

Une IA pour assister la gestion des déchets dans les bâtiments

Yassine Bouabdallaoui¹, Laure Ducoulombier², Pascal Yim³, Zoubeir Lafhaj^{4*}

¹ Centrale Lille, CNRS, Univ. Lille, UMR 9013 - LaMcube - Laboratoire de Mécanique, Multiphysique, Multi-échelle, F-59000 Lille, France; yassine.bouabdallaoui@centralelille.fr

² Bouygues Construction, 78280 Guyancourt, France; L.DUCOULOMBIER@bouygues-construction.com

³ Centrale Lille, Univ. Lille, CNRS, UMR 9189 - CRISTAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France; pascal.yim@centralelille.fr

⁴ Centrale Lille, CNRS, Univ. Lille, UMR 9013 - LaMcube - Laboratoire de Mécanique, Multiphysique, Multi-échelle, F-59000 Lille, France; zoubeir.lafhaj@centralelille.fr

RESUME Cette étude souligne l'importance et les défis de la durabilité et de la gestion des déchets, en s'appuyant sur le principe des 3R (Réduction, Recyclage, Réutilisation), dans le cadre de l'exploitation et de la maintenance des bâtiments. Elle présente le développement d'un assistant virtuel, AsTriD, qui facilite le tri des déchets grâce à la reconnaissance vocale et d'image. L'évaluation technique et de convivialité montre que 70% des utilisateurs sont très satisfaits. Néanmoins, plusieurs suggestions d'amélioration sont avancées, constituant une base solide pour une utilisation plus approfondie des assistants virtuels dans ce domaine.

Mots-clefs intelligence artificielle, durabilité, bâtiments, déchets, circularité

I. INTRODUCTION

La maintenance et l'exploitation durable a été identifiée comme un élément clé pour l'avenir des bâtiments (Nielsen et al., 2016). Ceci englobe de multiples aspects, notamment l'efficacité énergétique et l'économie en eau, la gestion des déchets, la qualité de l'air intérieur, la sécurité et le confort des occupants, la conception durable et l'engagement communautaire (Alfalah and Zayed, 2020). L'adoption de pratiques de durabilité dans l'exploitation des bâtiments est de plus en plus répandue. La réduction, le recyclage et la réutilisation, connus sous le nom de principe des 3R, sont à l'origine de solutions de gestion durable des déchets dans le monde entier (Goyal et al., 2018).

Pour établir le principe des 3R, l'organisation doit définir une politique de gestion efficace, mettre en place des systèmes de tri de bonne qualité et renforcer l'engagement personnel. En effet, une étude récente au Royaume-Uni montre qu'environ 90% des praticiens en exploitation-maintenance considèrent la gestion et le recyclage des déchets comme une question clé liée à leurs politiques de durabilité (Elmualim et al., 2012). La gestion des déchets est l'une des principales responsabilités des équipes, principalement dans le secteur tertiaire. Elles sont responsables de la définition des politiques de gestion des déchets, de la collecte et du tri des déchets dans le bâtiment avant qu'ils ne soient traités par les entrepreneurs locaux de la municipalité.

La gestion efficace des déchets dans divers établissements est un défi en raison de la variabilité des déchets solides générés, allant des bureaux aux commerces de détail, en passant par les établissements de santé et de loisirs. Le secteur tertiaire génère une quantité significative de déchets (Chateau and Lapillonne, 1982), par exemple, les bureaux en France produisent plus de 2,4 millions de tonnes de déchets par an, soit environ 150 kg par employé de bureau en moyenne. Cependant, le taux de recyclage de ces déchets est relativement faible, notamment pour le papier, où seulement 35% est recyclé. Chaque établissement est unique, ce qui nécessite des politiques de gestion des déchets adaptées.

