



solvistas
ACADEMY. UPGRADE YOUR FUTURE.

DATA SCIENCE

Ausbildung zum
Data Scientist

Lehrgangsbeschreibung

Alles, was ein Data Scientist wissen muss!

... Und noch mehr!



Dieser Lehrgang bietet sowohl für Personen mit IT-Erfahrung als Spezialisten (Entwickler, Datenbankadministratoren, Systemengineers,..), als auch für Personen aus dem Fachbereich (Statistiker:in, BI-Specialist, Business-Analyst,...) eine umfassende Weiterentwicklung in Richtung neuer Standards und Themengebiete und ist die Grundlage für eine Entwicklung Richtung Data Science.

Dabei werden Fragen beantwortet, die Sie sich in Ihrer Tätigkeit bisher vielleicht noch gar nicht gestellt haben:

- Was ist der Unterschied zwischen Business Intelligence und Data Science?
- Wie kann Data Science sinnvoll angewandt werden?
- Habe ich als Data Scientist nun Expertise in IT oder doch eher in Statistik?

Oder aber auch technische Fragen wie z.B.: Wie funktioniert die Modellierung von Daten? Wie integriere ich Daten in Systeme? Wie analysiere und visualisiere ich Daten richtig? Welche Gesetze und Vorschriften sind dabei zu beachten?

Lernformen

Mitarbeiter:innen fordern und fördern.
An individuelle Bedürfnisse angepasst.

Classroom Trainings

- ✓ Unterschiedliche Methoden zur Festigung des Inhalts
- ✓ Praxisübungen aus realen Projekten
- ✓ Real Life Experts als Trainer:innen
- ✓ Reflexionen und die Bearbeitung konkreter Fragestellungen von Teilnehmer:innen



Video Trainings (online)

- ✓ Jederzeit verfügbar zum Selbststudium in kleinen Einheiten
- ✓ Übungsaufgaben zur Festigung des Inhalts inkl. Best Practices und Lösungsvorschlägen
- ✓ Ein Forum zum Austausch von Fragen und Ideen



Virtuelle (Classroom-) Trainings

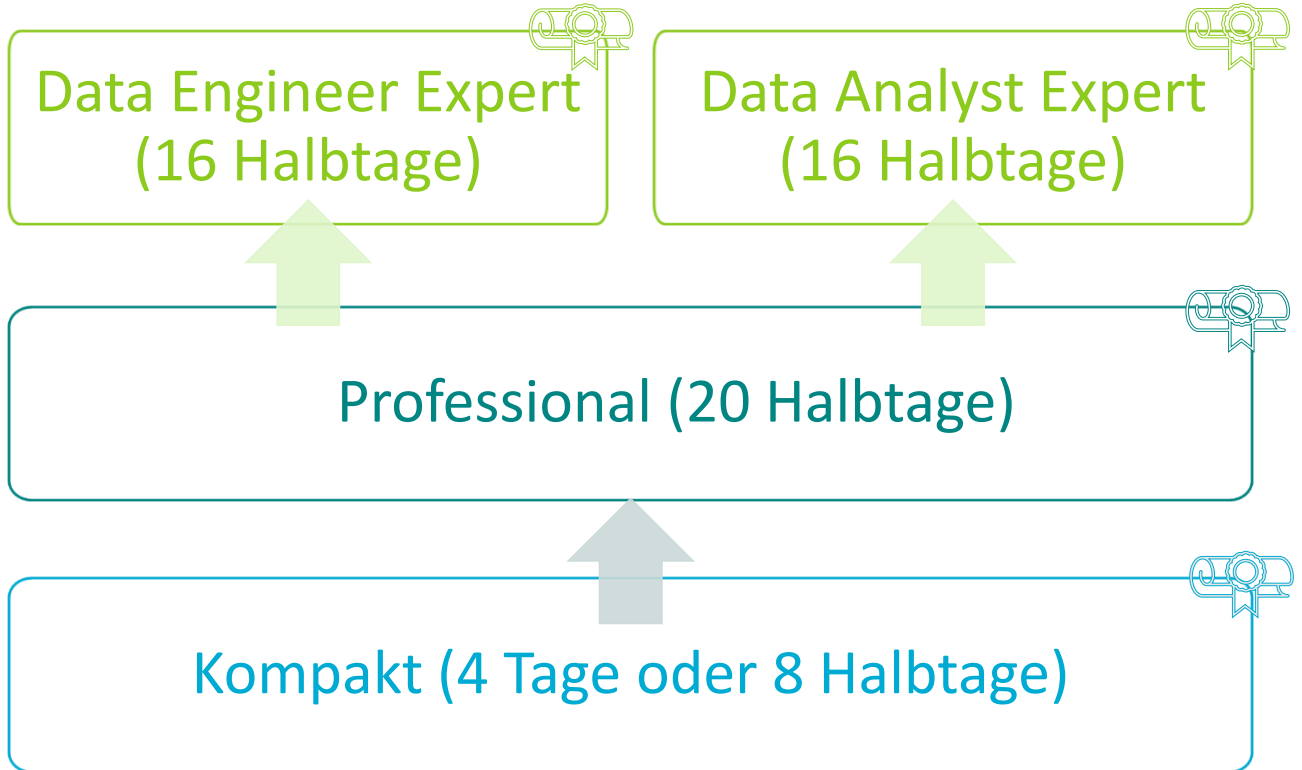
- ✓ Zu einem definierten Termin an einem beliebigen Ort
- ✓ Lernen durch gemeinsame Übungen und Reflexionen mit anderen Teilnehmer:innen
- ✓ Mit realen Trainer:innen, die durch das Training führen

Blended Learning

- ✓ Eine effizienzsteigernde Kombination verschiedener Methoden und Medien
- ✓ Zeitlich flexible Wissensvermittlung
- ✓ Wissensfestigung durch Reflexionen und Präsenzübungen
- ✓ Praktische Lösungen mit Coaching und Mentoring

Ausbildungspfad

Wissen – Verstehen - Anwenden



Lehrgangsbleitung

Inhaltliche Professionalität und persönliche Betreuung stehen im Vordergrund



Melissa Schmidt, MSc

Inhaltliche Leiterin der solvistas Academy

melissa.schmidt@solvistas.com

+43 664 2465 264

Sie ist seit vielen Jahren im IT-Bereich tätig mit dem Schwerpunkt Datenbanken, Visualisierung, Projektleitung und im konzeptionellen Bereich.

