



HAL
open science

Un assistant virtuel dédié à l'accompagnement de la personne, averti de son contexte de vie, grâce au smartphone

Florian Szczepaniak, Jérôme Boudy, Gérard Chollet, Mossaab Hariz,
Christophe Lohr

► To cite this version:

Florian Szczepaniak, Jérôme Boudy, Gérard Chollet, Mossaab Hariz, Christophe Lohr. Un assistant virtuel dédié à l'accompagnement de la personne, averti de son contexte de vie, grâce au smartphone. 9ème Colloque sur la Télésanté: "Dispositifs biomédicaux et technologies numériques en santé: des besoins aux usages" (JETSAN 2023), Université Paris 8; Sorbonne Université, Jun 2023, Paris, France. hal-04181543

HAL Id: hal-04181543

<https://hal.science/hal-04181543v1>

Submitted on 16 Aug 2023

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.



Un assistant virtuel dédié à l'accompagnement de la personne, averti de son contexte de vie, grâce au smartphone

Florian Szczepaniak¹, Jérôme Boudy¹, Gérard Chollet¹,
Mossaab Hariz¹, Christophe Lohr²

¹ SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France

² IMT Atlantique Bretagne-Pays de la Loire, 29238 Brest, France

Correspondance: florian.szczepaniak@telecom-sudparis.eu

Abstract

Les publications des dernières années décrivent une forte utilisation des systèmes portés (en particulier des smartphones) dans le domaine de la reconnaissance de l'activité humaine (HAR). Le smartphone peut être considéré à la fois comme un ensemble de capteurs et comme une interface pour l'utilisateur. Le projet européen Horizon 2020 e-VITA a pour but de promouvoir le vieillissement actif chez les populations séniors. A ces fins, et grâce aux Technologies de l'Information et de la Communication (TIC), les travaux actuels sont concentrés sur le développement d'un assistant virtuel permettant d'aider l'utilisateur dans ses tâches quotidiennes. L'une des TIC utilisée dans ce projet est le smartphone. En effet, les capteurs du smartphone permettent de mesurer et de reconstituer une partie du contexte de l'utilisateur, grâce à la communication avec un serveur distant. En outre, ce serveur est capable d'offrir différents services (dont la fusion des différentes modalités) afin de reconstituer le contexte. La fusion de données multimodales implique à son tour différents modules de machine learning. Enfin, afin de restituer efficacement le contexte à l'utilisateur et de dialoguer avec lui, un module de chatbot sous le modèle chatGPT développé par openAI y est intégré. La communication avec l'utilisateur sera alors établie de manière textuelle ou vocale.

Keywords— e-Health, Devices, Signal Processing, Data Fusion, Machine Learning

1 Introduction

Le contexte actuel du développement de l'IoT [1] (Internet of Things) rend la centralisation et l'enrichissement de ces sources de données cruciales dans de nombreux domaines, en particulier dans celui de la e-Santé. Cette tendance se confirme aussi lorsqu'il s'agit d'utiliser les smartphones en

tant que capteur, plus précisément jeu de capteurs [2]. En effet, les smartphones sont équipés de multiples capteurs embarqués (GPS, accéléromètres 3D, magnétomètre, gyroscope, microphone, etc.): il s'agit d'un outil entièrement programmable et communicant. De plus, son utilisation s'est fortement démocratisée, ce qui facilite la mise en service

de solutions basées sur ce système. Les travaux de V.P Cornet et R.J Holden [2] renseignent aussi sur le fait que les études portant sur l'utilisation des smartphones sont réalisées dans des échantillons de populations diverses (âge), avec des conditions identifiées (pathologies neurodégénératives ou autisme). Ainsi, la diversité des cas d'usage et les fonctionnalités natives du smartphone font de cet appareil un excellent capteur IoT pour la e-Santé.

