

# Anforderungsanalyse Weber AG

1	Inhaltsverzeichnis	
1	Einleitung	2
2	Kundenvorgaben:	2
2.1	Verkäufer	2
2.2	Objekte	2
2.3	Lage	3
2.4	Kunden	3
2.5	Kudentypen	3
2.6	Kontakte mit den Kunden	3
2.7	Kundenwünsche	3
2.8	Besichtigungen	4
2.9	Exposé	4
2.10	Kauf	4
3	Offene Fragen 1 (Kundentermin 06.04.2021)	6
3.1	Ergebnis	6
3.2	Lead Management	7
3.2.1	Felder bei Leads Interessenten:	7
3.2.2	Felder bei Leads Verkäufer:	7
3.2.3	Felder im bestehenden Formular Kaufinteressent:	8
3.2.4	Felder zu Kaufinteressent Wunsch:	8
3.2.5	Felder im bestehenden Internetformular Verkäufer:	8
3.2.6	Kommunikationskanäle:	9
3.2.7	Mögliche Aktionen bei Aufnahme Kundenanforderungen:	9
3.3	Use Cases	9
3.3.1	Use Case Altdaten einlesen	10
3.3.2	Use Case Lead Daten laden	10
3.3.3	Use Case Lead Daten automatisch laden	11
4	Offene Fragen 2 (Kundentermin 07.04.2021)	11
4.1	Ergebnis	11
5	Das Modell der operativen Datenbank	12
5.1	Das ERM-Modell	12
5.2	TabellendefinitionVorgaben	12
6	Datawarehouse	12

6.1	DWH Architektur	12
6.1.1	DWH Architektur für Weber AG, Abteilung Real Estate	10
6.2	Das Modell	12
7	Import von externen Daten (Leads)	12
7.1	Schematische Darstellung	12
7.2	Vorgaben	12

## 1 Einleitung

Aufgrund der Covid-19 Situation ist das Unternehmen wirtschaftlich belastet und möchte prüfen, ob es auch in andere Geschäftsmodelle investieren kann. Gerade werden neue Möglichkeiten der Umsatzgenerierung geprüft. Aufgrund des anhaltenden Immobilienbooms ist gerade eine Abteilung ‚Real Estate‘ mit ca. 16 Mitarbeitern im Aufbau. Es sollen Immobilienobjekte an Kunden vermittelt bzw. verkauft werden. Die Real Estate hätte dann die Rolle eines Maklers / Vermittlers inne.

## 2 Kundenvorgaben:

Es geht um die Konzeption eines Datenmodells / Datenbank für die operative Abwicklung des Tagesgeschäfts.

Alle Mitarbeiter sollen auf einen konsistenten, validen und aktuellen Datenbestand zugreifen.

Er beschreibt die folgenden Punkte:

### 2.1 Verkäufer

Es gibt Verkäufer von Objekten. Mit den Attributen

- Titel
- Vorname
- Nachname
- Straße
- PLZ
- Ort
- Telefon geschäftlich
- Telefon privat
- Telefon mobil
- Fax
- eMail

Ein Verkäufer bietet ein oder mehrere Objekte zum Verkauf an.

Verkäufer ist nicht unbedingt identisch mit Eigentümer.

Für die Belange hier ist diese Unterscheidung aber nicht tragend.

### 2.2 Objekte

Es gibt Objekte (Häuser: 1 FH, 2-3 FH, Mehrfamilienhäuser, Gewerbeobjekte, Wohnungen, Lofts, etc.) mit den Attributen

- Straße
- PLZ
- Ort
- Art des Objektes

- Preis
- Aktuell
- offen/verkauft
- Lage

Ein Objekt ist genau einer Lage zuzuordnen.

Ein Objekt wird von einem Verkäufer angeboten.

Ein Objekt ist aktuell oder nicht.

Alle Objekte sind Kaufobjekte.

Vermietungen werden hier nicht abgebildet.

## 2.3 Lage

Einer Lage können mehrere Objekte zugeordnet werden.

Jede Lage hat eine Lagebezeichnung (hier gemeint: der jeweilige Stadtteil in dem sich das Objekt befindet). Mit Lage ist hier keine Bewertung (i.S.v. Top, Mittel o.ä. gemeint.)

## 2.4 Kunden

Es gibt Kunden mit den Attributen

- Anrede
- Titel
- Vorname
- Nachname
- Straße
- PLZ
- Ort
- Telefon geschäftlich
- Telefon privat,
- Telefon mobil
- Fax
- eMail
- sonstige Bemerkungen
- aktuell
- (ein) Ansprechpartner.

Ein Kunde ist aktuell oder nicht.

## 2.5 Kundentypen

Es gibt die Kundentypen Privat, Makler, Bauträger.

Jeder Kunde ist genau einem Kundentyp zugeordnet

## 2.6 Kontakte mit den Kunden

Es gibt Kontakte mit den Kunden seitens der Real Estate mit den Attributen

- Kontaktart
- Datum
- Bemerkungen Kontakt

Ein Kunde hat mindestens einen bzw. mehrere Kontakte mit der Real Estate.