Dabei verantwortet Sie die Lehrgänge der solvistas Academy GmbH und legt dabei großen Wert auf fundierte Inhalte und deren Anwendung im praktischen sowie beruflichen Einsatz. Durch mehrere Jahre Erfahrung im IT - Management, weiß sie, worauf zu achten ist, um alle Teilnehmer:innen perfekt auf das Arbeitsumfeld vorzubereiten.

Sie ist ausgebildete Statistikerin, IT-Managerin und hat mehrere Zertifizierungen im Bereich Requirements Engineering, Data Science sowie Agilität.



Ing. Hans-Jörg Steffe, MSc MSc

Organisatorischer Leiter der solvistas Academy

hans-joerg.steffe@solvistas.com

+43 664 2465 254

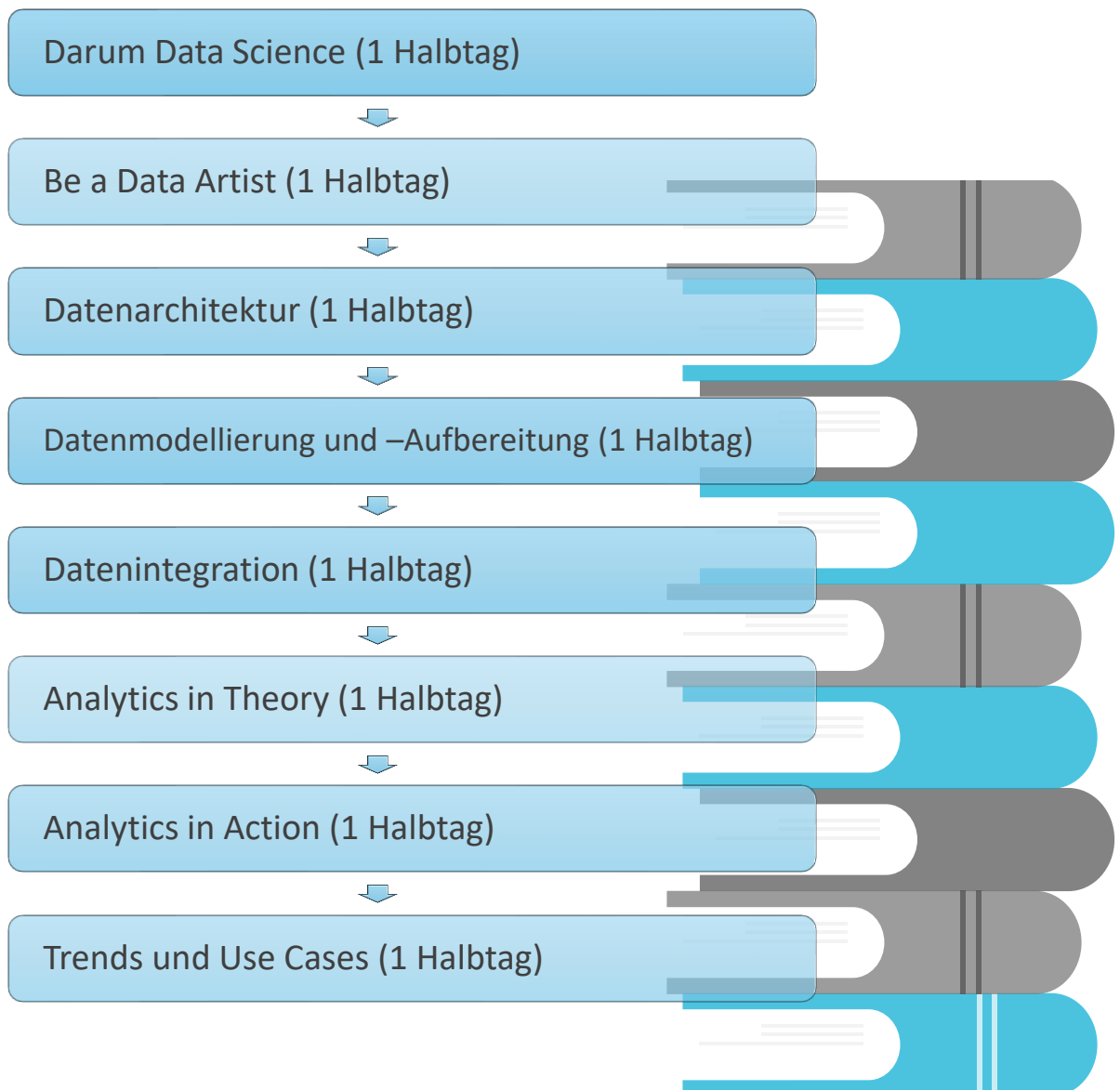
Seit über 20 Jahren in der IT-Beratung, der Aus- und Weiterbildung, der Personalentwicklung sowie der Mitarbeiterführung tätig.