II. Méthodologie de développement d'un assistant de tri de déchets Virtuel : AsTriD

Les assistants IA façonnent de plus en plus divers aspects de la société moderne, de l'automatisation résidentielle et du commerce électronique au service à la clientèle et au tourisme. Dans le domaine de l'exploitation et de la maintenance des bâtiments, le développement d'assistants virtuels IA offre un potentiel pour soutenir les politiques de recyclage, surveiller les objectifs de réduction des déchets et promouvoir les pratiques de réutilisation et de don. Les assistants virtuels, les agents intelligents et les chatbots sont des entités logicielles capables d'interagir avec les humains par la voix, le texte ou des interfaces visuelles (Hoy, 2018; Iannizzotto et al., 2018). Ils peuvent accomplir des tâches telles que répondre à des questions, gérer des appareils intelligents ou effectuer des actions spécifiques telles que la réservation de billets ou l'achat de billets. Ce domaine a connu une évolution significative depuis ELIZA en 1966 (Weizenbaum, 1966) jusqu'à SIRI en 2011, avec les techniques d'apprentissage automatique jouant un rôle central dans leur développement.

De la même manière, les progrès de la vision par ordinateur ont permis à l'utilisateur d'interagir avec un chatbot en utilisant des flux visuels et des images. Cette adoption va de l'utilisation individuelle et des appareils domestiques intelligents aux applications dans l'administration et les installations culturelles. L'objectif des travaux est de :

- Explorer l'utilisation d'assistants virtuels pour fournir des services de gestion des déchets dans les installations.
- Reconnaître et classer les déchets selon les règles de gestion des déchets de l'établissement.
- Fournir aux utilisateurs les bonnes instructions pour recycler, réutiliser ou donner leurs déchets.
- Ajoutez et personnalisez la liste des déchets pris en charge.
- Ajouter des instructions personnalisées et des informations utiles.
- Afficher et extraire les données de suivi.

Pour cela, un Assistant Virtuel, appelé AsTriD a été développé. Cet assistant aide un utilisateur à trier ses déchets. L'utilisateur peut demander à l'assistant virtuel de lui fournir des instructions en questionnant oralement l'assistant. Il peut également montrer à l'assistant les déchets dont il veut se débarrasser.

Pour développer cet Assistant, la première étape concerne la traduction des spécifications en fonctionnalités techniques. La seconde étape consiste à développer la solution. La dernière étape a porté sur l'évaluation et la validation de la solution. Deux niveaux d'évaluation technique et d'utilisabilité ont été abordés. L'évaluation technique a porté sur l'exactitude des instructions fournies par la solution. Au niveau de l'utilisabilité, l'évaluation porte sur la satisfaction de l'utilisateur.

Suite au processus d'évaluation, l'amélioration du modèle peut être évaluée sur l'application afin d'améliorer ses fonctionnalités. La conception et le développement de l'assistant virtuel ont été réalisés à l'aide de la méthodologie AGILE (Ahmed et al., 2010). AGILE est une méthode de gestion de projet qui est principalement utilisée dans le développement de logiciels. Elle implique un effort de collaboration continu entre les clients et les développeurs tout au long du projet.

Le choix de cette méthode a été motivé par le fait que le développement logiciel représente une part importante de ce travail et par la volonté d'intégrer l'utilisateur tout au long du projet. Plus précisément, le cadre Scrum a été utilisé dans le processus de développement. Scrum est un cadre AGILE bien connu dans la communauté du développement logiciel, mais il tire son origine d'un article de 1986 intitulé « The New New product development game ».

