

Développeurs et data scientists : les 5 principales façons de collaborer

Ne craignez pas les data scientists

Afin de créer des applications efficaces basées sur l'intelligence artificielle (IA), il vous faut connaître le mode de travail des data scientists, mais aussi savoir comment travailler à leurs côtés. Vous devez donc maîtriser les bases de l'IA et être capable de collaborer avec vos collègues spécialistes en science des données. Voici les cinq principaux éléments à connaître en cas de collaboration avec des data scientists et de création d'applications intelligentes guidées par l'IA. Servez-vous de cette liste de contrôle comme un guide pour établir des relations professionnelles solides et mettre en place des collaborations d'exception pour le développement d'applications.

1 Comprenez le mode de travail des data scientists

En général, les data scientists s'occupent davantage de la création et du perfectionnement de leurs modèles que du développement des applications ou de l'intégration de leurs modèles dans un logiciel. Il est rare qu'ils souhaitent être impliqués dans la création de pipelines d'intégration et de distribution continues (CI/CD) ou dans la rédaction du code d'application. Par ailleurs, ils utilisent des outils qui peuvent vous être étrangers, comme Python, R ou encore Jupyter Notebook.

Ils ne seront vraisemblablement pas les premiers à vous suggérer une collaboration ouverte, même si vous visez tous les mêmes objectifs. Toutefois, un travail d'équipe sérieux et cohérent entre vous et votre équipe de spécialistes en science des données est essentiel à la création d'applications efficaces. Une collaboration active se traduit par :

- ▶ le déploiement d'applications intelligentes axées sur les données qui tirent pleinement parti de l'IA ;
- ▶ une occasion pour les data scientists de mettre à contribution leur travail de modélisation pour des solutions prêtes au déploiement qui apportent de la valeur ajoutée à votre entreprise et à vos clients.

Il vous incombera probablement d'établir le premier contact et de faciliter l'expérience de collaboration avec vos collègues spécialistes en science des données. Suivez les recommandations fournies dans cette liste de contrôle pour découvrir comment forger un lien bénéfique avec votre équipe de data scientists.

2 Trouvez un terrain d'entente

Afin de mettre en place une collaboration fluide et une expérience qui convient à tous, il est important d'expliquer vos pratiques de développement et de constater dans quelle mesure elles complètent les efforts déployés par vos data scientists. Pour ce faire, suivez les conseils ci-dessous.

- ▶ **Encouragez les points de contact fréquents :** organiser des points de contact de façon fréquente et régulière est le meilleur moyen de garantir la bonne progression de vos projets communs en cours.
- ▶ **Respectez les limites :** il n'est peut-être pas désiré ou nécessaire que les data scientists sachent de quelle manière les applications passent en production. Bien que les pratiques MLOps soient un concept prisé, certains scientifiques préfèrent vous transférer leurs notebooks Jupyter par e-mail. Respectez leurs intérêts et leur façon de travailler pour qu'ils en fassent autant avec vous.
- ▶ **Partagez vos processus entre vous :** en plus de vous renseigner sur le mode de travail des data scientists, partagez vos processus et les outils que vous utilisez en production, comme Git, Tekton ou Kubernetes. Dans l'esprit de l'Open Source, donnez-leur un aperçu de vos processus.
- ▶ **Utilisez une plateforme de collaboration commune :** les plateformes communes de développement d'applications IA cloud-native telles que Red Hat® OpenShift® Data Science appuient et encouragent la collaboration entre vous et votre équipe spécialisée en science des données. La plateforme démocratise l'utilisation d'outils d'IA et permet aux équipes de mettre en œuvre et d'accélérer le développement d'applications intelligentes.

3 Apprenez à travailler avec des outils d'entraînement des modèles

Apprenez au moins les bases de certains des outils d'entraînement des modèles que les data scientists utilisent régulièrement. Avec une connaissance pratique de ces outils, vous serez en mesure de comprendre comment les modèles sont créés. Parmi les bibliothèques et les outils d'entraînement des modèles les plus prisés, on peut citer :

- ▶ les environnements de développement Jupyter et PyCharm ;
- ▶ PyTorch ;
- ▶ scikit-learn ;
- ▶ TensorFlow.

En vous familiarisant avec ces outils, mais aussi d'autres, vous mettez toutes les chances de votre côté pour réussir à créer des applications de modèles déployés. Vous pourrez, par la même occasion, mieux comprendre le travail qu'implique la création des modèles et résoudre plus facilement les problèmes liés à un manque de fluidité de l'intégration des modèles dans vos applications intelligentes.

4 N'abandonnez pas vos outils et processus de prédilection

Le travail avec des data scientists et l'IA nécessite d'apprendre à se conformer à un grand nombre de nouveaux processus et à utiliser quelques nouveaux outils. Néanmoins, vous pouvez continuer à utiliser la plupart de vos outils de prédilection pour appliquer votre logique. La modélisation et le code des applications peuvent se faire dans le langage et le framework souhaités.

Par exemple, en tant que développeur Quarkus, vous pouvez effectuer votre logique d'application dans Quarkus et faire en sorte qu'elle appelle une interface de programmation d'application (API) vers un point de terminaison REST (Representational State Transfer). Vos data scientists, quant à eux, se chargent du traitement des données et des prédictions à l'aide d'un outil comme Python ou R. Malgré le caractère complexe de l'IA et de la science des données, vous pouvez simplifier votre travail en ayant recours à des outils et à des processus que vous maîtrisez.

5 N'oubliez pas : le modèle fait partie de l'application

Si le modèle est important, les pratiques MLOps qui le soutiennent le sont tout autant. Pour garantir la continuité du bon fonctionnement de vos modèles lorsqu'ils sont mis en production, suivez ces quatre conseils :

- ▶ Créez une infrastructure de mise à disposition des modèles adaptée à l'application que vous développez.
- ▶ Concevez de nouveaux pipelines CI/CD ou étendez les pipelines existants afin de gérer à la fois l'entraînement et la mise à disposition du modèle.
- ▶ Mettez à l'échelle l'application de mise à disposition du modèle.
- ▶ Intégrez des services de données de diffusion en continu comme Apache Kafka et d'autres composants de collecte de données.

En fin de compte, le déploiement de votre application n'est que le début. Les modèles évoluent constamment et nécessitent une surveillance. Travaillez en étroite collaboration avec votre équipe de data scientists afin de définir les indicateurs de mesure que vous ou vos homologues de l'équipe d'exploitation devez surveiller pour éviter tout écart du modèle. Si un problème ou un changement se produit, collaborez avec vos data scientists pour affiner et améliorer vos modèles.

En savoir plus

Consultez la page dédiée à [Red Hat OpenShift Data Science](#) et lisez notre [livre numérique](#).

Tester la solution

Essayez [Red Hat OpenShift Data Science](#) et regardez une démonstration de la plateforme.



À propos de Red Hat

Red Hat aide ses clients à standardiser leurs environnements, à développer des applications cloud-native et à intégrer, automatiser, sécuriser et gérer des environnements complexes en offrant des services d'assistance, de formation et de conseil [primés](#).

f facebook.com/redhatinc
t [@RedHatFrance](https://twitter.com/RedHatFrance)
in linkedin.com/company/red-hat

**Europe, Moyen-Orient
et Afrique (EMEA)**
 00800 7334 2835
europe@redhat.com

France
 00 33 1 4191 23 23
fr.redhat.com