Er ist Leiter mehrerer Lehrgänge und Seminare zu den Themen Digitalisierung, Business-Analyse/Requirements Engineering, Data Science sowie agile Herangehensweisen. Zudem unterstützt er Unternehmen und Führungskräfte dabei, Konzepte und Vorgehensweisen erfolgreich und nachhaltig einzufügen. Herr Steffe ist Systemischer Coach und Berater.

Hans-Jörg ist Mitglied des Vorstands des International Requirements Engineering Board (IREB) e.V. sowie Mitglied der Arbeitsgruppen agiles Requirements Engineering und Prüfungsgruppe. Er ist auch im Domain Steering Committees (DSCs) von ITEDAS® zur Zertifizierung von SCRUM Experten und Mitautor des itedas® Lehrplans.

Kompakt Level

8 Halb-Tage



Darum Data Science (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training bietet den optimalen Einstieg in die Welt von Data Science (DS). Sie erhalten in diesem Training einen Überblick über die wichtigsten Begriffe und ein Verständnis davon, wie ein Data Science Projekt idealtypisch verläuft. Nach diesem Training sind Sie in der Lage Data Scientists unterschiedliche Rollen zuzuschreiben und das Datenmanagement Ihres Unternehmens zu hinterfragen, sowie den potentiellen Mehrwert von DS anhand von konkreten Anwendungsfällen zu erkennen.

INHALT:

- Die Grundlagen von Data Science
- Der Mehrwert beim Einsatz von Data Science
- Data Science Rollen
- Das Berufsbild Data Scientist
- Die Unterschiede zwischen Data Lake, Data Engineering, Data Mining, Digitalisierung und Big Data
- Unterschied und Klärung der Begriffe IoT und Cloud Computing
- Data Science Projektablauf
- Wert der Daten verstehen

Be a Data Artist (1 Halbtage)

BESCHREIBUNG UND ZIELE:

In dem Training wird vermittelt, wie aus gut strukturierten Daten ein wirklicher Mehrwert erzeugt wird. Dabei lernen Sie wie man die Daten so aufbereitet, dass Sie von der Zielgruppe verstanden und eingesetzt werden können. Nach diesem Training sind Sie in der Lage mehrere Tools zur Visualisierung zu benennen, professionelles UX-Design zu beurteilen, und die Datenaufbereitung gewinnbringend in Ihrem Unternehmen einzusetzen. Sie haben mit diesem Modul die Basis für einen Data Artist gelegt.

INHALT:

- Grundlagen und Vorteile von Visualisierungen
- Allgemeine Designrichtlinien
- UX-Design Richtlinien
- Reports und Dashboard Designs
- Analyse von Reports

Datenarchitektur (1 Halbtage)

BESCHREIBUNG UND ZIELE:

In diesem Training erhalten Sie einen Überblick über verschiedene Architektur-Ansätze im Bereich Data Science. Die behandelten Themen reichen von einer klassischen mehrschichtigen Data Warehouse (DWH) Architektur über Lösungsansätze beim Umgang mit Big Data bis hin zu Cloud-basierten Architekturen. Im Zuge des Trainings werden unter anderem die Begriffe Data Warehouse und Data Lake sowie deren Unterschiede näher betrachtet. Sie werden dabei in die Lage versetzt, eine passende Architektur und Strategie für Ihr Projekt zu wählen.

INHALT:

- Überblick über das Thema Datenarchitektur (Big Picture Data Architecture)
- Klassische DWH-Architektur
- Herausforderungen und Lösungsansätze im Big Data Bereich
- Unterschiede zwischen Data Lakes, Data Hubs, Data Meshes und Data Fabrics
- Vorteile des Cloud Computing
- Die Schritte von der Vision bis zur Umsetzung (Checkliste zum Projekt Setup)

Datenmodellierung und Aufbereitung (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training bietet die optimale Vorbereitung in der praktischen Anwendung der Datenaufbereitung, sowie den theoretischen Hintergrund zu wichtigen Themen und Theorien in diesem Bereich. Sie erhalten in diesem Training Know-How über die Modellierung von Daten, deren Historisierung Change Data Capture (CDC), sowie Qualitätssteigerung. Ebenso erhalten Sie einen Überblick über die wichtigsten Begriffe und ein Verständnis davon, wie sich Extract, Transform, Load (ETL) von Extract, Load, Transform (ELT) abgrenzt. Nach diesem Training sind Sie in der Lage verschiedene Ansätze der Modellierung anzuwenden und Daten für weitere Analyse oder Visualisierungstätigkeiten geeignet aufzubereiten.

INHALT:

- DWH Modellierungsgrundlagen Inmon, Kimball, Linstedt
- Was ist der Unterschied zw. ETL und ELT?
- Wie historisiert man Daten?
- Welche Arten der Datenbereinigung gibt es, um die Qualität zu steigern?
- Was ist CDC und welche Implementierungsarten gibt es?

Datenintegration (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training beschäftigt sich mit der Integration von Daten in unterschiedlichen Umgebungen. Sie erhalten Einblick in die Welt der NoSQL Datenbanken und den Umgang mit ebendiesen. Ein weiteres Thema ist die Datenintegration in der Cloud. Im Rahmen des Erstellens von ETL-Strecken wird mit der Überprüfung der Datenqualität und Transformationsqualität experimentiert und gängige Konzepte besprochen. Ein weiterer Punkt in diesem Training ist der Einsatz geeigneter Scheduling Konzepte für den regelmäßigen Datenintegrationsprozess.

