

CHALLENGE DE LA DATA

Data Paper

De quelle manière l'analyse des interactions du public à travers la FAQ et le ChatBot permet-elle d'identifier les attentes et préoccupations majeures lors de grands événements ?

Table des matières

Introduction	3
Description des données	4
Méthodologie et traitement des données	5
Analyse des résultat et interprétation	6
Perspective d'amélioration	7
Conclusion	8

Introduction

L'organisation des Jeux Olympiques et Paralympiques représente un défi logistique et humain de grande ampleur. Assurer une expérience optimale aux spectateurs, aux athlètes et aux organisateurs nécessite une compréhension fine de leurs attentes et préoccupations. Pour répondre à cet enjeu, les Jeux de Paris 2024 ont mis en place des FAQ et des ChatBots destinés à informer le public en temps réel. Ces outils numériques ne se contentent pas de fournir des réponses: ils constituent également une source précieuse de données permettant d'affiner la gestion de l'événement et d'identifier des pistes d'amélioration pour les éditions futures.

Ce travail a pour objectif d'exploiter les données ouvertes des JO 2024, en particulier celles liées aux échanges du public via la FAQ et les ChatBots, afin d'identifier les principales préoccupations et attentes des usagers lors de ces grands événements. En mettant en lumière les tendances et les insights dégagés de ces échanges, nous cherchons à fournir des recommandations pour optimiser l'organisation et l'impact des événements, qu'ils soient sportifs, culturels ou éducatifs. Plus largement, cette étude cherche à valoriser l'héritage des Jeux Olympiques et Paralympiques de Paris 2024 en exploitant les données ouvertes issues des interactions du public. L'objectif est d'identifier des enseignements permettant d'optimiser l'organisation et l'impact d'événements éducatifs, sportifs, culturels ou de grande ampleur sur les territoires, en améliorant leur accessibilité, leur logistique et l'engagement du public.

Notre réflexion s'articule autour de la question suivante : ***De quelle manière l'analyse des interactions du public à travers la FAQ et le ChatBot permet-elle d'identifier les attentes et préoccupations majeures lors de grands événements ?*** Cette question vise à explorer la manière dont ces outils numériques peuvent offrir des informations pertinentes pour anticiper les besoins du public et améliorer l'expérience globale des événements.

Par ailleurs, nous avons choisi le nom de notre groupe, Heraia, en référence aux concours gymniques antiques réservés aux femmes et organisés à Olympie en l'honneur de la déesse Héra.

Le choix de ce jeu de données repose sur l'importance des échanges en ligne dans la communication et l'interaction avec le public lors des grands événements. Les FAQ et les ChatBots sont des interfaces numériques qui permettent de recueillir les préoccupations des usagers. En exploitant ces données, nous pouvons non seulement tirer des conclusions pour les JO 2024, mais aussi offrir des recommandations pour l'organisation de futurs événements de grande envergure, en nous appuyant sur une base solide d'interactions réelles et tangibles avec le public.

Ce travail s'inscrit dans une logique d'exploitation intelligente des données pour renforcer la prise de décision dans l'organisation des grands événements.

Description des données

FAQ

Le jeu de données intitulé "FAQ des JO 2024" représente un ensemble riche d'interactions entre le public et les organisateurs des Jeux Olympiques et Paralympiques de Paris 2024. Ce jeu inclut les questions posées par les usagers concernant divers aspects de l'événement ainsi que les réponses fournies par le centre de contact officiel de Paris 2024. Ces données constituent une ressource précieuse pour comprendre les préoccupations, attentes et besoins du public avant l'événement de grande ampleur.

Le jeu de données "FAQ des JO 2024" est organisé sous forme de tableau, chaque ligne représentant une question posée par un usager. Il contient plusieurs colonnes pour structurer les interactions :

Identifiant unique: Chaque question est associée à un identifiant unique, permettant de suivre et tracer précisément les interactions.

Langue: Les questions et réponses sont disponibles en français et en anglais. Une phase de traduction a été effectuée afin d'assurer l'unité linguistique et d'optimiser l'exploitation des données.

