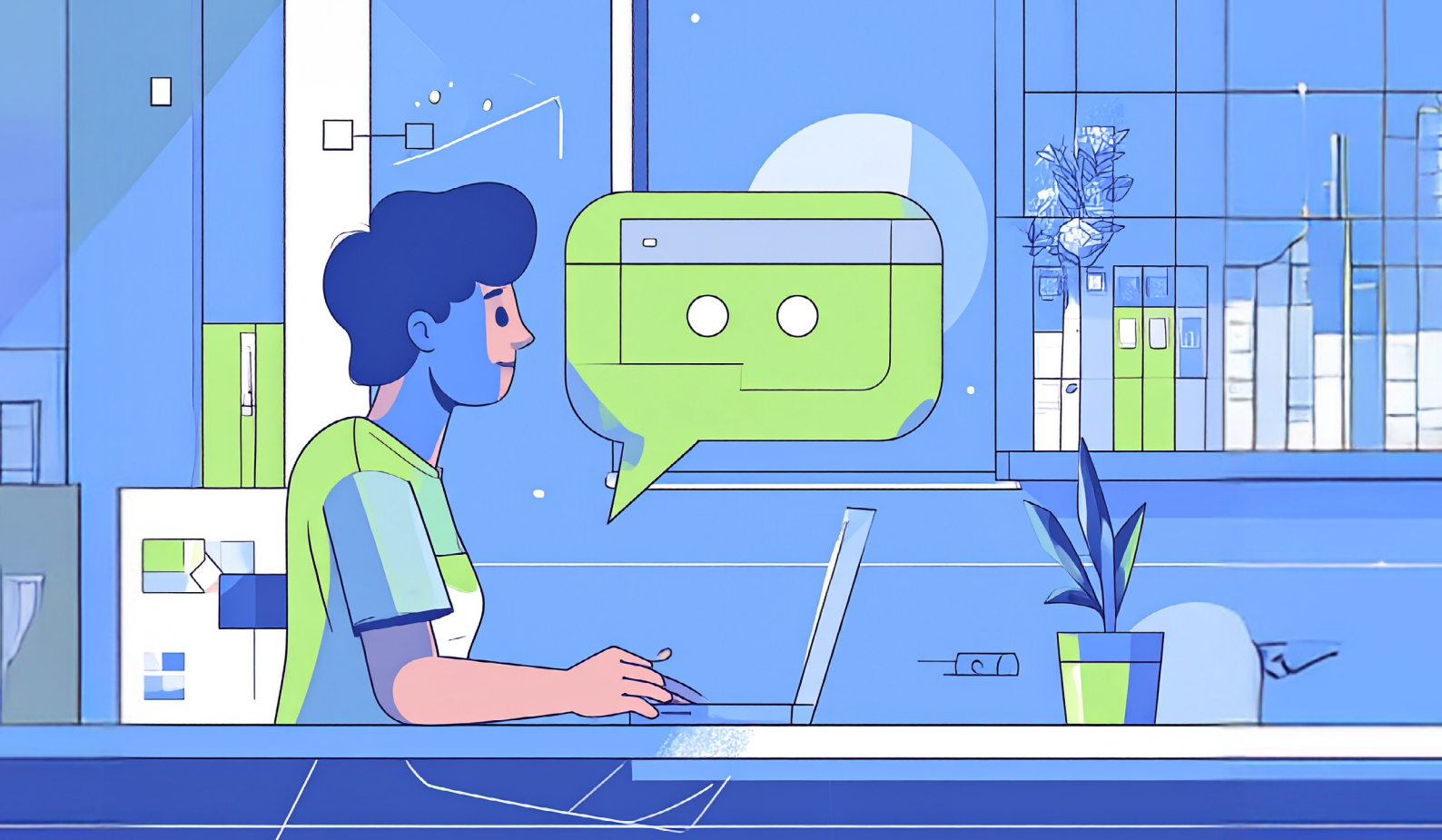


UND ACTION!