INHALT:

- NoSQL
- Datenintegration mit NoSQL
- Datenintegration im Big Data Umfeld und in Cloud Umgebungen
- ETL-Strecken bauen und worauf dabei zu achten ist
- Die Qualität bestehender Datenintegrations-Strecken einordnen
- Scheduling Möglichkeiten

Analytics in Theory (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training betrachtet was mit Begriffen wie Machine Learning und Artificial Intelligence eigentlich genau gemeint ist und wie sind diese im Data Science Kontext einzuordnen sind. Zwischen welchen Arten von Zielsetzungen und Aufgaben unterschieden wird, ist ein weiterer Schwerpunkt. Methoden und Algorithmen, die zum Erreichen dieser Tasks verwendet werden, können Sie nach dem Training konkret einordnen und verstehen deren Funktionsweise. Sie lernen auch unterschiedliche Techniken aus Machine Learning und Statistik kennen und wissen, wofür diese verwendet werden.

INHALT:

- Grundlagen von Machine Learning und Artificial Intelligence
- Arten von Zielsetzungen und Aufgaben im Bereich Data Analytics
- Methoden und Algorithmen
- Einsatz von Methoden aus dem Machine Learning und deren Einsatz

Analytics in Action (1 Halbttag)

BESCHREIBUNG UND ZIELE:

In diesem Training lernen Sie einige der größten Schwerpunkte von Data Science in der Umsetzung kennen: Was sind die gängigsten Tools für Analytics und Machine Learning (ML)? Welche Vor- und Nachteile haben diese? Gemeinsam betrachten, analysieren und bewerten Sie einfache Use Cases und bearbeiten diese mit Hilfe von in Python. Mithilfe von Machine Learning Algorithmen werden konkrete Fragen aus der Praxis beantwortet. Zusätzlich erhalten Sie einen Überblick über Verbesserungsmöglichkeiten (Model-Tuning, Variablenselektion) für die Datenanalyse. Nach diesem Training kennen Sie den ganzen Prozess der Datenanalyse mit den darin vorkommenden Schritten und können durch Datenanalyse in Ihrem Projekt Mehrwert generieren.

INHALT:

- Überblick über Tools für Datenanalyse
- Einsatz von ML-Algorithmen
- Einfaches Anwenden von Algorithmen in Python
- Tuning - wie können bessere Ergebnisse erzielt werden?
- Ablauf einer Datenanalyse

Trends und Use Cases (1 Halbttag)

BESCHREIBUNG UND ZIELE:

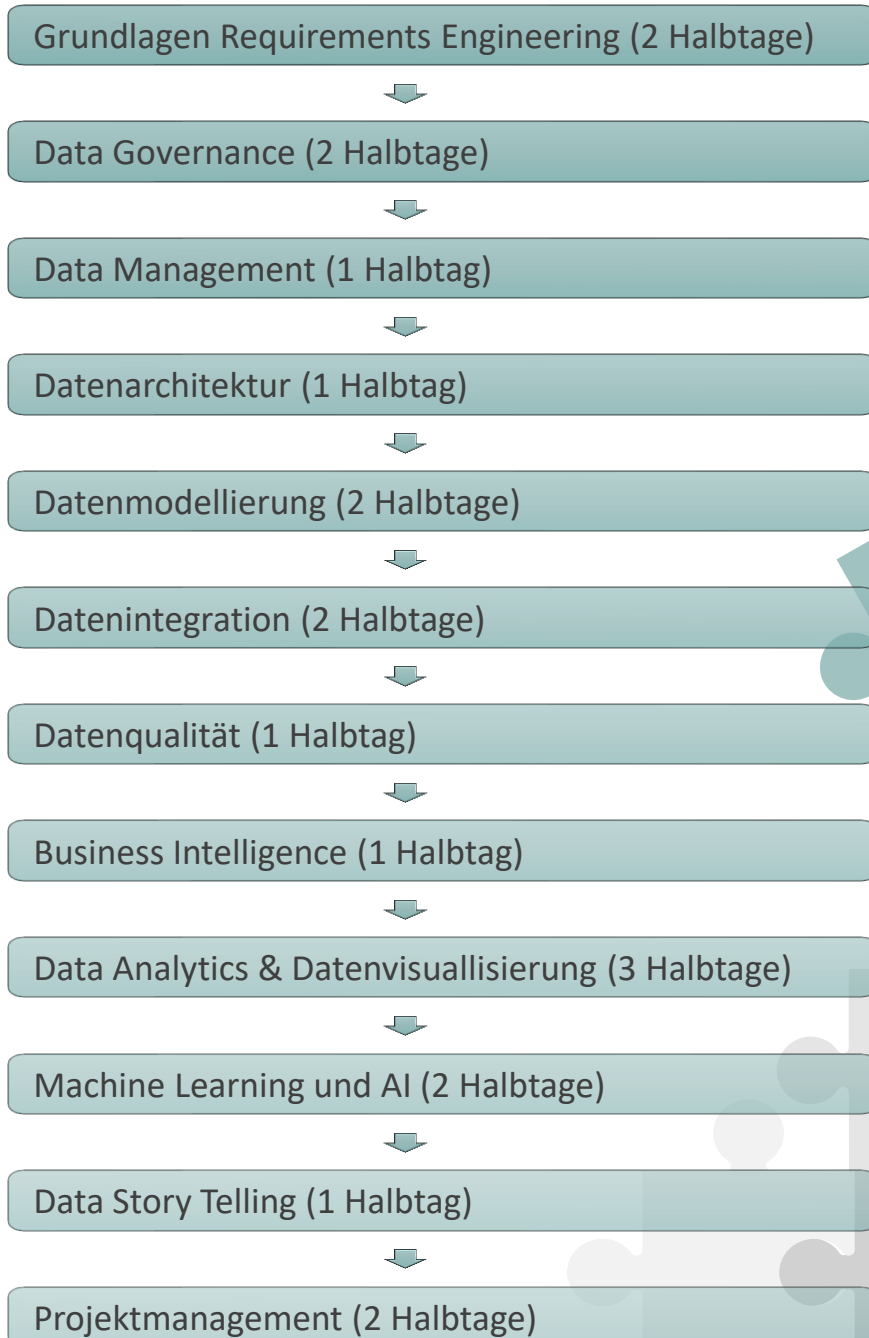
Trends in DataScience sind schnelllebig, um hier Up-to-Date zu bleiben bietet dieses Training "Trend und Use Cases" die Einführung in die diversen Trends, die aktuell in der Data Science Welt diskutiert werden. Von Augmented Reality über Artificial Intelligence bis hin zu neuen Methoden der Datenspeicherung. Nichts ist so wichtig, wie aus der Praxis zu lernen. Die behandelten Use Cases decken Branchen wie Industrie, Gesundheit und Handel ab und bieten Ihnen einen Überblick über "Lessons Learned" und Projektabläufe wie aus dem Bilderbuch.