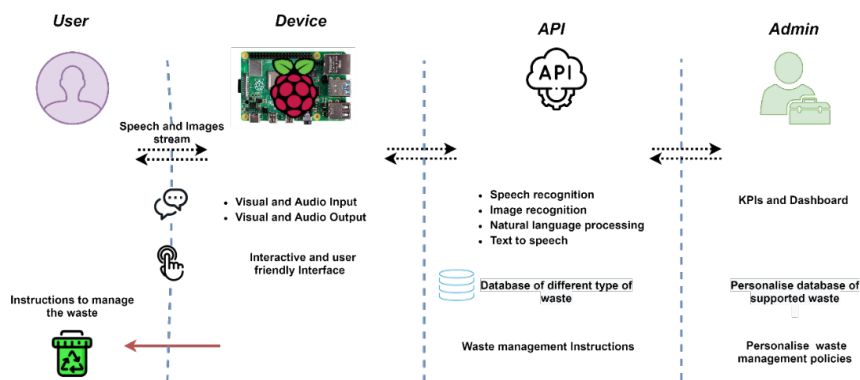


FIGURE 1. Architecture globale de l'Assistant Virtuel AsTriD

La vision par ordinateur est une discipline qui s'intéresse à l'extraction, à l'analyse et à l'interprétation de l'information contenue dans les images. La vision par ordinateur vise à générer une compréhension semblable à celle des humains à partir d'images ou de vidéos (Xu et al., 2019). La vision par ordinateur englobe un large éventail d'applications, allant de la détection et de la classification d'images aux applications de suivi des mouvements dans les flux vidéo jusqu'aux applications de reconstruction d'images. L'apprentissage automatique a pleinement participé à l'amélioration des applications de vision par ordinateur et de leur précision, notamment les algorithmes d'apprentissage profond, tels que les réseaux de neurones convolutifs.

Ces algorithmes ont atteint l'état de l'art dans les applications de reconnaissance d'objets en remportant le célèbre Large Scale Visual Recognition Challenge depuis l'année 2012. Ces dernières années, le « Natural Language Processing (NLP) » ou « Traitement du Language Naturel (TLN) » a connu un tournant avec le développement de l'apprentissage profond et des réseaux neuronaux artificiels. Les techniques d'apprentissage profond ont obtenu des résultats de pointe dans de nombreuses tâches de TLN. De nombreux secteurs utilisent les

applications du TLN dans leur industrie ou leur processus commercial. Cela inclut les assistants virtuels, qui deviennent une partie essentielle des services à la clientèle.

L'API est le back-end de l'Assistant Virtuel (AV). Elle gère les fonctionnalités de reconnaissance vocale, de reconnaissance d'images et de traitement du langage naturel pour classer le texte. L'API est hébergée dans un service « cloud » et est connectée à l'appareil final et à l'interface d'administration. L'interface d'administration permet au gestionnaire de l'établissement d'ajouter, de modifier et de personnaliser les consignes en fonction de l'établissement. L'interface d'administration fournit à l'administrateur les principales statistiques sur la gestion des déchets.

L'API est le cœur de l'AV, elle contient l'agent PLN et l'agent de vision par ordinateur qui traitent le texte et l'image en entrée. L'API est connectée aux API externes de Google Cloud Dialogflow et Cloud Speech. L'API est connectée au robot et à l'interface d'administration comme expliqué précédemment. Tous ces modules sont reliés par l'interface API qui gère les connexions externes et orchestre les agents d'IA.

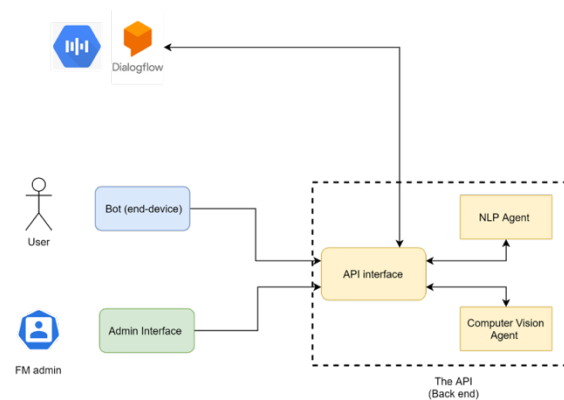


FIGURE 2. les composants de l'API AsTriD, Traitement des données textuelles et audio.

Le traitement et la compréhension des données textuelles et audio sont l'une des principales fonctionnalités de l'AV. L'AV doit être capable de transformer un fichier audio en données textuelles à l'aide de la reconnaissance vocale, puis de traiter le texte à l'aide de techniques NLP avant de générer une réponse à l'aide de la génération vocale. Certaines tâches, comme la génération et la reconnaissance vocale, nécessitent beaucoup de développement.

Dialogflow permet le développement de technologies d'interaction homme-machine basées sur la parole et le texte (Chien and Yao, 2020). La compréhension du langage naturel de Dialogflow fonctionne en utilisant les concepts d'intentions et d'entités. Dialogflow essaie de faire correspondre la demande de l'utilisateur avec la liste d'intentions définie en se basant sur la similarité entre l'expression et les phrases d'entraînement dans chaque intention. Le schéma fourni par la documentation de Google Dialogflow reprend le flux du processus de correspondance des intentions et des réponses.