Question: Cette colonne contient la question formulée par l'utilisateur, qui peut varier en termes de clarté et de complexité.

Réponse: La colonne des réponses fournit des informations détaillées, souvent accompagnées de liens vers des ressources externes ou des sites web afin d'aider l'utilisateur à trouver plus de précisions.

Thématique: Les questions sont classées selon des catégories thématiques spécifiques.

URL d'accès: Chaque réponse peut inclure des liens vers des pages externes.

Les thématiques abordées dans ce jeu de données sont diverses et couvrent une large gamme de préoccupations du public. Parmi elles, on retrouve des questions pratiques liées à la billetterie, les transports, l'accessibilité, les engagements environnementaux, les conditions de sécurité et les événements olympiques et paralympiques.

ChatBot

Le jeu de données ChatBot du Centre de Contact de Paris 2024 se compose des échanges entre le public et l'outil automatisé conçu pour répondre aux questions des usagers. Contrairement à la FAQ, qui repose sur des questions spécifiques, le ChatBot fonctionne par l'utilisation de groupes de mots pour catégoriser les messages reçus et les associer à des thématiques prédéfinies.

Les informations disponibles dans ce jeu de données comprennent la date de l'interaction, le contenu du message et la catégorie thématique associée à chaque message.

Le volume d'interactions dans ce jeu de données est considérablement plus important que celui de la FAQ, avec près de 200 000 lignes de données, en comparaison des 632 lignes de la FAQ. Chaque message du ChatBot a été catégorisé selon la thématique abordée, ce qui permet de réaliser une analyse approfondie des préoccupations du public. Les catégories thématiques couvertes par le ChatBot sont majoritairement similaires à celles de la FAQ, mais avec des ajouts notables, comme des thématiques liées au recrutement / candidature ou à l'impact des JOP sur l'espace public, qui ne sont pas présentes dans le jeu de données de la FAQ.

Méthodologie et traitement des données

Le travail d'analyse a débuté par un nettoyage approfondi des données afin d'assurer leur fiabilité. Ce nettoyage a consisté à éliminer les doublons et les données incohérentes, ainsi qu'à standardiser les formats et harmoniser les libellés des catégories de questions. Cette étape préliminaire a permis de rendre les données prêtes pour une analyse détaillée et homogène.

Une première étape dans le jeu de données FAQ a consisté à exclure les questions et réponses strictement liées aux JO, comme celles concernant le label "Terre de Jeux 2024". Ces questions étaient trop spécifiques à cet événement pour être utilisées dans notre analyse qui visait à extraire des informations pouvant être généralisées à d'autres événements de grande envergure.

Ensuite, les questions ont été regroupées selon des thématiques générales, telles que la billetterie, l'accessibilité, et les transports. Certaines sous-catégories spécifiques (par exemple, "Billetterie SAV" ou "Billetterie demande d'informations") ont été fusionnées sous des catégories plus larges pour simplifier l'analyse.

Une fois les données nettoyées et préparées, nous avons procédé à l'analyse des préoccupations majeures du public, en utilisant différentes techniques selon le jeu de données.

Nous avons commencé par une analyse lexicale pour le jeu de données FAQ. Nous avons extrait les mots-clés les plus fréquemment utilisés dans chaque catégorie de question. Cela nous a permis de repérer des thématiques importantes, comme la billetterie, les boutiques, et les transports.

Nous avons également créé des tableaux croisés dynamiques pour comptabiliser le nombre de questions dans chaque catégorie. Ces tableaux nous ont permis de faire ressortir les thématiques les plus fréquemment abordées par les usagers. Par exemple, le thème des boutiques est rapidement apparu comme un sujet central, suivi de la billetterie et de l'accessibilité.

Une des caractéristiques intéressantes du jeu de données ChatBot est sa dimension temporelle, qui permet d'analyser l'évolution des préoccupations du public au fil des mois. Par exemple, en juillet 2024, la billetterie est devenue la thématique la plus fréquemment abordée. Grâce à l'analyse des données mensuelles, nous avons pu isoler les principales préoccupations de chaque période et observer leur évolution dans le temps.