INHALT:

- aktuelle/ neue Schlagwörter aus der Welt von Data Science
- Wichtigkeit der Datenstrategie
- Aktuelle Anwendungsfälle von Data Science
- neue Arbeitsformen durch Data Science

Professional Level

20 Halbtage



Grundlagen Requirements Engineering (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Grundlagen Requirements Engineering bietet allen in einem Data Science Projekt Beteiligten die Möglichkeit, einen Einstieg in die Welt des Anforderungsmanagement zu erhalten. Sie lernen dabei nicht nur von Anforderungserhebung, Strukturierung und Verwaltung, sondern auch wie die Steuerung der Stakeholder:innen erfolgreich funktionieren kann. Dieses Training orientiert sich am offiziellen Standard des International Requirements Engineering Boards (IREB).

INHALT:

- Grundlegende Begriffe von Requirements Engineering
- Arten an Anforderungen
- Stakeholdermanagement und Kontextabgrenzung
- Quellen von Anforderungen und Erhebungsmethoden
- Die Strukturierung von Anforderungen
- Anforderungsdokumente

Data Governance (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training bietet eine Einführung in Data Governance. Dabei werden die Datensicherheit, die Datennutzung, Datenownership sowie die Richtlinien zu Daten genauer betrachtet. Nach dem Training sind Sie in der Lage, Datenschutzrichtlinien richtig anzuwenden, verstehen die rechtlichen Anforderungen an Daten und wissen, wie ein Genehmigungsprozess in einem Unternehmen verlaufen kann.

INHALT:

- Unterschiedliche Nutzungsmöglichkeiten von Daten verstehen
- Datennutzungskatalog einsetzen
- Data Compliance: Was sind die rechtlichen Anforderungen an Daten?
- Datenschutzrichtlinien
- Unternehmensgenehmigungsprozesse für Daten

Data Management (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Data Management umfasst ein breites Spektrum an Themenbereichen, von denen einige als eigene Module des Data Science Lehrgangs angeboten werden. In diesem Training liegt der Fokus auf Grunddaten-, Referenzdaten- und Metadatenmanagement. Sie lernen die Unterschiede zwischen Stamm-, Bewegungs-, Referenz- und Metadaten kennen und erhalten einen Einblick in den professionellen Umgang mit diesen Datenarten. Dieses Training orientiert sich an dem offiziellen Lehrplan zum Data Excellence Professional (www.cdxp.org). Nach diesem Training sind Sie in der Lage ein erstes Data Excellence Konzept zu entwickeln.

INHALT:

- Überblick über die Teilgebiete des Data Managements
- Grunddatenmanagement (Stamm- und Bewegungsdaten)
- Metadatenmanagement
- Referenzdatenmanagement

Datenarchitektur (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Eine moderne Datenarchitektur muss in der Lage sein, sämtliche fachlichen Anforderungen rund um das Thema Daten möglichst effizient erfüllen zu können. In diesem Training werden die Komponenten einer modernen Datenarchitektur vorgestellt und Konzepte zur Rollen- und Rechtevergabe sowie Speicherkonzepte näher betrachtet. Nach dem Training sind Sie bereit, für Ihr Projekt die optimale Datenarchitektur zu definieren.

INHALT:

- Komponenten einer modernen Datenarchitektur
- Was gibt es bezüglich Schnittstellen zu beachten?
- Erstellung eines Architekturbilds
- Welche Konzepte gibt es bzgl. Rollen- und Rechtevergabe?
- Speicherkonzepte: was, wo wie lange speichern? Zentral/verstreut?
- Lizenzmodelle
- Wie viel kostet der Betrieb pro Jahr? (Anzahl User, Anzahl Umgebungen)

Datenmodellierung (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Von einer Kommunikationsgrundlage im Zuge der Anforderungsanalyse bis hin zur technischen Umsetzung spielen Datenmodelle in jedem Data Science Projekt eine wichtige Rolle. In diesem Training erhalten Sie einen Einblick in die Definition fachlicher, logischer und physischer Datenmodelle. Nach dem Training sind Sie im Stande, unterschiedliche Modellierungstechniken, die auf dem neuesten Stand des Praxiseinsatzes sind, zu unterscheiden, den jeweiligen Mehrwert zu erkennen und in Ihrem Projekt anzuwenden.

INHALT:

- Unterschiede zwischen fachlichem, logischem und physischem Modell
- 3. Normalform
- Data Vault 2.0 Modellierung
- Dimensionale Modellierung
- Modellierung semi-strukturierter Daten (z.B. XML-Schema- Definitionen)
- Überblick über verschiedene Tools
- Skalierbarkeit und paralleles Arbeiten mit Daten
- Der agile Aufbau von Datenmodellen

Datenintegration Hands On (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Im Training Datenintegration Hands On werden die häufigsten Vorgänge der Datenintegration in realitätsnahen Anwendungsbeispielen erarbeitet und gängige Vorgehensweisen erlernt. Inhalte sind unter anderem die Implementierung einer Zeitdimension und auch die nötigen Schritte zur Berechnung von Kennzahlen im Verlauf der Zeit. Was hierbei in der Datenintegration zu beachten ist, erlernen Sie in diesem Modul. Weiterführend wird die Integration unterschiedlicher Quellen, sowie unterschiedlicher Formate wie XML oder Json beleuchtet.