Dans le cadre du programme Horizon 2020, le projet Européen et Japonais e-VITA a pour but de promouvoir le vieillissement actif chez les populations seniors. Ce consortium interculturel et pluridisciplinaire développe un assistant (ou coach virtuel) ayant pour but d'aider un utilisateur dans ses activités de la vie quotidienne (AVQ). Physiquement intégré à l'environnement de la personne, le coach virtuel bénéficie d'un réseau de capteurs IoT permettant de contextualiser les activités de l'utilisateur. Pour y parvenir, ce dernier aura le choix de s'équiper et d'équiper son habitation de multiples capteurs, tels que des détecteurs de présence infrarouge (DeltaDore ou EnOcean), des détecteurs d'ouverture de portes (DeltaDore ou EnOcean), un robot compagnon (Nao), un hologramme (Gatebox), des thermomètres connectés (NetAtmo), etc. De plus, pour assurer une disponibilité à chaque instant de l'assistant, le smartphone de l'utilisateur pourra s'ajouter à la liste des capteurs IoT présentés plus haut. Cet ensemble de capteurs sera par la suite désigné comme l'écosystème IoT du projet e-VITA.

L'initiative de ce papier s'inscrit dans l'écosystème du projet e-VITA. À ce stade de l'expérience, seul le smartphone est utilisé comme capteur IoT. La contextualisation de l'activité de l'utilisateur repose sur une quantification du contexte (grâce à l'écosystème IoT) qui sera ensuite qualifiée (introduction de l'aspect sémantique). Cette démarche s'inscrit dans le domaine de la recherche sur la fusion de données qui

peut se définir comme suit: «*La fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses. Elle vise à l'obtention d'information de plus grande qualité; la définition exacte de plus grande qualité dépendra de l'application.*» [3]

De nature numérique, le cadre formel peut être soit architectural soit algorithmique, tel que conceptualisé dans l'article [4]. La différence principale entre ces deux approches se trouve au niveau de la couche de centralisation de l'information imposée par l'approche architectural, ainsi que par un déploiement orienté de type Edge- ou Cloud-Computing. L'approche architecturale permet, entre autres, une gestion des flux de manière asynchrone et de ce fait, l'articulation d'un ensemble d'approches algorithmiques coordonnées par la couche de centralisation. L'application est ici un assistant vocal ayant une connaissance sur le contexte de l'utilisateur. La représentation sémantique du contexte de vie de l'utilisateur est l'information valorisée résultant de la fusion de données.

Les récents travaux sur les Transformers [5], en particulier pour la génération automatique de texte et les agents conversationnels [6], sont de bons candidats pour la génération du contexte de vie. Afin d'obtenir un contexte de vie *de plus grande qualité*, l'approche architecturale suppose qu'il existe des sous-informations pouvant l'augmenter. Ainsi, l'étude des différents flux d'information du smartphone donne lieu à la détermination de ces sous-contextes. L'hypothèse est alors qu'un nombre plus important de sous-contextes, ou avec une valorisation plus importante, résulte un contexte de vie de plus grande qualité. Le mécanisme client-serveur, avec une application dédiée sur smartphone (appelée *MyADL*) et une architecture informatique dédiée *Data Fusion Platform* (DFP), est l'approche qui a été retenue.

2 Prototypage

2.1 Ressources et compatibilité

Le mécanisme présenté dans les prochaines sections fonctionne grâce à un smartphone et un serveur. Pour le smartphone, il s'agit d'un **Google - Pixel 6 Pro**. Pour le serveur, il s'agit d'un **Dell - Precision 7920 Tower**, Intel Xeon Gold 6230R x2 4GHz, mémoire vive 64 GB et NVIDIA Quadro RTX 8000. À noter que le dimensionnement du serveur peut différer de celui utilisé ici, en fonction du nombre d'utilisateurs et de la complexité des traitements d'analyse mis en œuvre. Pour les systèmes sous Linux, avec une complexité relative (maximum trois services nécessitant plus de 2 GB par utilisation est un indicateur correct) et une dizaine d'utilisateurs, 16 GB de RAM, 8 cœurs ~3GHz sont suffisants. Enfin, bien que le smartphone choisi pour l'expérience soit sous Android, il est possible d'utiliser l'application avec un smartphone sous iOS. En effet, le développement de cette dernière a été réalisé avec le framework **Flutter** qui assure une compatibilité entre les différentes plateformes.

2.2 Smartphone

L'application dédiée *MyADL* transmet de manière continue de l'information au *DFP* et permet à l'utilisateur de dialoguer avec l'assistant.

