

Entwicklung eines Informationsbausteines für das Data-Mining der Post CH AG

Studiengang: MAS Data Science

Durch die weltweit zunehmende Digitalisierung in den unterschiedlichsten Bereichen und Branchen fallen immer grössere Mengen an Daten an. So auch bei der Schweizerischen Post, die sich zunehmend mit dem möglichen Nutzen und den neuen Erkenntnissen beschäftigt, die daraus entstehen könnten. Denn dass sich aus den gesammelten Daten ein Mehrwert ergeben kann, ist nicht mehr von der Hand zu weisen.

Ausgangslage und Zielsetzung

Durch den Aufbau einer Analytics-Plattform, ist innerhalb der Post der Begriff der Informationsbausteine entstanden. Das allgemeine Verständnis und eine einheitliche Definition dafür sind durch die Aktualität noch nicht vorhanden.

Deshalb soll in dieser Arbeit eine Definitionsgrundlage geschaffen werden. Das Hauptziel der Arbeit konzentriert sich aber auf den Aufbau eines solchen Bausteines.

Umsetzung

Der Begriff und das Konzept der Informationsbausteine, in Zusammenhang mit Data Analytics, findet sich heute in der Fachliteratur mit dieser Namensgebung nicht. Deshalb wurden zuerst bekannte Methoden und Konzepte aus dem Umfeld von (Big-Data-) Analytics-Plattformen, Data Warehouse und Business Intelligence untersucht. Informationen über die Methode und die Idee der Informationsbausteine wurden durch Interviews mit unterschiedlichen Personen aus dem engen Umfeld der neuen Analytics-Plattform gewonnen. Nun war es möglich, die Gemeinsamkeiten und Unterschiede zwischen den klassischen Konzepten und dem der Informationsbausteine aufzuzeigen. Die praktische Umsetzung befasste sich mit einer Data-Mining-Aufgabe aus dem Bereich der Klassifikation.

Ergebnis

Das Grundkonzept der Informationsbausteine ist nicht neu, sondern stark von klassischen Business-Intelligence-, Data-Warehouse- und Data-Lake-Ansätzen abgeleitet. Speziell der Business-Intelligence-Aspekt, bei welchem mittels Datenanalysen neue Erkenntnisse gewonnen werden, ist ein zentrales Element. Während Business Intelligence jedoch dazu eingesetzt wird, durch die neuen Erkenntnisse eine Unterstützung bei operativen und strategischen Entscheidungen zu schaffen, eignen sich die Informationsbausteine besonders dazu, neues Wissen zu generieren.

Der wesentliche Unterschied zwischen den klassischen Konzepten und den Informationsbausteinen ist die Kapselung der Informationen als eigene Datenobjekte, wobei diese als wiederverwendbare «Services» mit wertschöpfendem Charakter betrachtet werden können. Der systemische Aufbau ist vergleichbar mit einem Domänen-Datawarehouse.

Bei der durchgeführten Klassifikation konnte ein sehr gutes Ergebnis erzielt werden. Als abschliessendes Ergebnis wurden die Erkenntnisse mit anderen Informationen verknüpft und in einer Informationsgrafik visualisiert. Wodurch auch wieder neues Wissen entstanden ist.



Daniel Suter