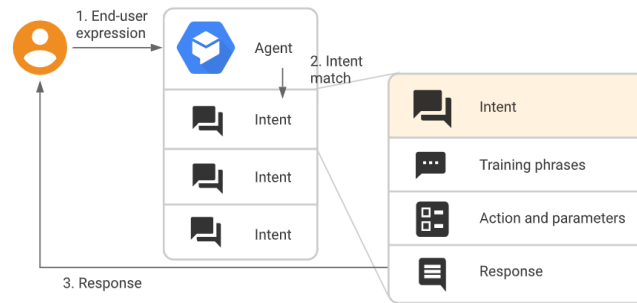


FIGURE 3. Aperçu du flux de correspondance des intentions et de réponse dans DialogFlow, extrait de la documentation Google Cloud.

Le processus de traitement des requêtes audio est illustré dans la figure 4. L'utilisateur émet une demande sous la forme d'un fichier audio en utilisant le microphone du robot. Le fichier audio est envoyé à l'API qui le normalise et le transfère à l'API Dialogflow.

L'API Dialogflow contient l'agent qui a été préformé pour détecter les intentions et les entités pour ce cas d'utilisation de la gestion des déchets. Ensuite, les intentions et entités détectées sont renvoyées à l'API, où la réponse de l'agent Dialogflow est envoyée au service de conversion de la parole en texte afin de synthétiser une réponse audio. La réponse audio est envoyée à l'API, puis au robot, et l'utilisateur reçoit la réponse via le haut-parleur du robot. Ce processus est répété tout au long de la conversation entre l'utilisateur et l'AV.

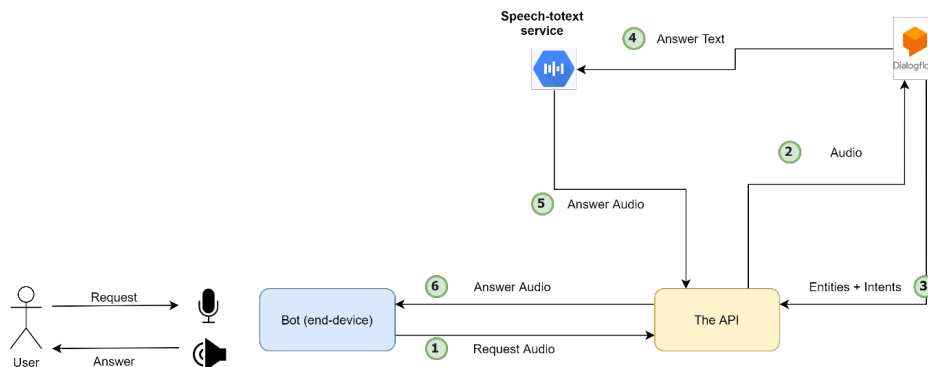


FIGURE 4. Aperçu du processus de traitement des requêtes audio.

Concernant le processus de reconnaissance d'image, le déclencheur audio est traité de la même manière que toute demande de texte. Ensuite, la caméra du robot est activée et un message audio est émis par le robot pour avertir l'utilisateur.

L'agent envoie le résultat de la classification à l'API qui utilise Dialogflow pour synthétiser la réponse audio basée sur le résultat de la classification de l'image en suivant le même processus que celui expliqué dans la section précédente. La fonctionnalité de classification des images a été développée sous la forme d'un script interne hébergé dans l'API de l'assistant virtuel. Inception est un réseau de neurones convolutifs qui a atteint les performances de l'état de l'art pour la classification et la détection lors du défi de reconnaissance visuelle à grande échelle ImageNet 2014 (Szegedy et al., 2015).

Afin de réentraîner le modèle à classer les déchets. Dans cet article, une première liste contenant 10 classes de déchets a été identifiée.