La transmission concerne les signaux inertiels (accéléromètre, magnétomètre, gyroscope), le nombre de pas et les coordonnées GPS. Ces données permettent, par des mécanismes de calcul et d'agrégation, de déterminer le contexte de vie de l'utilisateur. De plus, la transmission de ces modalités s'effectue en temps réel. Le temps réel étant défini comme le traitement d'une grandeur dans un temps plus court que son évolution. Le premier sous-contexte prototypé s'intéresse alors à l'activité physique et la marche de l'utilisateur grâce aux signaux inertiels [7]. L'étude de ce con-

texte en temps réel impose certaines spécifications sur le *DFP* (voir section 2.3). À noter que pour assurer le dialogue, la transmission doit être bidirectionnelle.

Le dialogue entre l'assistant et l'utilisateur est, soit textuel, soit vocal. Afin d'être facilement utilisable, les fonctionnalités *Speech To Text* (STT) et *Text To Speech* (TTS) ont été ajoutées. Plusieurs solutions orientées SDK ou des bibliothèques peuvent être utilisées pour la partie STT, **TakeNote**, **Flutter STT** ou encore le STT natif des smartphones. Ce constat est aussi valable pour la partie TTS. La solution retenue est d'utiliser le STT natif du smartphone combiné à la bibliothèque **Flutter TTS** afin d'assurer le dialogue avec l'utilisateur. Les différents visuels sont illustrés en figure 4.

2.3 Data Fusion Platform

Protocole de communication La première étape consiste à déterminer une méthode de communication permettant de suivre l'évolution des modalités pouvant atteindre une fréquence de mise à jour de l'ordre de la dizaine de Hertz. En effet, il est nécessaire d'obtenir une analyse en temps réel de la démarche. Trois protocoles de communication client-serveur peuvent être mis en concurrence: HTTP/1.1, HTTP/2 et WebSocket. Durant la phase expérimentale, la communication via HTTP/1.1 a été rapidement écartée, d'une part à cause d'une forte consommation de ressources sur le système, d'autre part à cause de la nature unidirectionnelle de la communication. Bien que les performances de HTTP/2 semblaient intéressantes, en particulier avec le framework gRPC, la communication via WebSockets a été choisie en première approche afin de faciliter le développement. À noter que les WebSockets et HTTP/2 supportent la communication bidirectionnelle. Cette dernière peut permettre au système de facilement notifier l'utilisateur afin de rendre l'interaction dynamique.

Collecter, valoriser et historiser La seconde étape consiste à collecter, valoriser et faire un historique des différentes modalités afin de préparer la création du contexte. Les différentes modalités du smartphone (signaux inertiels, GPS et nombre de pas) sont transmises via la communication WebSocket. Les coordonnées GPS et le nombre de pas sont envoyés au *context broker* Fiware Orion [8] (équivalent de la couche de centralisation dans l'approche algorithmique). Pour suivre l'évolution des signaux inertiels et évaluer la démarche de l'utilisateur, les données sont transmises au *message broker* Apache Kafka [9] pour une consommation par le service lié à la détermination de la démarche du *DFP*. Pour respecter la centralisation, ce service de machine learning transmet la démarche de l'utilisateur à Orion, après calcul. Les entités courantes présentent sur Orion voient leurs états stockés dans un historique grâce à l'intégration de Fiware Cygnus.

Agrégation et sémantique La troisième étape consiste en une agrégation des différentes modalités obtenues puis de les transformer sous une forme dite sémantique. L'agrégation est basée sur une période temporelle variable qui peut être ajustée pour s'adapter à l'utilisateur. Le mécanisme repose, en partie, sur Fiware STH-Comet permettant de requêter les séries temporelles, les données sous forme d'historique, créés par Fiware Cygnus. Ainsi, le *DFP* fonctionne largement sous le standard Fiware; de fait, la communication entre les services du *DFP* et l'accès aux données sont standardisés selon la norme NGSi-v2. La figure 1 propose une représentation de l'entité d'agrégation. La forme sémantique, quant à elle, est le résultat de la création d'un texte procédural à partir des différents champs à disposition dans l'entité agrégée, une représentation de l'entité de la forme sémantique est proposée en figure 2.

```
{
  "id": "urn:ngsi-v2:ContextAggregation:0000-0000-0000",
  "type": "ContextAggregation",
  "username": {"type": "Str", "value": "Albert"},
  "age": {"type": "Int", "value": 70},
  "delta_hours": {"type": "Int", "value": 24},
  "avg_steps_count": {"type": "Int", "value": 2000},
  "avg_gait": {"type": "Dict", "value": {"sit": 0.3, "lay": 0.65, "walk": 0.05}},
  "move_outside_home": {"type": "Int", "value": 3},
  "avg_time_outside_home": {"type": "Float", "value": 0.3}
}
```