Tool Execution im
Companion for First



Modernen Goldgräbern auf der Spur

Daten gelten so manchem heute als das Gold des 21. Jahrhunderts. Ganz sicher existieren in den Tiefen moderner Computersysteme mehr als genug davon. Worauf es ankommt, ist, die entscheidenden Bröckchen aufzuspüren und zum Strahlen zu bringen. Mit dem Companion for First geben wir unseren First Cloud Kunden aus der Finanzindustrie dafür ein maßgeschneidertes Werkzeug an die Hand.



Der Companion steht Ihnen als kompetenter Sparringspartner und unermüdlicher Assistent bei der täglichen Arbeit mit First Cloud zur Seite. Ohne sich mit technischen Details, komplexem IT-Know-how und einer Reihe manueller Ausführungsschritte zu belasten, können Sie mit dem Companion tief in Ihre Datenbestände abtauchen und die gewünschten Informationen zutage fördern.

Schritt für Schritt

In den vorangehenden drei Whitepapers haben wir die Grundlagen der KI-Technik beleuchtet und uns auf die Spuren der Nutzereingaben und Datenströme durch das Innenleben des Companions begeben. Wir haben die sogenannte „**RAG Pipeline**“ als Rückgrat des gesamten Prozesses kennengelernt und die Mechanismen zur Interpretation und zum Verständnis von Nutzeranfragen ans Licht geholt. Auf Basis einer simplen Nutzeranfrage sind dabei komplexe Handlungsanleitungen mit teils Dutzenden von Zwischenschritten erwachsen. Das war die Aufgabe des Query Processings.

Im vorliegenden Whitepaper geht es darum, wie die so entstandenen Handlungsanleitungen im Rahmen der Tool Execution umgesetzt, ausgeführt und miteinander verknüpft werden, um sich Schritt für Schritt den gesuchten Antworten zu nähern. Wir verlassen damit weitgehend das Feld der Künstlichen Intelligenz, denn in dieser Stufe der RAG Pipeline werden die meisten Prozessschritte mit den Mitteln klassischer, deterministischer Informatik abgewickelt. Die hier zutage geförderten Daten verbleiben deshalb auch vollständig innerhalb des Companion for First. Nur autorisierte Nutzer bekommen sie zu Gesicht.



KI VERSTEHEN

In dieser Reihe bisher erschienen:

- Grundlagen der KI-Technik
- Die RAG Pipeline im Companion for First
- Das Query Processing im Companion for First

Sie möchten mehr erfahren?

Alle Infos zum Companion for First auf einen Blick gibt es hier.