Au départ, nous avons tenté de compter les mots-clés par thématique et par jour. Cependant, l'énorme volume de données rendait cette analyse difficile à gérer. Nous avons donc ajusté notre approche en utilisant des tableaux croisés mensuels, qui ont permis d'agréger les données par mois grâce à des formules de somme dans Excel. Cela nous a permis de regrouper les occurrences des mots-clés par thématique sur une période donnée, rendant ainsi l'analyse plus cohérente et compréhensible.

Le ChatBot a ajouté une dimension dynamique à notre analyse. Nous avons utilisé les données du ChatBot pour compléter celles de la FAQ, ce qui a permis de mieux cerner les préoccupations évolutives du public et d'avoir une vue plus précise de l'évolution des demandes sur une période donnée. Les données du ChatBot ont enrichi l'analyse en offrant une perspective temporelle et en fournissant une vision complémentaire aux données de la FAQ.

Nos datavisualisations :

[La répartition des questions du public lors des JO : Types, fréquences, mots clés](#)

[Évolution Mensuelle des Demandes d'Information par Catégorie – Jeux Olympiques 2024](#)

Descripti on des données

FAQ

Le jeu de données intitulé "FAQ des JO 2024" représente un ensemble riche d'interactions entre le public et les organisateurs des Jeux Olympiques et Paralympiques de Paris 2024. Ce

Le jeu inclut les questions posées par les usagers concernant divers aspects de l'événement ainsi que les réponses fournies par le centre de contact officiel de Paris 2024. Ces données constituent une ressource précieuse pour comprendre les préoccupations, attentes et besoins du public avant l'événement de grande ampleur.

Le jeu de données "FAQ des JO 2024" est organisé sous forme de tableau, chaque ligne représentant une question posée par un usager. Il contient plusieurs colonnes pour structurer les interactions :

Identifiant unique: Chaque question est associée à un identifiant unique, permettant de suivre et tracer précisément les interactions.

Langue: Les questions et réponses sont disponibles en français et en anglais. Une phase de traduction a été effectuée afin d'assurer l'unité linguistique et d'optimiser l'exploitation des données.

Question: Cette colonne contient la question formulée par l'utilisateur, qui peut varier en termes de clarté et de complexité.

Réponse: La colonne des réponses fournit des informations détaillées, souvent accompagnées de liens vers des ressources externes ou des sites web afin d'aider l'utilisateur à trouver plus de précisions.

Thématique: Les questions sont classées selon des catégories thématiques spécifiques.

URL d'accès: Chaque réponse peut inclure des liens vers des pages externes.

Les thématiques abordées dans ce jeu de données sont diverses et couvrent une large gamme de préoccupations du public. Parmi elles, on retrouve des questions pratiques liées à la billetterie, les transports, l'accessibilité, les engagements environnementaux, les conditions de sécurité et les événements olympiques et paralympiques.

Analyse des résultats et interprétation

L'analyse des données révèle plusieurs préoccupations majeures exprimées par le public. Parmi les thématiques les plus fréquemment abordées, on trouve la billetterie, avec des demandes concernant principalement le service après-vente (SAV) et les demandes d'informations générales.

En effet, la catégorie "Billetterie - SAV" a cumulé 42 675 interactions, tandis que "Billetterie - Demande d'information" a enregistré 20 568 demandes.

Cette forte demande dans le domaine de la billetterie est confirmée par les données mensuelles : en juillet 2024, la billetterie représentait à elle seule un volume impressionnant de 21 473 messages dans cette catégorie.

Les préoccupations relatives aux informations pratiques, notamment sur l'accessibilité, les transports, et la préparation à l'événement, figurent également parmi les sujets les plus traités. Le nombre de questions dans ces domaines montre un intérêt soutenu pour les conditions de participation et l'organisation des événements. De plus, les questions concernant la boutique et le merchandising ont généré 3 797 discussions, ce qui témoigne d'un intérêt pour les produits associés aux événements.