Figure 1: Exemple d'une entité d'agrégation

```
{
  "id": "urn:ngsi-v2:ContextAggregationSemantic:0000-0000-0000",
  "type": "ContextAggregationSemantic",
  "description": {"type": "Str", "value": "The following is a description of living habits. You are a fair and honest wellness coach and you have to write a report about the following facts. The person is named Albert and he is 70 years old. Since last 24 hours he has spent 72 minutes being physically active. Since last 24 hours his most common position has been to lay for 15.6 hours. Since the last 24 hours he has made 6000 steps. Since last 24 hours he has left his habitation for 30% of his time. Since last 24 hours he has left and returned to his habitation 3 times."}
}
```

Figure 2: Exemple d'une entité sémantique

Générer le contexte de l'utilisateur La quatrième et dernière étape consiste à valoriser l'entité précédente (figure 2). Le processus de traitement de cette entité est réalisé par des transformers, en particulier les **Generative Pre-trained Transformers** (GPT). Actuellement, il s'agit du modèle GPT-3 d'openAI [10]. Pour accélérer la phase de développement, le *DFP* intègre un module de requêtes à l'API ChatGPT d'openAI. Néanmoins, cette solution s'inscrit dans l'utilisation de solution du cloud computing, ce qui n'est pas souhaitable au regard des considérations RGPD (voir [11]).

```

{
  "id": "urn:ngsi-v2:ContextAwareness:0000-0000-0000",
  "type": "ContextAwareness",
  "characteristic": "Albert, a 70-year-old man, has been living a fairly sedentary lifestyle in the last 24 hours. He has spent 72 minutes being physically active, 15.6 hours laying down, and only 5% of his time outside of his habitation. He has taken 2000 steps and left and returned to his habitation 3 times. As his wellness coach, it is important to note that Albert's current lifestyle is not conducive to good health. He needs to increase his physical activity and reduce the amount of time he spends laying down. Additionally, he should aim to spend more time outside of his habitation, as this can help to improve his overall wellbeing. \n\nIt is recommended that Albert increase his physical"
}

```

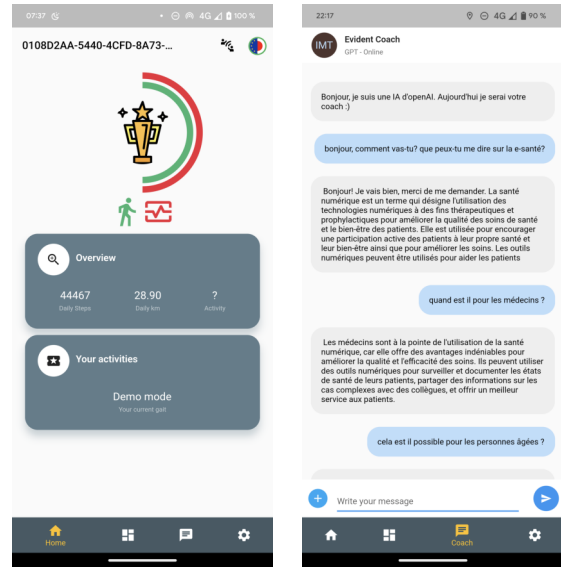
Figure 3: Exemple d'une entité de contexte averti

La nouvelle entité est alors une caractéristique du contexte de vie de l'utilisateur sur la période d'agrégation choisie au préalable. Cette caractéristique sera utilisée dans le chatbot (produit par le modèle GPT-3) afin d'apporter le soutien le plus pertinent à l'utilisateur. Enfin, c'est cette dernière qui constitue l'étape finale de la fusion de données multimodales, permettant de passer d'une approche quantitative (données issues des capteurs et de la fusion des données des capteurs) à une approche qualitative sur le contexte de vie de l'utilisateur.