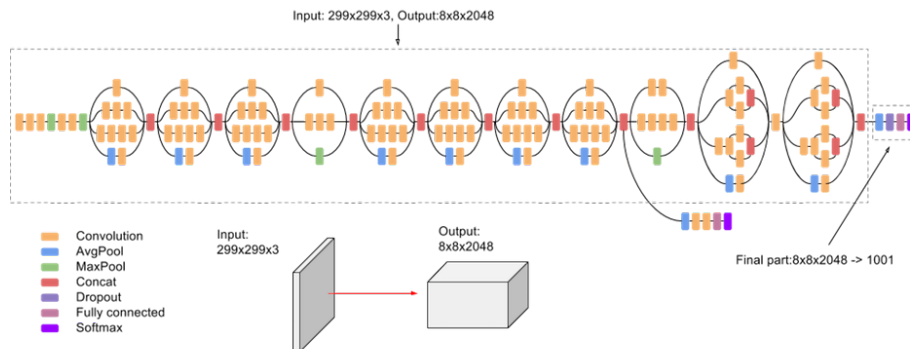


FIGURE 5. Illustration de l'architecture d'Inception V3 (Szegedy et al., 2016).

Afin de réentraîner le modèle pour classer les déchets dans ce document, une première liste contenant 10 classes de déchets a été identifiée. Les images de ces classes ont été collectées sur ImageNet et Google Image.

Contrairement à la fonctionnalité audio, la fonctionnalité de détection d'image prend en charge un nombre limité de classes et l'administrateur ne peut pas ajouter de nouvelles classes sans réentraîner le modèle de classification d'image. L'interface d'administration se présente sous la forme d'une plateforme web connectée à l'appareil final via l'API de l'assistant virtuel. Pour faciliter la tâche de l'administrateur, plusieurs scénarios intégrés sont proposés dans l'interface, notamment un scénario de tri classique, un scénario de déchets de bureau, un scénario de déchets industriels et un scénario de réutilisation et de don.

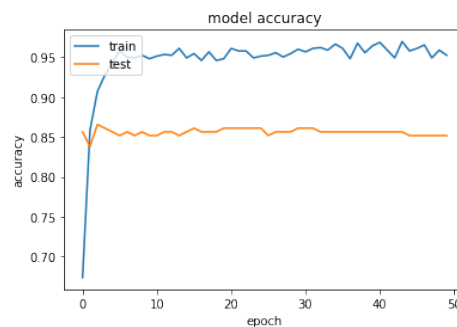


FIGURE 6. Précision de la classification.

Les catégories de déchets font référence au type de déchets pris en charge dans ce scénario. Chaque catégorie contient une liste de synonymes qui seront utilisés pour entraîner automatiquement l'API Dialogflow. Ces synonymes sont générés automatiquement à l'aide de l'agent NLP.

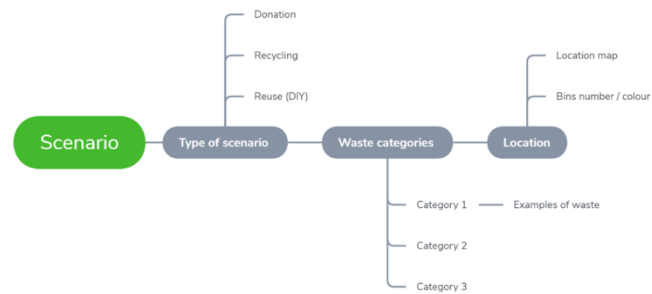


FIGURE 7. Composants du scénario sur les déchets.

Cette fonctionnalité est dédiée à l'administrateur via l'interface d'administration. Elle propose des synonymes potentiels que l'administrateur peut ajouter. Ce processus a été mis en place pour améliorer la précision de la solution Dialogflow. L'utilisation de synonymes permet d'entraîner l'API Dialogflow à la configuration la plus exhaustive possible.

Chaque fois que l'administrateur ajoute un nouveau type de déchet, cette fonctionnalité propose une liste de synonymes similaires au mot qu'il a utilisé. Le modèle de génération de synonymes a été construit sur la base de la bibliothèque FastText (Grave et al., 2019). Cette fonctionnalité est hébergée dans l'API. L'interface d'administration propose un tableau de bord de suivi qui permet à l'administrateur de contrôler la quantité de déchets générés dans l'établissement et la situation des poubelles.

