de la quatrième passe, le modèle atteint une précision d'entraînement de 93,5 %, avec une précision de validation d'environ 69,5 %.

Ces valeurs ont été retenues comme point d'équilibre pour l'entraînement du modèle, et servent de base à l'évaluation présentée dans la section suivante.

Résultats et interprétations

L'analyse des verbalisations issues de dix-neuf participants montre que le modèle de *machine learning* peut classifier l'intention de conception à partir des énoncés des concepteurs avec une fiabilité notable. Sur l'ensemble du corpus traité, 73,69 % des classifications automatiques correspondent aux codages manuels de référence. Cependant, cette performance globale masque des variations importantes selon les dimensions de l'intention de conception et entre participants, révélant à la fois des différences dans la manière dont les intentions sont exprimées et les limites des observables verbaux.

Différences de performance selon les dimensions de l'intention de conception

Le tableau 2 illustre les résultats du modèle de *machine learning* dans l'interprétation de l'intention de conception, par rapport au codage manuel, pour un extrait des énoncés verbaux. Les performances de classification varient fortement selon les dimensions étudiées :

- Matérialité : 85,47 %
- Fonction : 82,15 %
- Forme : 51,44 %

Cette hiérarchie suggère que certaines dimensions se prêtent mieux à une analyse basée sur les verbalisations. Les références à la matérialité mobilisent un lexique explicite et stable (par exemple : pierre, bois, texture, revêtement), facilitant la classification automatique. Les intentions fonctionnelles sont exprimées via des actions, des objectifs d'usage ou des relations spatiales explicites, fournissant des indices linguistiques suffisamment discriminants.

À l'inverse, la forme est souvent implicite, peu verbalisée ou exprimée de manière comparative ou métaphorique (ex. : "*cubic*", "*half in, half out*"), ce qui accroît l'ambiguïté sémantique et explique les écarts observés. De plus, la forme étant fréquemment représentée par le support graphique plutôt que verbale, le modèle rencontre des limites intrinsèques pour cette dimension.

Tableau 2 – Extrait d'interprétation de l'intention de conception par le modèle de *machine learning*, par rapport au codage manuel.

P	Énoncé verbal	Classification Manuelle	Machine Learning
P2	Je pourrais peut-être explorer tout ce qui est matériaux.	material	Material
P2	Je pourrais travailler la matérialité. Parce que là, à ce stade d'esquisse,	function, form	Material
P2	j'ai mon aménagement intérieur, j'ai mon volume extérieur.	material	Form
P2	Je veux de la pierre, mais je veux aussi un autre matériau.	material	Material
P2	Je vois qu'il a mis beaucoup de béton sur cette partie,	material	Material
P2	ou peut-être qu'il va mettre du bois ici.	material	Material

P2	Mais d'un autre côté, tu as utilisé une pierre un peu sombre.	material	Material
P2	Oui, ça c'est bien. Et puis il y a l'acrotère, qui est différent.	material	Material
P2	Ou alors peut-être un autre matériau à l'intérieur. Mais je ne sais pas si on voit l'intérieur.	material	Material
P2	Ben, à l'intérieur, c'est comme ça qu'on va le voir. Enfin, il t'a montré de la pierre,	material	Function
P2	non ? Sur le premier. Si tu mets du bois à l'intérieur,	material	Material
P2	Il va appliquer du bois à l'intérieur. En fait,	material	Material
P2	l'objectif ici, c'est de voir avec quoi je pourrais associer la pierre.	material	Material
P2	Tu as la façade en ardoise.	material	Material

Analyse des écarts entre codage manuel et classification automatique

L'examen des divergences montre que les écarts ne relèvent pas uniquement d'erreurs du modèle. Environ la moitié des discordances correspondent à des énoncés pouvant légitimement appartenir à plusieurs dimensions. Par exemple, la description d'un espace en termes de proportions ou de disposition peut être interprétée comme forme et fonction simultanément. Dans le corpus, environ 11% des désaccords relèvent de ce recouvrement catégoriel, proportion variable selon les participants.

Certaines divergences s'expliquent également par les propriétés de l'apprentissage automatique : le modèle a été fine-tuné sur un petit corpus avec quatre *epochs*, équilibre entre apprentissage suffisant et risque d'*overfitting*. Au-delà de quatre passes, le modèle aurait mémorisé le corpus plutôt qu'apprendre des motifs généralisables, ce qui aurait dégradé la performance sur de nouvelles données. Ainsi, certaines erreurs de classification reflètent des contraintes techniques de la méthodologie.

Variabilité inter-participants et explicitation du processus de conception

Les performances varient selon le style de verbalisation : certains participants produisent des énoncés structurés et explicites, facilitant la classification, tandis que d'autres s'expriment de manière plus elliptique ou s'appuient davantage sur le support graphique. Par exemple, P2 présente 33 % de ses énoncés pouvant appartenir à plusieurs dimensions, ce qui complique l'assignation univoque.