3 Résultats et Analyse

3.1 Analyse de la démarche

Tous les apprentissages de modèles réalisés dans cette section utilisent la base de données KU-HAR [12]. Cette dernière a été retravaillée afin de correspondre à l'usage désiré et se limite alors à 5 labels considérés comme suffisants (allongé, courir, assis, debout, marcher) pour la détermination



Page d'accueil

Chatbot

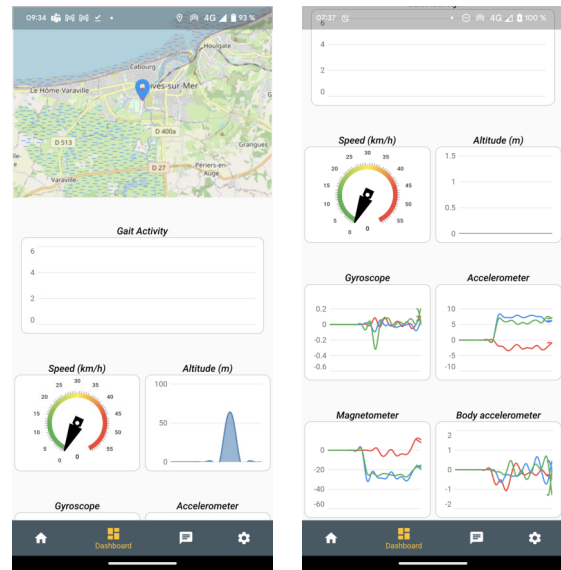


Tableau de bord

Figure 4: Visuels de MyADL

de la démarche en temps réel. Les modèles de deep learning (Multi Layer Perceptron - MLP) ont été développés avec Keras & Tensorflow et les forêts aléatoires (Random Forest, RF) ont été développées avec les bibliothèques du projet NVIDIA Rapids. Ainsi, indépendamment du modèle retenu,

Modèle	Précision	Temps exécution	Utilisateurs
RF-100	91.1	1.734 s	≥ 8000
RF-300	91.3	2.612 s	≥ 5000
RF-600	91.4	4.772 s	≥ 2000
MLP-50	94.2	18.413 s	≥ 800

Figure 5: Performances machine learning

il est possible d'avoir une exécution du code sur GPU.

Les résultats sont présentés en figure 5. À noter que le nombre d'utilisateurs est calculé à partir de la fréquence de calcul général ($f_{gen} = \frac{437269}{temps\ execution}$), de la fréquence d'envoi ($f_{envoi} = 10$ Hz), du temps alloué à la détermination de l'activité dans le mécanisme de traitement du streaming ($t_{calc} = 0.05$ s) et du nombre de données par batch ($data_{batch} = f_{envoi} \cdot batch_{time}$, avec $batch_{time} = 0.3$ s). Ainsi le nombre d'utilisateurs est donné par: $n = \frac{f_{gen} \cdot t_{calc}}{data_{batch}}$

Au regard des temps d'exécution, les RF proposent un moyen efficace de déterminer la démarche de l'utilisateur de manière continue, même pour un grand nombre d'utilisateurs. Néanmoins, il est clair que l'approche MLP semble être plus efficace en termes de performance, mais pour 3 points de performance en plus le temps d'exécution est multiplié par 10. Enfin, l'augmentation du nombre d'estimateurs n'augmente que très légèrement la précision au détriment d'un temps d'exécution qui ne cesse d'augmenter. En somme, le modèle plus pertinent pour l'analyse en temps de réel de la démarche est l'approche RF avec 100 estimateurs.

Néanmoins, le modèle n'est que l'étape de valorisation dans le processus d'analyse de la démarche. Comme précisé en section 2.3 l'information doit être collectée, valorisée et mise sous forme d'historique. La collecte est réalisée par l'envoi d'information du smartphone via le canal WebSocket au service de collecte du DFP. Ce service accepte un nombre fini de connexions, en fonction de la charge des connexions établies. Au re-

gard des différentes données transmises, seuls les signaux inertiels peuvent être considérés. Ainsi, pour un enregistrement de 2 secondes des signaux inertiels (accéléromètre, gyroscope), avec une fréquence d'échantillonnage de 100 Hz, transmis à une fréquence de 10 Hz, le taux d'occupation du service de WebSocket est de l'ordre de 10%, soit 10 utilisateurs par service. En termes de taille de paquet échangé, c'est environ 200 kB qui sont échangés toutes les 0.1 s. La latence du système mériterait une étude précise, mais 1 seconde est une borne supérieure, qui nous semble correcte. La fréquence du signal de la démarche de l'utilisateur est de l'ordre de la dizaine de hertz.