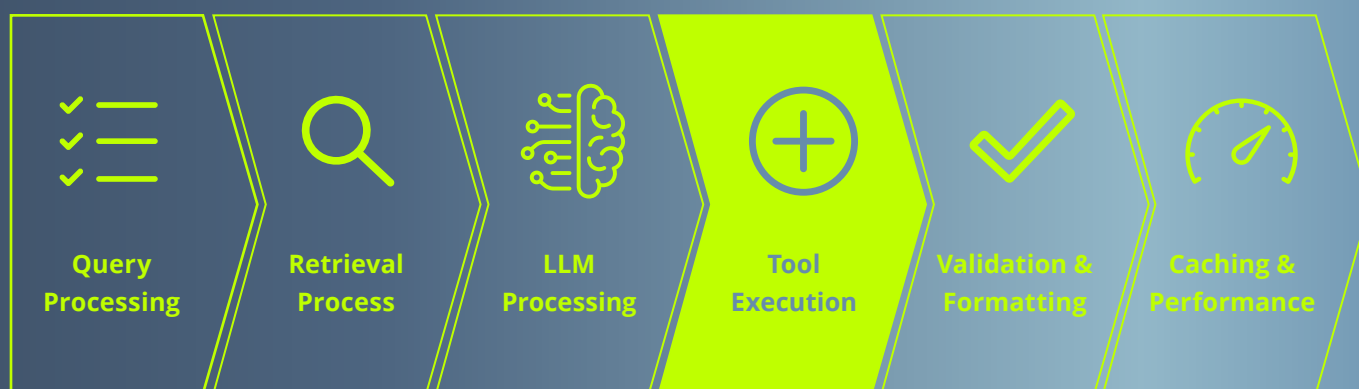
Alles in der richtigen Reihenfolge

Die im Query Processing als Antwort auf eine Nutzeranfrage erstellten Handlungsabläufe basieren auf kleinteiligen Arbeitsschritten, die jeweils mit ganz spezifischen Tools als deren „Ausführungsorgane“ verbunden sind. Diese Tools können Depotbestände nach Kriterien durchsuchen, Umsatz- und Stückzahlen aggregieren, einen einzelnen Aktienkurs zu einem gegebenen Zeitpunkt auslesen und vieles, vieles mehr. Sie definieren damit die Fähigkeiten und den Handlungsspielraum des Companion bei der inhaltlichen Generierung von Antworten, also beim Auflisten oder Zusammenrechnen konkreter Zahlen, Tabellen und Grafiken.

Zum Stand Januar 2026 sind im Companion mehr als 600 solcher Tools definiert, die teils ganz allgemein auf den mehr als 1500 Datenbanktabellen eines First Systems operieren, häufig jedoch ganz spezifisch – je nach konkreter Aufgabe.

Entscheidend ist, dass die Tool-Aufrufe vor dem Start der Ausführung in eine abgestimmte Reihenfolge gebracht werden, damit nicht der zweite Schritt vor dem ersten genommen wird. Das Aufsummieren von Beständen etwa macht ja nur Sinn, nachdem man zuvor die einzelnen Bestandswerte ermittelt hat. Und vergleichbares gilt für viele andere Vorgänge. Die unverzichtbare Sortierung vom Großen zum Kleinen, vom Output zum Input, erfolgt, bevor es richtig losgeht, beim Übergang vom Query Processing zur **Tool Execution**.

Im Übergang von Query Processing zu Tool Execution werden die Aufgaben zuerst sortiert.



Vom Tool zur Datenbank

Viele der Tools setzen ihre Aufgaben als klassische Datenbankoperationen in Form von **SQL-Befehlen** um. Es sind dies die Werkzeuge, die auch innerhalb der Benutzeroberfläche von First Cloud hinter den Kulissen zum Einsatz kommen, wenn Benutzer entsprechende Menübefehle aufrufen oder umfassendere Aktionen anstoßen.

Die Tools erhalten bei ihrem Aufruf einerseits Parameter, die ihre konkrete Vorgehensweise bestimmen, andererseits aber auch Filter, die den Output für die nachfolgende Verarbeitung eingrenzen. Nehmen Sie als Beispiel etwa die Suche nach den Anleihen eines gegebenen Emittenten, die bestimmte Kriterien bezüglich Laufzeit, Rating oder Konditionen erfüllen müssen. Von den gefundenen Anleihen werden stets nur jene zurückgeliefert, auf die die gegebenen Filterkriterien zutreffen.

In den meisten Fällen definiert der Output des einen Tools den Input des nachfolgend aufgerufenen. Die Daten werden also entlang der Tool Execution von Verarbeitungsschritt zu Verarbeitungsschritt weitergereicht. Der Companion selbst ist größtenteils in der modernen Programmiersprache **Python** verfasst und die Daten werden deshalb in Form sogenannter **„Python Dataframes“** verpackt und übermittelt.

So erwächst auf dem Weg durch den Ausführungsplan eine Art großes zweidimensionales Datenfeld mit den verarbeiteten Informationen, ein wenig vergleichbar mit einer Excel-Tabelle. All jene Spalten und Zeilen, die für das Ergebnis nicht unmittelbar relevant sind, werden am Ende der Verarbeitungskette abgeschnitten.

Der Output eines Tools definiert den Input des Nachfolgenden.



Validierung und Protokollierung

Eine Art Querschnittsfunktion, die sich entlang der gesamten Prozesskette durch alle Stufen der RAG Pipeline zieht, ist der Aspekt der Datenvalidierung.

Es beginnt beim Query Processing mit der Fragestellung, ob sich die Nutzeranfrage tatsächlich auf den Komplex der Kapitalanlageverwaltung mit First Cloud bezieht. Bei allen anderen Themen wird der Nutzer höflich um eine Neuformulierung der Fragestellung gebeten, denn Fragen kostet zwar nichts, der Companion versteht sich jedoch nicht als „Plaudertasche“ für alle Fälle. Schuster, bleib bei deinem Leisten, lautet hier das Motto.