INHALT:

- Zeitdimensionen implementieren
- Berechnungen auf Grund der Zeitdimension - Kennzahlen Jahresvergleich
- Datenbeladung Flat Table
- Erweiterung um Historisierung
- Unterschiedliche Quellen anbinden
- Json oder XML File verarbeiten
- Umgang mit fehlenden oder mehrfachen Daten
- Validierung von Business Keys und deren Integration
- Big Data Techniken

Datenqualität (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Wenn es um die Verarbeitung von Daten und den Informationsgewinn aus diesen geht, ist Datenqualität essentiell, um validierte Aussagen treffen zu können. Nach dem Prinzip “garbage in garbage out” ist es notwendig die Daten passend aufzubereiten und die nötigen Qualitätsindikatoren auszuwählen. Sie erfahren in diesem Modul nicht nur etwas über das Datenqualitätsmanagement, sondern beschäftigen sich auch mit der Ermittlung und Auswertung unterschiedlicher Qualitätskennzahlen. Im Abschluss werden noch mögliche Tools, um die Datenqualität zu wahren, vorgestellt.

INHALT:

- Motivation für Datenqualität
- Datenqualitätsindikatoren (DQI)
- DQIs ermitteln und auswerten
- Datenqualitätsmanagement
- Robustheit von Kennzahlen und Modellen/Algorithmen
- Augmented Data Quality
- Tools

Business Intelligence (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Ein schneller Überblick über die Daten ist ihr Ziel? Wie sollen Sie die Daten, die Sie täglich brauchen, nutzen? In diesem Kurs lernen Sie zu verstehen, wie sie im Bereich Data Science aktuell aufgestellt sind und welche Potenziale in Ihren Unternehmensdaten stecken. Sie erhalten auch anwendungsorientierte Insights und lernen die Probleme der diversen Usergruppen kennen.

INHALT:

- BI Definition
- Datenlandschaft und Workflows
- Praktische Tools und deren Einsatz
- Bericht, Dashboard, Cube
- Abfragen aus einem Cube

Data Analytics & Datenvisualisierung Hands On (3 Halbtage)

BESCHREIBUNG UND ZIELE:

Sie haben viele Daten, wissen auch, dass in Ihren Daten viel Information steckt, aber haben nicht das richtige Werkzeug, um doch alles an Information zu gewinnen? In diesem Training lernen sie von der Theorie bis zur Praxis die wichtigsten Methoden zur Analyse und Auswertung von Daten. Sie probieren sich in Advanced Analytics und werden mit Hilfe von Statistik einen Blick in die Zukunft werfen. Der zweite Halbtage beschäftigt sich vor allem mit der Struktur der Daten, um so noch mehr Erkenntnisse zu gewinnen. Data Mining ist hier das Schlagwort. Dies ist in aller Munde, und Sie lernen es zu verstehen!

Das Wichtigste ist es, das Gelernte auch anzuwenden und so wird der letzte Halbtage genutzt, um Ihre neu erworbenen Kenntnisse von der explorativen Datenanalyse über passende Analysemethoden bis hin zur Visualisierung der Ergebnisse praktisch anzuwenden.

INHALT:

- Begrifflichkeiten zu BI und Analytics
- Techniken und Prozesse im Data Mining
- Beispielhafte Anwendungen
- Muster in Daten finden
- Passende Analysemethoden anwenden
- Explorative Datenanalyse und deskriptive Statistiken
- Ergebnisse visualisieren

Machine Learning und AI (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Machine Learning und Artificial Intelligence sind sehr leistungsstarke Felder, mit denen man den Wert von Daten erschließen kann. Mit Hilfe dieser Methodiken ist es möglich, das Denken in unterschiedlichsten Branchen und Geschäftsmodellen zu verändern und zu verbessern. Als datengesteuertes Unternehmen lässt man die Daten für sich arbeiten und Entscheidungen treffen. Aber wie genau wird man so ein datengesteuertes Unternehmen? Welche Unternehmen sind Vorreiter? Neben all diesen allgemeinen Themen werden in dieser Schulung ihre Kenntnisse zu Supervised und Unsupervised Learning vertieft und an Beispielen angewandt. ML und AI kommen ohne Gradient Descent und Loss Functions nicht aus, so auch nicht diese Schulung. In dieser Schulung erhalten sie also einen ersten Einstieg in Richtung AI fürs eigene Unternehmen.

INHALT:

- Begrifflichkeiten / Überschneidungen / Abgrenzungen
- Machine Learning Tasks
- Gradient Descent & Loss
- Neurale Netzwerke
- Best Practices
- Social Impact of AI
- Einstieg in AI und ML im eigenen Unternehmen

Data Story Telling (1 Halbttag)

BESCHREIBUNG UND ZIELE:

Was bringen Ihnen die Erkenntnisse in Ihren Daten, wenn die Personen, die Sie davon überzeugen wollen, ihnen aufgrund der Abstraktheit des Themas nicht folgen können? Data Science, Machine Learning und Digitalisierung sind für Nicht-Eingeweihte oftmals eine unverständliche Black Box. Genau hier setzt Data Storytelling an und versucht die Welt der Daten mit der Welt der fachlichen Thematik zu kombinieren und eine Geschichte zu schaffen, die Ihre Stakeholder von ihren Ergebnissen überzeugen. Zum einen erlernen Sie die Präsentationserstellung, um Ihre Ergebnisse und Erkenntnisse gut aufzubereiten. Danach wird die Datastory behandelt, in der es darum geht, warum Geschichten wichtig sind. Dabei lernen Sie ein Storyboard anhand eines Beispiels kennen. Die zu überzeugenden Stakeholder werden zum einen vom Gehörten und zum anderen auch durch visuelle Reize überzeugt. Welche Reize sind wann am besten eingesetzt?

