

# Die fünf besten Ansätze für die Zusammenarbeit zwischen Entwicklern und Data Scientists

## Data Scientists beißen nicht.

Damit Sie leistungsfähige, auf künstlicher Intelligenz (KI) basierende Anwendungen entwickeln können, müssen Sie wissen, wie Data Scientists arbeiten und wie Sie mit ihnen zusammenarbeiten können. Das bedeutet, dass Sie über die Grundlagen von KI Bescheid wissen sollten und darüber, wie Sie mit Ihren Kollegen aus dem Bereich Data Science zusammenarbeiten können. Im Folgenden finden Sie die fünf wichtigsten Punkte, die Sie bei der Zusammenarbeit mit Data Scientists und der Entwicklung KI-gesteuerter intelligenter Anwendungen beachten sollten. Nutzen Sie diese Checkliste als Guide für gute Arbeitsbeziehungen und Zusammenarbeit bei der Anwendungsentwicklung.

## 1 Verstehen Sie, wie Data Scientists arbeiten

Data Scientists sind meist mehr mit der Erstellung und Optimierung ihrer Modelle beschäftigt als mit der Anwendungsentwicklung oder der Integration ihrer Modelle in eine Software. Nur selten sind sie in den Aufbau von CI/CD-Pipelines (Continuous Integration/Continuous Delivery) oder das Schreiben von Anwendungscode involviert und verwenden Tools wie Python, R und Jupyter Notebooks, mit denen Sie vielleicht nicht vertraut sind.

Obwohl Sie alle auf die gleichen Ziele hinarbeiten, ergreifen Ihre Kollegen von der Data Science möglicherweise nicht die Initiative für eine offene Zusammenarbeit. Jedoch ist eine solide, konsistente Teamarbeit zwischen Ihnen und Ihrem Data-Science-Team für die Erstellung gut funktionierender Anwendungen unerlässlich. Eine aktive Zusammenarbeit kann zu folgenden Ergebnissen führen:

- ▶ das Deployment von intelligenten, datengesteuerten Anwendungen, die die Vorteile der KI effektiv nutzen
- ▶ die Möglichkeit für Data Scientists, ihre Modellierungserfahrung in einsatzfähige Lösungen einfließen zu lassen, die einen Mehrwert für Ihr Unternehmen und seine Kunden darstellen

Vermutlich müssen Sie die ersten Kontakte knüpfen und die Zusammenarbeit mit Ihren Kollegen von der Data Science fördern. Orientieren Sie sich an dieser Checkliste, um herauszufinden, wie Sie mit Ihrem Data-Science-Team am besten in Kontakt treten können.

## 2 Finden Sie eine gemeinsame Basis

Erklären Sie Ihre Entwicklungsabläufe und zeigen Sie auf, wie diese die Arbeit der Data Scientists ergänzen. So können Sie eine reibungslose Zusammenarbeit und ein gutes Erlebnis für alle Beteiligten schaffen. Um dies zu erreichen, können Sie:

- ▶ **Häufige Kontaktpunkte fördern.** Ein häufiger und regelmäßiger Kontakt hilft die gemeinsamen Projekte auf Kurs zu halten.
- ▶ **Grenzen respektieren.** Data Scientists sind vielleicht nicht daran interessiert oder müssen nicht wissen, wie Sie Anwendungen in die Produktion einführen. Obwohl MLOps ein beliebtes Konzept ist, bevorzugen einige Data Scientists den Versand ihrer Jupyter Notebooks per E-Mail. Respektieren Sie die Interessen und die Arbeitsmethoden Ihrer Kollegen für eine produktive Zusammenarbeit.
- ▶ **Sich über Abläufe austauschen.** Lernen Sie nicht nur, wie Data Scientists arbeiten, sondern teilen Sie auch Ihre Arbeitsabläufe und die Tools, die Sie in der Produktion verwenden, wie Git, Tekton oder Kubernetes. Gewähren Sie ihnen im Rahmen des Open Source-Konzepts einen Einblick in Ihre Abläufe.
- ▶ **Gängige Plattform zur Zusammenarbeit verwenden.** Gängige cloudnative KI-Entwicklungsplattformen wie Red Hat® OpenShift® Data Science können Sie bei der Zusammenarbeit mit Ihrem Data-Science-Team unterstützen und fördern. Die Plattform demokratisiert die Nutzung von KI-Tools und ermöglicht Teams die Implementierung und Beschleunigung der Entwicklung intelligenter Anwendungen.