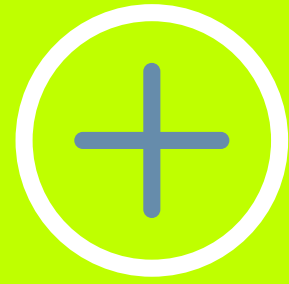
Weiter geht es mit der Datenvalidierung beim Query Processing. Dort gilt es, den Schritt-für-Schritt-Ausführungsplan unter diversen Gesichtspunkten zu prüfen: allen voran im Hinblick auf die Vollständigkeit und Ausführbarkeit. Liegt im Ausführungsplan kein Zirkelschluss vor, sind bei den anstehenden Tool-Aufrufen im Rahmen der Tool-Execution stets alle Pflichtparameter gesetzt?

Die individuelle **Parameterprüfung** innerhalb der einzelnen Tools beginnt bei der obligatorischen User-ID und erstreckt sich bis zum gewünschten Bezugsdatum für die abzurufenden Daten. Jedes Tool prüft selbständig, ob die mit der User-ID verbundenen Zugriffsrechte für die angeforderten Operationen ausreichend sind. Falls nicht, werden die gesuchten Daten in der Regel zwar nicht vollständig zurückgehalten, wohl aber entscheidende Details geschwärzt, bzw. in der Form „X.XXX,XX“ zurückgeliefert (Stichwort „Vorstandszahlen“).

Bei Tools, die potenziell sehr viele Daten liefern, findet intern auch jeweils eine technische Abschätzung im Hinblick auf die zu erwartende Datenmenge statt. Die Datengenerierung wird blockiert, falls das gesamte System durch übergroße Datenmengen in der Weiterverarbeitung beeinträchtigt werden könnte.

Wann immer ein Ausführungsplan an einer dieser Bedingungen scheitert, besteht die Option, stattdessen auf Plan B oder C zurückzugreifen – vorausgesetzt, das Query Processing hat mehr als einen möglichen Ausführungspfad zur Beantwortung einer Nutzeranfrage erbracht. Denn auch das gehört zur nicht-deterministischen Natur der KI-Technik.

Alle Tool-Aufrufe, ihre Parameter und Resultate werden übrigens revisionssicher protokolliert. Damit man auch später jederzeit nachvollziehen kann, wie welche Zahlen zustande gekommen sind, wer diese angefragt hat und wie der Companion überhaupt zu seiner Antwort gelangt ist.



TECHNISCHE ANBINDUNG DER TOOLS

Alle Tools werden als eigenständige und in sich geschlossene Web-Services umgesetzt. Das folgt einem Leitgedanken der Cloud Technik, der die gesamte Architektur von First Cloud durchzieht. Der große Vorteil: Für den Aufrufer ist es unerheblich, wo diese Dienste konkret ausgeführt werden, ob auf der aktuellen Maschine oder in einem entfernten Rechenzentrum. Das Ergebnis fließt unabhängig davon zu ihm zurück. Die passenden Aufrufe erfolgen über sogenannte **REST-APIs** und nutzen das bewährte HTTP-Protokoll, das auch für den Austausch zwischen Web-Clients und Web-Servern im World Wide Web zum Einsatz kommt.

Für den Datentransport zwischen den Tools und ihren Aufrufern – aber auch untereinander – wird **JSON** genutzt, die **JavaScript Object Notation**. Diese dient im Cloud-Bereich als eine Art Lingua franca für den datenbasierten Austausch zwischen unterschiedlichsten Computersystemen über die Grenzen von Programmiersprachen, Betriebssystemen und physischen Standorten hinweg. Ganz praktisch betrachtet werden quasi die Inhalte verschachtelter Datenstrukturen mit Programmvariablen, Zahlen, Zeichenketten und Listen ausgetauscht, in die sich alle Arten von Informationen hineinpacken lassen. Dank JSON versteht man sich.