INHALT:

- Data Storytellings anwenden
- Datengetriebene Ergebnisse effektiv in Grafiken und Präsentationsfolien umwandeln
- Storyboards erstellen
- Datensätze effektiv und leicht verständlich präsentieren

Projektmanagement Basics (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training bietet einen optimalen Einstieg in die Grundlagen des Projektmanagements und der Organisation von Projektteams. Sie erhalten sowohl einen Einblick in die Phasen eines Projekts, welche Rollen dabei eingebunden sind, als auch wann ein Projekt startet und wie es abgeschlossen wird. Nach dem Training sind Sie in der Lage, einen Projektauftrag anzunehmen, das Projekt zu planen und ordnungsgemäß abzuschließen.

INHALT:

- Grundlagen Projektmanagement
- Projektphasen
- Projektrollen
- Der Projektauftrag
- Wann ist ein Projekt zu Ende?
- Konflikte erkennen

Expert Level Data Engineering

16 Halbtage

Data Engineering

Projektleitung und Anforderungsmanagement (4 Halbtage)



Hochskalierbare Architekturen (2 Halbtage)



Die professionelle Infrastruktur (4 Halbtage)



Data Engineering Hands On (6 Halbtage)

Projektleitung und Anforderungsmanagement (4 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training baut auf „Projektmanagement Basics“ auf. Sie erhalten in diesem weiterführenden Training einen Überblick über Anforderungsmanagement, sowie eine große Anzahl an Anforderungen zu verwalten und zu priorisieren. Ebenso lernen Sie die wichtigsten Werkzeuge zur Anforderungsverwaltung kennen. Nach diesem Training sind Sie in der Lage ein Projekt zu planen und Risiken abzuschätzen, Sie wissen zu welchen Konflikten es kommen kann und lernen die Konfliktlösungsmöglichkeiten anzuwenden.

INHALT:

- Eine große Anzahl an Anforderungen verwalten
- Priorisieren von Anforderungen
- Grundlagen des Projektmanagement
- Projektsteuerung, Fokus auf Kostenkontrolle und Projektcontrolling
- Konfliktmanagement, Konfliktarten
- Konfliktlösungsmöglichkeiten verstehen

Hochskalierbare Architekturen (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training zeigt weiterführend den Mehrwert von Data Science, Big Data und Cloud Lösungen und komplexe Anwendungen im Bereich der Datenarchitektur. Sie erhalten einen Einblick in den Aufbau unterschiedlicher Architekturen und deren Anwendungsfälle. Nach dem Training können Sie abschätzen, was es heißt, mit hochskalierbaren Daten zu arbeiten. Außerdem lernen Sie die Orchestrierung von Container Architekturen und deren Einsatzmöglichkeiten kennen.

INHALT:

- Mehrwert von Data Science
- Architekturen in Big Data und Cloud Umgebungen
- Weiterführende Datenarchitektur
- Container Orchestrierung

Die professionelle Infrastruktur (4 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training bietet den optimalen Einstieg in die Welt der Infrastruktur und Virtualisierung im Data Science Bereich. Sie erhalten in diesem Training eine Einführung in das Konzept von Multi Tenant und beschäftigen sich mit den notwendigen Vorbereitungen in der Infrastruktur für reibungslose Deployments. Zusätzlich wird der Unterschied und die Vorteile von pervasive und embedded Systemen gemeinsam erarbeitet. Nach diesem Training sind Sie in der Lage das Scheduling und Monitoring in unterschiedlichen Betriebssystemen aufzusetzen. Zusätzlich werden Best Practises für Memory Management, Error Handling und Bottle Neck Analysen vorgestellt.

INHALT:

- Verstehen eines Multi Tenant Ansatzes
- Infrastruktur Deployment
- Pervasive/Embedded Systems
- Scheduling und Monitoring in unterschiedlichen Betriebssysteme (Windows, Linux)
- Virtualisierung unter Linux/Microsoft/Container
- Memory Management
- Error Handling und Bottle Neck Analyse

Data Engineering Hands On (6 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training besteht aus drei großen Themen, dem Kennenlernen und Vertiefen von Methoden im Bereich Data Lake, der Datenintegration und dem Aufbau eines Data Warehouse. Sie vertiefen ihr Wissen über die Verarbeitung von Datenströmen in Echtzeit, Cloud Computing, sowie verteilte und parallele Systeme. Dieses Training ermöglicht Ihnen, eine analytische Datenverarbeitung umzusetzen, und das 4-Schichten Modell zu verstehen. Ebenso können Sie können Ihre Daten in einem Data Lake mit verschiedenen Methoden speichern und aufbereiten.

INHALT:

- Rolle Data Engineer
- Grundlagen von Internet of Things (IoT)
- Big Data und warum es immer falsch verstanden wird
- Azure/AWS - Entwicklung und Implementierung
- Cloud Computing
- Kommunikationssysteme/Tiering
- CDC, Historisierung/Versionierung
- DWH Aufbau – 4-Schichten Modell
- Deployment Konzepte

Expert Level Data Analytics

16 Halbtage

Data Analytics



Projektleitung und Anforderungsmanagement (4 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training baut auf „Projektmanagement Basics“ auf. Sie erhalten in diesem weiterführenden Training einen Überblick über Anforderungsmanagement, sowie eine große Anzahl an Anforderungen zu verwalten und zu priorisieren. Ebenso lernen Sie die wichtigsten Werkzeuge zur Anforderungsverwaltung kennen. Nach diesem Training sind Sie in der Lage ein Projekt zu planen und Risiken abzuschätzen. Sie wissen zu welchen Konflikten es kommen kann und lernen die Konfliktlösungsmöglichkeiten anzuwenden.