Il peut également définir et surveiller les indicateurs de performance tels que les indicateurs de réduction des déchets. La possibilité de suivre les journaux de l'AV bot permet à l'administrateur de définir des règles pour surveiller et suivre l'évolution de l'élimination et du recyclage des déchets. En outre, il permet à l'utilisateur de définir et de suivre des indicateurs de réduction des déchets afin de motiver les utilisateurs de l'établissement à réduire la production de déchets.

III. Résultats et analyses

Le processus d'évaluation est une phase importante pour valider la solution ainsi que pour recueillir des informations sur la façon d'améliorer la solution. L'évaluation technique vise à valider l'exactitude des instructions fournies par la solution. L'évaluation technique a été réalisée pour évaluer à la fois la fonctionnalité audio et la fonctionnalité de reconnaissance d'image (Iannizzotto et al., 2018). Pour la fonctionnalité audio, le processus d'évaluation s'est inspiré des travaux d'une précédente étude sur les assistants virtuels. Une liste de tâches a été demandée au système à l'aide d'une entrée audio, chaque tâche a été réalisée 12 fois sur l'ensemble de l'expérience.

L'évaluation du modèle de classification d'images a été réalisée dans la section en suivant une mesure de précision de machine learning classique. Cependant, une autre évaluation était nécessaire pour évaluer le module lorsqu'il fonctionne comme un système complet. L'arrière-plan derrière les objets était un mur blanc pour simplifier l'évaluation. Selon le même concept que

l'évaluation audio, pour chaque objet, des mesures ont été collectées concernant le nombre de réponses entièrement réussies, le nombre d'échecs partiels et le nombre d'échecs.

L'évaluation de la convivialité ne visait que les utilisateurs finaux. Cependant, les administrateurs vivent des expériences différentes à travers l'interface d'administration, ils ont également eu un retour différent sur la convivialité de l'application. Les résultats de l'évaluation technique de l'AV montrent une précision globalement satisfaisante. L'utilisation de l'IA pour générer des synonymes afin d'entraîner le Dialogflow avec la liste la plus exhaustive a contribué à l'augmentation de la précision des demandes vocales.

Cependant, la mauvaise qualité de l'audio ou la présence de bruit peut affecter la précision des requêtes. Cette étude n'a pas couvert l'impact du matériel et de l'environnement sur la performance de l'AV. La fonctionnalité de reconnaissance d'images présente, comme prévu, une précision moyenne similaire à celle de l'évaluation des tests effectuée lors de la phase de développement. Comme pour la fonctionnalité vocale, l'environnement et le fond devant la caméra peuvent affecter la précision globale de la reconnaissance d'images.

Comme pour l'évaluation technique, les résultats de l'évaluation de la convivialité ont montré une satisfaction significative des utilisateurs. Environ 70% des répondants étaient « Très satisfaits », 15% « Satisfaits », 7% « Neutres », 9% « Pas satisfaits » et 0% « Très insatisfaits ». Les résultats de cette évaluation confirment que l'expérience interactive est le moteur de l'adoption de l'AV (Duguleană et al., 2020). De même, les résultats de l'évaluation de l'utilisabilité ont été satisfaisants avec 70% de réponses « Très satisfait ».

IV. Conclusion

Un assistant virtuel a été conçu pour interagir avec les utilisateurs d'un bâtiment en utilisant des flux vocaux et visuels pour les aider à trier leurs déchets. Plusieurs techniques d'IA ont été utilisées pour permettre les fonctionnalités de reconnaissance vocale et d'image.

La méthodologie AGILE a été utilisée pour diriger le développement de la solution. Puis, une évaluation des performances de la solution a été réalisée en effectuant une évaluation technique et une évaluation de la convivialité.