Externe Tools in Reichweite

Die überwiegende Mehrzahl der Tools stammt direkt aus First Cloud, aber auch externe Tools werden genutzt. Beispielsweise für die Abfrage von Echtzeitdaten und historischer Kursverläufe von Aktientiteln und anderen Wertpapieren. First Cloud bedient sich dazu des anerkannten Dienstes **Alpha Vantage**, der ebenfalls über ein **REST-API** angesprochen wird. Alternativ können jederzeit weitere vergleichbare Dienste eingebunden werden.

First Cloud Kunden werden darüber hinaus eigene Tools integrieren können und diese auch durch den Companion for First abrufbar machen. Beispielsweise um den Zugriff auf eigene Datenbanken, hausinterne IT-Systeme oder externe Systeme von Kunden oder Dienstleistern einzubeziehen.

Entscheidend ist, dass dies bereits beim Auflösen einer Nutzeranfrage durch den Companion geschehen kann, also bei der Planerstellung auf der Ebene des Query-Processings. Diese Tools können deshalb auch zur Auflösung von Anfragen in den anderen KI-basierten Abschnitten der RAG Pipeline beitragen. Wir bewegen uns dabei im derzeit stark wachsenden Feld **agentischer KI-Funktionen**. Deshalb nutzt der Companion das populäre **MCP-Protokoll**, das vom US-amerikanischen KI-Anbieter Anthropic entwickelt, aber mittlerweile in die Hände der offenen Linux Foundation übertragen wurde (siehe Kasten).

Bevor sich der Vorhang hebt

Am Ende der Tool Execution liegt nach zahlreichen Verarbeitungsstufen und Datenoperationen schließlich die gewünschte Antwort auf die Anfrage des Nutzers vor. Doch die ermittelten Resultate in Form einer zweidimensionalen Datentabelle sind nicht das, was letztendlich auf dem Bildschirm erscheint. Im letzten großen Schritt steht noch die Aufbereitung dieser Daten zu aussagekräftigen Tabellen, Grafiken und Reports an, um sie in einen passenden fachlichen Rahmen zu setzen. Erst dadurch werden aus den Rohdaten greifbare Entscheidungshilfen und Präsentationsvorlagen für die Kapitalanlageverwaltung. Und genau darum wird es im folgenden Whitepaper gehen, wenn sich alles um die Ergebnisaufbereitung im Companion for First dreht. Dann kommt auch wieder die KI ins Spiel, so viel sei an dieser Stelle schon einmal verraten.

BAUSTEINE AGENTISCHER KI

Das **Model Context Protocol (MCP)** ist ein offener Standard, über den große KI-Sprachmodelle (**Large Language Models, LLM**) sicher und standardisiert mit externen Anwendungen und Diensten kommunizieren können. Für die KI geht es darum, über ihr statisches Wissen hinauszugehen und zu einem dynamischen Agenten zu reifen, der aktuelle Informationen abrufen und passende Maßnahmen je nach Aufgabenstellung und Datenlage ergreifen kann.

Den Kern des MCP bildet ein **Client-Server-Modell**, bei dem ein **MSC-Server** als Bindeglied zwischen dem KI-Modell (als MCP-Host) und den von ihm angebotenen Diensten und Funktionen fungiert. Das KI-Modell muss nicht alle verfügbaren Funktionen der bereitgestellten MCP-Server a priori kennen, weil es die Server nach ihren individuellen Diensten und Fähigkeiten fragen und sich so dynamisch neue Möglichkeiten erschließen kann. Das ist ein zentraler Aspekt dieses Ansatzes.



Sind bei der Tool Execution KI-Funktionen im Spiel?

Nein, die Tool Execution ist ein deterministischer Prozess nach den Methoden der klassischen Informatik. Dieser Teil der Verarbeitungskette kommt vollständig ohne KI-Funktionen aus und ist daher auch frei von Unschärfen oder der Gefahr von „Halluzinationen“.

Ist die Tool Execution komplett reproduzierbar?

Ja, eben weil es sich um einen vollständig deterministischen Prozess handelt. Natürlich können sich zwischen zwei wiederholten Durchläufen aktuelle Bestände und Marktwerte ändern. Beim Bezug auf historische Kurs- und Bestandsdaten werden jedoch keinerlei Abweichungen zu vorherigen Ergebnissen auftreten.