3.2 Détermination des contextes

La détermination des contextes est le résultat d'un processus en trois temps. Tout d'abord, il faut procéder à l'agrégation des différentes données d'intérêts durant une période définie. Les données d'intérêts se définissent comme les données à valeur ajoutée pour la détermination du contexte. Il s'agit du suivi de la démarche, du nombre de pas, de ses déplacements, représentés en figure 1, à cela est ajouté des informations telles que le prénom et l'âge. Puis en listant les différents attributs, la forme sémantique de cette entité est générée, représenté en figure 2. Enfin, ce texte est transmis à l'API du chatGPT afin d'obtenir une représentation du contexte de vie de l'utilisateur, représenté en figure 3.

Afin de valider l'hypothèse de départ, la méthode consiste à comparer les contextes issus de différentes formes sémantiques. Les deux formes sémantiques comparées sont celle de la figure 2 et la même dont les informations sur la démarche sont supprimées. Dans le premier cas (figure 3), l'utilisateur avait un contexte de vie plutôt défavorable pour sa santé alors que dans le second cas (6), ce dernier semble plutôt favorable. En effet, les informations liées à la démarche mon-

```
{
  "id": "urn:ngsi-v2:ContextAwareness:0000-0000-0001",
  "type": "ContextAwareness",
  "characteristic": "Albert, a 70-year-old man, has been living a healthy lifestyle over the past 24 hours. He has taken 2000 steps and has left his habitation for 5% of his time. He has left and returned to his habitation 3 times in the last 24 hours. Overall, Albert is making good progress towards a healthier lifestyle. He is taking an adequate amount of steps and is getting out of his habitation regularly. However, he could benefit from increasing the amount of time he spends outside of his habitation. This would help him to get more exercise and fresh air, which can help to improve his overall health and wellbeing."
}
```

Figure 6: Exemple d'une entité de contexte averti

trent que l'utilisateur est le plus souvent allongé et n'alloue que peu de temps à avoir une démarche active (marcher, courir, etc.). Ainsi, le contexte de vie de l'utilisateur est plus proche de la réalité grâce à l'utilisation de sous-contexte comme celui de la démarche. Enfin, cette caractérisation est intégrée dans le service relatif au dialogue afin que le chatbot soit renseigné par ces informations.

3.3 Assistant vocal - MyADL

L'assistant vocal est une fonctionnalité de l'application MyADL. Cette application fonctionne conjointement avec le *DFP*. Cette dernière profite du caractère ambivalent du smartphone, capteur / interface homme - machine (IHM). En effet, comme décrit précédemment, l'objectif est d'abord de transmettre de l'information. Toutefois, la création d'une IHM permet une incitation à l'utilisation du système complet et donc un accompagnement de l'assistant plus présent. Ainsi, *MyADL* est composé (voir figure 4) d'une page d'accueil résumant les informations importantes, d'un dashboard pour assurer la transparence entre les données collectées et transmises, et le module de

chatbot pour l'assistant. Bien que la qualité de la synthèse vocale du module TTS soit souvent monocorde, la conversation verbale ou textuelle avec le chatbot est fonctionnelle.

4 Conclusion

En conclusion, l'approche architecturale de la fusion de données ayant comme source de données un smartphone hébergeant de multiples capteurs actimétriques est un moyen d'identifier le contexte de vie de l'utilisateur. L'enrichissement des données, par le biais de l'analyse de la démarche, montre que l'étude des sous-contextes est une méthode fonctionnelle afin d'améliorer la qualité de la génération de ce contexte de vie. Inspiré par le paradigme de la lambda architecture [13], le *DFP* est capable de suivre l'utilisateur dans ses activités en temps réel et d'en agréger les résultats sur des périodes temporelles. Enfin, le support STT et TTS de *MyADL* permet un dialogue avec l'assistant et offre un ensemble de fonctionnalités à l'utilisateur. Ainsi, grâce au smartphone, il est possible de suivre l'activité avec comme perspectives la détermination du contexte de l'utilisateur tout en étant un assistant vocal dédié à la personne.