Plusieurs pistes d'amélioration ont été identifiées :

- Améliorer l'expérience utilisateur en incluant un avatar virtuel, qui peut être conçu et ajouté à l'écran (Ren et al., 2017). L'avatar virtuel peut encore améliorer les interactions entre l'AV et les utilisateurs en rendant la conversation plus interactive.
- Améliorer la convivialité de l'AV en ajoutant davantage de contenu à l'écran de l'AV, notamment des informations pratiques sur l'installation (Duguleană et al., 2020).
- Évaluer la convivialité pour les utilisateurs administrateurs
- Augmenter le nombre d'objets pris en charge par l'outil de reconnaissance d'images

Le présent travail peut constituer une base solide pour étudier une mise en œuvre plus poussée des assistants virtuels et de l'IA afin d'améliorer les pratiques de dans l'exploitation et la maintenance des Bâtiments. D'autres travaux peuvent être effectués pour rendre cet AV opérationnel.

REFERENCES

- Ahmed, A., Ahmad, S., Ehsan, N., Mirza, E., Sarwar, S.Z., 2010. Agile software development: Impact on productivity and quality, in: 5th IEEE International Conference on Management of Innovation and Technology, ICMIT2010. <https://doi.org/10.1109/ICMIT.2010.5492703>
- Alfalalah, G., Zayed, T., 2020. A review of sustainable facility management research. *Sustain. Cities Soc.* 55. <https://doi.org/10.1016/j.scs.2020.102073>
- Chateau, B., Lapillonne, B., 1982. *Energy Demand in the Residential and Tertiary Sector*. Springer, Vienna, pp. 20–72. https://doi.org/10.1007/978-3-7091-8639-8_3
- Chew, M.Y.L., Conejos, S., Asmone, A.S., 2017. Developing a research framework for the green maintainability of buildings. *Facilities* 35, 39–63. <https://doi.org/10.1108/F-08-2015-0059>
- Chien, Y.H., Yao, C.K., 2020. Development of an ai userbot for engineering design education using an intent and flow combined framework. *Appl. Sci.* 10, 1–14. <https://doi.org/10.3390/app10227970>
- Duguleană, M., Briciu, V.A., Duduman, I.A., Machidon, O.M., 2020. A virtual assistant for natural interactions in museums. *Sustain.* 12, 1–17. <https://doi.org/10.3390/SU12176958>
- Elmualim, A., Valle, R., Kwawu, W., 2012. Discerning policy and drivers for sustainable facilities management practice. *Int. J. Sustain. Built Environ.* 1, 16–25. <https://doi.org/10.1016/j.ijbsbe.2012.03.001>
- Goyal, S., Esposito, M., Kapoor, A., 2018. Circular economy business models in developing economies: Lessons from India on reduce, recycle, and reuse paradigms. *Thunderbird Int. Bus. Rev.* <https://doi.org/10.1002/tie.21883>
- Grave, E., Bojanowski, P., Gupta, P., Joulin, A., Mikolov, T., 2019. Learning word vectors for 157 languages. *Lr. 2018 - 11th Int. Conf. Lang. Resour. Eval.* 3483–3487.
- Hoy, M.B., 2018. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Med. Ref. Serv. Q.* 37, 81–88. <https://doi.org/10.1080/02763869.2018.1404391>
- Iannizzotto, G., Bello, L. Lo, Nucita, A., Grasso, G.M., 2018. A vision and speech enabled, customizable, virtual assistant for smart environments. *Proc. - 2018 11th Int. Conf. Hum. Syst. Interact. HSI 2018* 50–56. <https://doi.org/10.1109/HSI.2018.8431232>
- Nielsen, S.B., Sarasoja, A.L., Galamba, K.R., 2016. Sustainability in facilities management: an overview of current research. *Facilities* 34, 535–563. <https://doi.org/10.1108/F-07-2014-0060>
- Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: *Proceedings of the IEEE*

Computer Society Conference on Computer Vision and Pattern Recognition.
<https://doi.org/10.1109/CVPR.2015.7298594>

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the Inception Architecture for Computer Vision, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. <https://doi.org/10.1109/CVPR.2016.308>

Weizenbaum, J., 1966. ELIZA-A computer program for the study of natural language communication between man and machine. Commun. ACM.
<https://doi.org/10.1145/365153.365168>

Xu, S., Wang, J., Wang, X., Shou, W., 2019. Computer vision techniques in construction , operation and maintenance phases of civil assets : A critical review.