La célébration et les événements spéciaux ont également été des sujets d'attention, enregistrant 7 696 messages. En revanche, des sujets comme le recrutement / candidature (2 307 messages) ou l'hospitalité (1 560 messages) ont suscité moins d'intérêt, bien qu'ils restent des catégories pertinentes à prendre en compte pour une vision complète de la gestion de l'événement.

Perspective d'amélioration

Grâce à notre analyse des données issues du Chatbot et de la FAQ, nous avons identifié les principales préoccupations des spectateurs et participants des Jeux Olympiques et Paralympiques de Paris 2024. Cette analyse approfondie apporte des éléments de réponse précieux pour anticiper les attentes et le potentiel manque d'informations du public. Cela permet d'améliorer l'expérience utilisateur lors d'événements futurs.

Suite à notre travail, nous avons remarqué que les thèmes les plus recherchés concernent les boutiques officielles et la billetterie, ce qui en fait une priorité pour l'optimisation des plateformes d'information. Afin de répondre efficacement à ces demandes, il serait pertinent de prévoir une meilleure mise en avant des informations principales sur la billetterie et les boutiques directement sur le site officiel. L'intégration d'une FAQ interactive pourrait être une solution car elle permettrait un accès simplifié aux réponses sur les modalités de paiement, les retours, les remboursements et les étapes après achat. Il serait également intéressant de créer du contenu explicatif sous forme de vidéos et d'infographies pour guider les utilisateurs sur l'achat de billets, l'échange d'articles ou encore l'utilisation des billets aux entrées des sites de compétition de manière plus agréable et ludique.

De plus, l'accessibilité des transports en commun, l'adaptation des sites de compétition et les billets spécifiques pour les personnes en situation de handicap figurent parmi les principales préoccupations. Un manque d'informations centralisées oblige les spectateurs concernés à effectuer des recherches individuelles. Pour pallier cela, il serait pertinent de mettre en ligne un guide d'accessibilité dès l'ouverture de la billetterie regroupant toutes les

informations sur les transports adaptés, l'accessibilité des sites et les services dédiés aux personnes en situation de handicap. Une communication plus proactive via des campagnes d'emailing contenant les réponses aux questions les plus fréquentes serait également un levier efficace. Enfin, la création de supports interactifs comme des vidéos explicatives, des infographies et des cartes dynamiques permettrait de visualiser les accès, les transports accessibles et les services sur site.

Suite à cette anticipation, une expérience utilisateur plus fluide et personnalisée pourrait être développée avec un espace dédié permettant aux visiteurs d'accéder rapidement aux informations qui les concernent comme les billets, le programme et les recommandations logistiques. Il serait aussi intéressant d'intégrer des recommandations basées sur la localisation et les centres d'intérêt des spectateurs.

En fin de compte, un cadre structuré pourrait être développé, intégrant des outils comme des FAQ dynamiques, des chatbots avancés, des guides interactifs et des plateformes de communication.. Ce modèle adaptable aiderait à standardiser les réponses aux questions récurrentes tout en prenant en compte les spécificités de chaque événement.

Conclusion

L'analyse des échanges via la FAQ et le ChatBot des JO 2024 met en lumière les principales préoccupations du public, telles que la billetterie, l'accessibilité et les boutiques. Ces données fournissent des insights précieux pour améliorer l'organisation et l'expérience des futurs événements. En identifiant les sujets les plus fréquemment abordés, comme les informations pratiques, les datavisualisations permettent d'anticiper les besoins du public et d'optimiser la gestion des événements. Cette approche vous offre une base solide pour adapter les futurs événements aux attentes des participants et garantir une expérience plus fluide et personnalisée, tout en valorisant l'héritage des Jeux Olympiques et Paralympiques de Paris 2024 pour optimiser l'organisation et l'impact d'événements de grande ampleur grâce aux données ouvertes.

Bibliographie

Paris 2024 - FAQ. <https://data.education.gouv.fr/explore/dataset/paris-2024-faq/table/>. Consulté le 27 mars 2025.

Paris 2024 - Messages chatbot centre de contact grand public. <https://data.education.gouv.fr/explore/dataset/paris-2024-messages-chatbot-centre-de-contact-grand-public/table/>. Consulté le 27 mars 2025.