Pour l'étude relative au contexte ou encore pour le chatbot, il s'agit de l'API de chat-GPT. L'utilisation de transformers intégré *DFP* pourrait permettre une meilleure gestion des problématiques RGPD. De plus, le fine tuning [14] (réglages fins des modèles afin de se spécialiser) sur des modèles comme GPT-2 [15] pourrait donner lieu à la génération de meilleurs contextes de vie et à une plateforme orientée edge computing. Toujours dans cette optique d'amélioration et en considérant l'écosystème IoT du projet e-VITA, nombre de capteurs comme ceux intégrés à l'habitation peuvent être de nouvelles sources donnant lieu à de nouveaux sous-contextes à étudier.

Conflits d'intérêts: Les auteurs ne déclarent aucun conflit d'intérêt.

Data Availability Statement: KU-HAR: An open dataset for heterogeneous human activity recognition [12]

References

- [1] A. Ometov, V. Shubina, L. Klus, J. Skibińska, S. Saafi, P. Pascacio, L. Flueraștoru, D. Quezada-Gaibor, N. Chukhno, O. Chukhno, A. Ali, A. Channa, E. Svertoka, W. Bin Qaim, R. Casanova-Marqués, S. Holcer, J. Torres-Sospedra, S. Casteleyn, G. Ruggeri, and E. S. Lohan, “A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges,” *Computer Networks*, p. 108074, Apr. 2021.
- [2] V. P. Cornet and R. J. Holden, “Systematic review of smartphone-based passive sensing for health and wellbeing,” *Journal of Biomedical Informatics*, vol. 77, pp. 120–132, Jan. 2018.
- [3] L. Wald, “DEFINITIONS AND TERMS OF REFERENCE IN DATA FUSION,” *International Archives of Photogrammetry and Remote Sensing*, vol. 32, 1999.
- [4] M. McTear, K. Jokinen, M. M. Alam, Q. Saleem, G. Napolitano, F. Szczepaniak, M. Hariz, G. Chollet, C. Lohr, J. Boudy, Z. Azimi, S. D. Roelen, and R. Wieching, “Interaction with a Virtual Coach for Active and Healthy Ageing,” *Sensors*, vol. 23, p. 2748, Jan. 2023. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- [5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention Is All You Need,” Dec. 2017. arXiv:1706.03762 [cs].
- [6] G. Caldarini, S. Jaf, and K. McGarry, “A Literature Survey of Recent Advances in Chatbots,” *Information*, vol. 13, p. 41, Jan. 2022. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
- [7] A. K. Panja, A. Rayala, A. Agarwala, S. Negoy, and C. Chowdhury, “A hybrid tuple selection pipeline for smartphone based Human Activity Recognition,” *Expert Systems with Applications*, vol. 217, p. 119536, May 2023.
- [8] A. Hassani, A. Medvedev, A. Zaslavsky, P. Delir Haghighi, P. P. Jayaraman, and S. Ling, “Efficient Execution of Complex Context Queries to Enable Near Real-Time Smart IoT Applications,” *Sensors (Basel, Switzerland)*, vol. 19, p. 5457, Dec. 2019.
- [9] R. Wiatr, R. Słota, and J. Kitowski, “Optimising Kafka for stream processing in latency sensitive systems,” *Procedia Computer Science*, vol. 136, pp. 99–108, Jan. 2018.
- [10] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, “Language Models are Few-Shot Learners,” July 2020. arXiv:2005.14165 [cs].
- [11] B. L, “ChatGPT et vie privée : ne confiez surtout pas vos secrets à l’IA ! Voici pourquoi.” Section: Intelligence artificielle.
- [12] N. Sikder and A.-A. Nahid, “KU-HAR: An open dataset for heterogeneous human activity recognition,” *Pattern Recognition Letters*, vol. 146, pp. 46–54, June 2021.
- [13] R. Alghamdi and M. Bellaïche, “An ensemble deep learning based IDS for IoT using Lambda architecture,” *Cybersecurity*, vol. 6, p. 5, Mar. 2023.
- [14] J.-S. Lee and J. Hsiang, *Patent Claim Generation by Fine-Tuning OpenAI GPT-2*. June 2019.
- [15] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training,”