INHALT:

- Eine große Anzahl an Anforderungen verwalten
- Priorisieren von Anforderungen
- Grundlagen des Projektmanagement
- Projektsteuerung, Fokus auf Kostenkontrolle und Projektcontrolling
- Konfliktmanagement, Konfliktarten
- Konfliktlösungsmöglichkeiten verstehen

Be a Data Artist & Visual Analytics (4 Halbtage)

BESCHREIBUNG UND ZIELE:

In diesem Training werden sowohl verschiedenste Methoden zur Visualisierung von unterschiedlichsten Daten, sowie Designrichtlinien und Anwendungen, mit welchen man Daten für User bereit stellen kann, vermittelt. Sie erhalten einen Einblick in Geodaten und lernen die dazugehörigen speziellen Datenformate sowie Standards kennen. Nach dem Training sind Sie bereit, unterschiedliche Visualisierungstechniken richtig durch UX-Design einzusetzen. Außerdem können Sie Dashboards erstellen, analysieren und Visualisierungskonzepte hinterfragen.

INHALT:

- Vor- und Nachteile der Datenvisualisierung hinterfragen, bewerten.
- Geodaten – Datenformate und - standards
- UX-Design - Colour, Editorial Thinking, Designprozess
- Komponenten (Visuelles Design, Informationsarchitektur, Interaktionsdesign, Usability)
- Dashboards – erstellen, analysieren, hinterfragen, bewerten.
- Grundlagen der visuellen Wahrnehmung
- Visualisierungsmethoden (Vor-/Nachteile/ Beeinflussung)
- Data Life Cycle
- Unterschied zwischen human-centered design/user-centered design

Analyse all the Data! (3 Halbtage)

BESCHREIBUNG UND ZIELE:

Dieses Training legt den Fokus auf Machine Learning, Statistik und Artificial Intelligence. Sie erhalten hier vertiefte Kenntnisse über explorative Datenanalyse, Zeitreihenanalyse und verschiedene Prognoseverfahren. Auch für die Analyse von unstrukturierten Daten lernen Sie hier die passenden Methoden und Algorithmen kennen. Nach dem Training können Sie die gängigsten Methoden, um neue Datenströme zu clustern und Prognosen zu erstellen, anwenden. Außerdem kennen Sie die Grundlagen von Real Time Analytics und dessen Einsatzgebiete und Anwendungsbeispiele, genauso wie verschiedene Gütekriterien und die Anwendung von statistischen Tests. Außerdem werden Sie verstehen, wie Sie Text Mining und Natural Language Processing (NLP) voneinander abgrenzen und was mit diesen Methoden möglich ist. Nach diesem Training sind Sie zudem in der Lage Text Analytics und Information Extraction zu verstehen und im Kontext von Machine Learning Projekten einzuordnen.

INHALT:

- Zeitreihenanalyse
- Lebensdaueranalyse
- Versuchsplanung
- Real Time Analytics
- Evaluation von Modellen und statistischen Tests
- Gütekriterien – Aussagekraft von Machine Learning und statistischen Ergebnissen bewerten
- Gradient Descent und Loss Functions
- Perceptron Theorie – der Beginn von neuronalen Netzen
- Einführung in das Text Mining
- Maschinelles Lernen für die Textanalyse
- Semantische Netze
- Überblick Natural Language Processing und Information Extraction

Data Analytics – do it yourself! (2 Halbtage)

BESCHREIBUNG UND ZIELE:

Sie haben in den vorherigen Modulen die Theorie erlernt, nun geht es darum das Erlernte in die Praxis umzusetzen. Ganz nach dem Motto: „Bring your own dataset“, sollen passende und vorbereitete Daten mit Data Analytics verarbeitet und analysiert werden. Mit ihren eigenen Daten werden Sie mittels Python Ihre Daten analysieren und in weiterer Folge in PowerBI selbst grafisch aufbereiten und darstellen. Ziel des Moduls ist es, eigene – neue - Ergebnisse zu erhalten und Ihr gewonnenes Wissen aus den anderen Modulen an realen Daten anzuwenden.

INHALT:

- Anwendung diverser Funktionen in Python
- Anwendung der gelernten Methoden (Analyse und Visualisierung)
- Interpretation und Präsentation der Ergebnisse

Tech'n'Tools (1 Halbtage)

BESCHREIBUNG UND ZIELE:

Wie der Name des Moduls bereits sagt, liegt der Fokus dieses Halbtages auf Techniken und Tools. Analytics in der Cloud funktioniert etwas anders, als die klassische Analyse der Daten. Genau diese Unterschiede werden Ihnen in diesem Modul näher gebracht. Sie lernen über Python und R, zwei Programmiersprachen, die viele Packages zur Datenanalyse zur Verfügung stellen. Ein weiterer sehr wichtiger Aspekt in der Umsetzung von Data Analytics ist die Performanceoptimierung. Hier lernen Sie wie Ihre Analysen schneller und genauer werden können und durch Performanceoptimierung und Tuning auf ein neues Level gehoben werden.

INHALT:

- Python, wichtigste Libraries
- R, wichtigste Libraries
- Einführung in Analytics in der Cloud
- Performanceoptimierung



solvistas Academy GmbH
Graben 18
4020 Linz
+43 732 27 27 27 - 400
academy@solvistas.com

© solvistas Academy GmbH