De plus, les verbalisations révèlent des glissements cognitifs : même lorsque les consignes indiquent de se concentrer sur la fonction ou la forme à des moments précis, les concepteurs revenaient spontanément sur la matérialité ou la forme, illustrant la nature dynamique et contextuelle de l'intention de conception. Ces observations soulignent que l'efficacité du modèle dépend à la fois de ses caractéristiques internes et du style cognitif et discursif du concepteur.

Portée des résultats pour l'analyse de l'intention de conception

Dans l'ensemble, ces résultats indiquent que le *machine learning* peut identifier des motifs sémantiques dans les verbalisations des concepteurs, en particulier pour les dimensions fonction et matérialité. La forme reste plus difficile à capturer, reflétant les limites des verbalisations comme observables des processus cognitifs.

Ces écarts ne doivent pas être interprétés comme de simples faiblesses du modèle, mais comme des indicateurs méthodologiques : ambiguïté des catégories, recouvrement des dimensions cognitives, dépendance aux modalités de représentation mobilisées par les concepteurs, et contraintes techniques liées au fine-tuning et à la taille du corpus. Ils préparent ainsi le terrain pour une discussion approfondie sur les limites et apports des méthodes automatiques dans l'étude de l'intention de conception.

Discussion

Cette étude visait à examiner si le *machine learning* pouvait interpréter l'intention de conception à partir de verbalisations produites par les concepteurs lors de la génération de concepts dans un contexte multimodal. Nos résultats montrent que, même avec un corpus limité, le modèle est capable de classifier de manière fiable les intentions liées à la fonction et à la matérialité, tandis que les intentions formelles restent plus difficiles à saisir.

Compréhension des processus cognitifs à travers l'intention de conception

Nos résultats montrent que les verbalisations sont une fenêtre partielle mais informative sur l'activité cognitive. Les intentions liées à la fonction et à la matérialité sont souvent explicitées directement dans le discours (ex. “*Je veux placer le lit face à la vue de la piscine*”, “*Je vais utiliser du bois pour le sol*”), alors que la forme est plus difficile à verbaliser et souvent liée à des représentations graphiques ou à des métaphores. Cette observation confirme que certaines dimensions de la cognition de conception sont moins accessibles via le langage seul, et que la combinaison multimodale (modalités verbale et graphique) reste essentielle.

Le corpus comprend des verbalisations en français alors que le modèle est entraîné en anglais. Malgré ce décalage linguistique, le modèle a réussi à identifier les patterns cognitifs pertinents, ce qui suggère une robustesse du *machine learning* face aux variations linguistiques. Néanmoins, certaines nuances culturelles ou expressions spécifiques ont pu introduire des divergences dans la classification, soulignant l'importance d'un encadrement méthodologique adapté lors de l'application à des corpus multilingues. Les verbalisations montrent que l'intention de conception est fluide et évolutive. Par exemple, lors de la phase initiale centrée sur la fonction, le participant P1 évoque simultanément des considérations formelles et matérielles. De même, P15 rapporte : “*Il est difficile de séparer fonction et forme sans les mélanger.*” Ces exemples illustrent que les concepteurs adaptent leur processus au contexte et aux suggestions générées par l'outil multimodal. Le *machine learning*, en capturant ces variations, montre sa capacité à suivre l'évolution dynamique des motifs cognitifs.

Apports et limites méthodologiques de l'approche ML

Le recours au *machine learning* présente plusieurs avantages méthodologiques par rapport aux méthodes traditionnelles de codage manuel :

- Il permet d'analyser de grands volumes de verbalisations de manière rapide et reproductible.
- Il identifie automatiquement des motifs sémantiques qui peuvent échapper à un codage humain, notamment lorsque les énoncés sont formulés différemment mais véhiculent la même intention.
- Et il offre une base pour l'analyse en temps réel de l'activité cognitive, ce qui serait prohibitif avec des protocoles de *think-aloud* traditionnels, exigeant un codage manuel minutieux.

Cependant, l'approche présente également des limites :

- Ambiguïté des catégories : certains énoncés peuvent légitimement appartenir à plusieurs dimensions (forme/fonction, fonction/matérialité), limitant l'efficacité des modèles à classification unique.
- Recouvrement des intentions : les glissements cognitifs observés (p. ex. un concepteur revenant spontanément à la matérialité alors que l'attention devrait être sur la fonction) illustrent que l'intention est fluide et contextuelle, difficile à capturer entièrement par une analyse sentence par sentence.
- Dépendance au langage : la dimension formelle, souvent représentée graphiquement, est moins bien capturée par un modèle exclusivement textuel, montrant les limites des verbalisations comme observables uniques.