### 3 Erlernen Sie den Umgang mit Tools für Modelltraining

Lernen Sie zumindest die Grundlagen einiger der von Data Scientists regelmäßig verwendeten Tools für das Training von Modellen. Mit Kenntnissen über Tools zum Modelltraining können Sie besser verstehen, wie die Modelle aufgebaut sind. Die folgenden sind einige der beliebtesten Tools und Libraries für das Modelltraining:

- ▶ Jupyter und PyCharm Entwicklungsumgebungen
- ▶ PyTorch
- ▶ scikit-learn
- ▶ TensorFlow

Wenn Sie sich mit diesen und anderen Tools vertraut machen, sind Sie besser in der Lage, erfolgreich bereitgestellte Modellanwendungen zu erstellen. Außerdem erhalten Sie ein besseres Verständnis für die Arbeit, die mit der Erstellung der Modelle verbunden ist. So können Sie Probleme lösen, wenn sich die Modelle nicht reibungslos in Ihre intelligenten Anwendungen integrieren lassen.

### 4 Behalten Sie Ihre bevorzugten Tools und Abläufe bei

Wenn Sie mit Data Scientists und KI arbeiten, werden Sie viele neue Abläufe und ein paar neue Tools erlernen müssen. Verwenden Sie aber weiterhin viele Ihrer bevorzugten Tools, um Ihre Logik zu erstellen. Der Anwendungscode und die Modellierung können in jeder beliebigen Sprache oder jedem beliebigen Framework erstellt werden.

Als Quarkus-Entwickler können Sie beispielsweise Ihre Anwendungslogik in Quarkus erstellen und über eine API (Application Programming Interface) einen REST (Representational State Transfer)-Endpunkt aufrufen, während Data Scientists die eigentliche Datenverarbeitung und Prognosen mit einem Tool wie Python oder R durchführen. Da KI und Data Science komplexe Themen sind, sollten Sie Ihre Arbeit einfach gestalten und vertraute Tools und Abläufe verwenden.

### 5 Das Modell ist Teil der Anwendung

Das Modell ist wichtig, ebenso wie die MLOps, die hinter dem Modell stehen. Die folgenden vier Punkte sollten Sie beachten, damit Ihre Modelle in der Produktion weiterhin gut funktionieren:

- ▶ Erstellen Sie eine modellgestützte Infrastruktur, die für die von Ihnen entwickelte Anwendung geeignet ist.
- ▶ Erstellen Sie neue oder erweitern Sie bestehende CI/CD-Pipelines, um sowohl das Training als auch die Bereitstellung des Modells zu gewährleisten.
- ▶ Skalieren Sie die modellgestützte Anwendung.
- ▶ Integrieren Sie Streaming-Datenservices wie Apache Kafka und andere Datenerfassungskomponenten.

Schließlich ist das Deployment Ihrer Anwendung nur der Anfang. Die Modelle ändern sich ständig und müssen kontrolliert werden. Arbeiten Sie eng mit dem Data-Science-Team zusammen und legen Sie fest, welche Metriken Sie oder Ihre Kollegen aus dem operativen Bereich kontrollieren müssen, um eine Modellabweichung zu verhindern. Im Falle eines Problems oder einer Änderung arbeiten Sie mit den Data Scientists zusammen, um Ihre Modelle zu präzisieren und zu optimieren.

#### Mehr erfahren

Gehen Sie zu [Red Hat OpenShift Data Science](#) und [lesen Sie das E-Book](#).

#### Ausprobieren

[Testen Sie Red Hat OpenShift Data Science](#) und sehen Sie sich eine Demo an.



#### Über Red Hat

Red Hat unterstützt Kunden dabei, ihre Umgebungen zu standardisieren, cloudnative Anwendungen zu entwickeln und komplexe Umgebungen mit [vielfach ausgezeichnetem](#) Support, Training und Consulting Services zu integrieren, zu automatisieren, zu sichern und zu verwalten.

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

#### EUROPA, NAHOST, UND AFRIKA (EMEA)

00800 7334 2835  
[de.redhat.com](https://de.redhat.com)  
[europe@redhat.com](mailto:europe@redhat.com)

#### TÜRKEI

00800 448820640

#### ISRAEL

1 809 449548

#### VAE

8000-4449549