Verlassen die ermittelten Daten im Rahmen der Tool Execution den Companion?

Nein, die ermittelten Daten werden bei der Abarbeitung des Ausführungsplans zwar von Tool zu Tool weitergereicht. Sie verbleiben auf diesem Weg aber permanent im Hoheitsbereich von First Cloud. Erst das Endergebnis bzw. die daraus produzierten Tabellen und Grafiken werden dem Nutzer präsentiert und dringen so nach außen.

Werden die Ausführungspläne protokolliert?

Ja, alles in der RAG Pipeline fließt ins Audit-Log, von der Nutzeranfrage und deren Auflösung im Query Processing über die erstellten Ausführungspläne und die Zwischenergebnisse bei der Tool Execution bis hin zu den präsentierten Endergebnissen. Alle diese Daten werden revisionssicher archiviert.

Wie wird die Sicherheit beim Datenzugriff gewährleistet?

Indem bei jedem Verarbeitungsschritt innerhalb der RAG Pipeline die User-ID des initierenden Nutzers mitgeführt wird. Die im Rahmen der Tool Execution ausgeführten Tools prüfen jeweils anhand der damit verbundenen Nutzerrechte (auf Gruppenebene), ob der gewünschte Zugriff auf die jeweiligen Datenbankinhalte oder andere Informationen gestattet ist.

Warum werden die First-internen Tools nicht auch über MCP angebunden, so wie die externen Tools?

Weil bei den internen Tools gar nicht unbedingt eine Auswahl durch das externe LLM gewünscht wird. Die große Anzahl an Tools würde außerdem den MCP-Server überfrachten.



„Wir ertrinken in Information, hungern aber nach Wissen.“

Rutherford D. Roger

Neugierig geworden?

Sie möchten mehr über First Cloud, den **Companion for First** und den Betrieb in der Fact Cloud erfahren? Dann schreiben Sie uns oder sprechen Sie direkt mit unserem Experten Aleksandar Ivezić. Er freut sich auf Ihre Anfrage:



Aleksandar Ivezić

a.ivezic@fact.de

+49 2131 777 238



Hauptsitz Neuss

Fact Informationssysteme & Consulting GmbH
Hellersbergstraße 11 | 41460 Neuss | Tel.: +49 2131 777-0

Standort Frankfurt am Main

Fact Informationssysteme & Consulting GmbH
Wilhelm-Leuschner-Straße 81 | 60329 Frankfurt am Main | Tel.: +49 69 8740313-121

E-Mail: mail@fact.de

Wir haben Ihnen dieses Whitepaper gerne zu unverbindlichen Informationszwecken überlassen. Bitte beachten Sie aber, dass die darin enthaltenen Informationen allgemeiner Natur sind und eine Beratung im konkreten Einzelfall nicht ersetzen können.

Die Fact Informationssysteme & Consulting GmbH hat diese Unterlage nach bestem Wissen erstellt und die Inhalte sorgfältig erarbeitet. Die Informationen werden ständig geprüft und aktualisiert. Gleichwohl können Fehler nicht ausgeschlossen werden. Bitte haben Sie deshalb Verständnis dafür, dass wir keine Garantie und/oder Haftung für die Aktualität, Richtigkeit und Vollständigkeit übernehmen. Infolgedessen haften wir nicht für direkte, indirekte, zufällige oder besondere Schäden, die Ihnen oder Dritten durch die Verwendung der Informationen dieser Unterlage entstehen.

Inhalt, Darstellung und Struktur dieser Unterlage sind urheberrechtlich geschützt und eine Nutzung, Verwendung, Reproduktion oder Weitergabe an Dritte – ganz oder teilweise – ist nur mit unserer ausdrücklichen vorherigen schriftlichen Zustimmung zulässig.

Öffentlich – Alle Rechte an den Inhalten dieses Whitepapers liegen bei der Fact Informationssysteme & Consulting GmbH.

Stand Januar 2026



**Folgen Sie uns
auf LinkedIn**



Alle im Whitepaper verwendeten Bilder wurden mit Unterstützung des KI-Tools Adobe Firefly generiert. Sie dienen ausschließlich der visuellen Veranschaulichung und haben keinen dokumentarischen Charakter.