- Contraintes techniques : le fine-tuning sur un petit corpus a nécessité un compromis entre apprentissage et sur-apprentissage (*overfitting*), limitant la capacité du modèle à généraliser sur de nouvelles données.

Limites méthodologiques du corpus de conception collecté

Il convient de considérer les résultats à la lumière des limites des verbalisations think-aloud. Comme l'ont souligné Nisbett et Wilson (1977), ces protocoles permettent d'accéder uniquement aux intentions conscientes, et certains processus cognitifs demeurent implicites. Le modèle de *machine learning* analyse donc principalement les intentions explicitement formulées, et non l'ensemble des processus cognitifs sous-jacents. Le choix d'utiliser un protocole *think-aloud* dans cette étude se justifie par l'objectif spécifique d'observer et de classifier systématiquement les intentions individuelles lors de la génération de concepts. Mais cette approche méthodologique de recours au *machine learning* pourrait tout à fait être appliquée à des corpus d'interaction issus de réunions de projet ou de sessions de conception collaborative, pour capturer des interactions naturelles et révéler des motifs cognitifs complémentaires.

Implications pour la recherche et la pratique

Ces résultats ouvrent plusieurs perspectives. Tout d'abord, la combinaison de modèles multimodaux (texte et croquis) pourrait améliorer la capture de dimensions moins verbalisées comme la forme. Ensuite, l'utilisation de modèles ML pour l'analyse automatique des verbalisations pourrait réduire la charge de travail dans les études de cognition en conception, tout en fournissant des insights reproductibles et systématiques. De plus, la détection de motifs sémantiques dans les verbalisations peut informer le développement d'outils d'aide à la conception, capables de proposer des suggestions contextuellement pertinentes selon l'intention exprimée par le concepteur. Enfin, la variabilité inter-participants souligne que la méthodologie doit être flexible, et que l'interprétation humaine reste essentielle pour traiter les cas ambigus ou multi-dimensionnels.

Conclusion

Cette étude a exploré la faisabilité d'utiliser le *machine learning* pour interpréter l'intention de conception à partir des verbalisations produites par les concepteurs dans un contexte multimodal de génération de concepts.

Nos résultats montrent que le *machine learning* peut interpréter de manière fiable l'intention de conception à partir de verbalisations, en particulier pour les dimensions fonctionnelles et de matérialité. Les écarts observés pour la forme mettent en lumière la complexité et l'ambiguïté des intentions de conception, ainsi que les limites des seuls observables verbaux. Ces résultats suggèrent que l'apprentissage automatique peut compléter les méthodes *think-aloud* traditionnelles, en offrant un outil scalable pour analyser la cognition en conception. Enfin, ce travail ouvre des perspectives pour le développement d'approches multimodales et multi-étiquettes capables de capturer la richesse dynamique des processus créatifs des concepteurs.

Référencement

- Cross, N., Christiaans, H., & Dorst, K. (Eds.). (1996). *Analysing design activity*. Wiley.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North* (pp. 4171–4186). Association for Computational Linguistics.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). The MIT Press.
- Ganeshan, R., Garrett, J., & Finger, S. (1991). Representing and reasoning with design intent. In *Artificial Intelligence in Design '91* (pp. 737–755). Elsevier.

- Hay, L., Cash, P., & McKilligan, S. (2020). The future of design cognition analysis. *Design Science*, 6, e20.
- Hay, L., Duffy, A. H. B., McTeague, C., Pidgeon, L. M., Vuletic, T., & Greal, M. (2017). A systematic review of protocol studies on conceptual design cognition: Design as search and exploration. *Design Science*, 3, e10.
- Henderson, M. R. (1993). Representing functionality and design intent in product models. In *Proceedings on the second ACM symposium on Solid modeling and applications - SMA '93* (pp. 387–396). ACM Press.
- Liang, Y., Liu, Y., Chen, C., & Jiang, Z. (2018). Extracting topic-sensitive content from textual documents—A hybrid topic model approach. *Engineering Applications of Artificial Intelligence*, 70, 81–91.
- Moran, T. P., & Carroll, J. M. (1996). *Design rationale: Concepts, techniques, and use*. CRC Press.
- Nielsen, C. K. E. B. B., Daalhuizen, J., & Cash, P. J. (2021). Defining the behavioural design space. *International Journal of Design*, 15(1), 1.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Otey, J., Company, P., Contero, M., & Camba, J. (2014). A review of the design intent concept in the context of CAD model quality metrics. In *2014 ASEE Annual Conference & Exposition Proceedings* (p. 24.100.1–24.100.10). ASEE Conferences.
- Schön, D. A. (1992). Designing as reflective conversation with the materials of a design situation. *Knowledge-Based Systems*, 5(1), 3–14.
- Siddharth, L., Blessing, L., & Luo, J. (2022). Natural language processing in-and-for design research. *Design Science*, 8, e21.

• • •