## 2.7 Kundenwünsche

Es gibt Kundenwünsche mit den Attributen

- Lage
- Haustyp
- Preis Minimum
- Preis Maximum
- Bemerkungen

Ein Kunde kann einen oder mehrere Kundenwünsche haben.

## 2.8 Besichtigungen

Ein Kunde kann kein, ein oder mehrere Objekte besichtigen.

Ein Objekt wiederum kann von einem oder mehreren Kunden besichtigt werden.

## 2.9 Exposé

Für jedes Objekt gibt es ein Exposé (eine Art ausführliche Objektbeschreibung).

Zu einem Objekt kann einem Kunden ein Exposé zugesendet werden.

Ein Kunde kann auch mehrere Exposés von mehreren Objekten bekommen.

Das Exposé von einem Objekt kann auch an viele Kunde gesendet werden.

## 2.10 Kauf

Ein Kunde kann kein, ein oder mehrere Objekte kaufen.

Ein Objekt wiederum wird i.d.R. nur von einem Kunden gekauft.

Der Abteilungsleiter der Real Estate hat gleich noch typische Aktionen beschrieben, die mit dem neuen IT-System möglich sein sollen:

- Abfrage Anzahl Besichtigungen pro Objekt
- Abfrage Besichtigungsliste nach Objekt
- Abfrage mögliche Objekte für Kunden
- Abfrage Anzahl Exposés pro Kunde
- Abfrage Anzahl versandter Exposés nach Objekt
- Kundesuche nach Kunden Nummer
- Kundesuche nach Nachname
- Kundesuche nach Name Objekt
- Kundensuche nach Haustyp
- Kundensuche nach Lage
- Kundensuche nach Lage und Typ
- Kundensuche nach Lage, Typ und Preis
- Kundensuche nach Preis Maximum
- Kundensuche nach Nachname
- Kundenliste von Objekt
- Objektsuche nach Straße
- Objektsuche nach Lage und Typ
- Objektsuche nach Nummer
- Objektsuche nach Straße
- Objektsuche nach Verkäufer

- Objektsuche nach Straße
- Abfrage Schnittmenge Kundenwunsch Objekte
- Abfrage Verkäufer
- Abfrage zugesandte Exposés nach Name
- Mailing an alle Kunden
- Mailing an alle Kunden bzgl. neuem Objekt

### 3 Offene Fragen 1 (Kundentermin 06.04.2021)

- ☐ Kunde interessiert sich für mehrere Objekte. Wer ist sein einer Ansprechpartner (Kontakt oder Verkäufer)?
- ☐ Können Kunden auch direkt Real Estate telefonisch kontaktieren?
- ☐ Wie ist der Prozess der Kontaktaufnahme vorgesehen? Gibt es Vorstellungen vom Kunden?
- ☐ Wie ist der Prozess Verkaufsanbahnung geplant/definiert? (Besichtigung nötig?, Expose nötig?) Hat der Kunde bereits eine Vorstellung oder eine Vorgehensweise?
- ☐ Was ist mit dem Status „Aktuell“ gemeint? (Objekt und Kunde)
- ☐ Soll der Kauf eines Objektes an mehrere Kunden vorgesehen werden?
- ☐ DWH, sind dafür weitere, komplexere, Abfragen vorgesehen.

#### 3.1 Ergebnis

- Es wird nur in Deutschland vermittelt.
- Interessent und Kunden kommunizieren nur über Mitarbeiter der Weber AG.
- Derzeitiges Volumen beträgt ca. 1.000 Anfragen / Tag.
- Angepeiltes Volumen für die Zukunft ca. 20.000 Anfragen / Tag.
- Es gibt einen Datenaltbestand in Excel (80%) und Papierform (20%). Dieser soll mit in das neue System übernommen werden. Planung für die Umsetzung gefordert.
- Datenaustausch externer Leads im csv-Format für Verkäufer und Interessenten erforderlich. Die benötigte Datenqualität muss >90% betragen.
- Derzeit sollen 16 Mitarbeiter im System arbeiten. Es muss beachtet werden, dass das System auch beim angestrebten Volumen von 20.000 Anfragen / Tag noch performant arbeitet und auch erweiterbar sein sollte, da sich die Anzahl paralleler Transaktionen vergrößern wird.
- Exposés können sowohl im Dateisystem oder als BLOBs (Binary Large Objects) abgelegt werden und in der Datenbank entsprechend verlinkt werden.
- Die gesamte Kommunikation, auch Versuche, mit Kunden und Interessenten wird in der Tabelle „Kontakte mit den Kunden“ abgelegt. Diese ist eine essentielle Quelle für spätere Auswertungen und Optimierung des bestehenden Geschäftsprozesses.
- Das Merkmal „Aktuell“ wird benötigt, wenn ein Kunde oder Objekt gerade verhindert oder nicht verfügbar ist und zu einem späteren Zeitpunkt wieder aktiviert werden soll.

- Status „abgeschlossen“ bedeutet, dass der Verkauf des Objektes durch notarielle Bestätigung abgeschlossen wurde. (Status: offen, verkauft, abgeschlossen)  
Offen Objekt steht zur Verfügung  
Verkauft Objekt wurde verkauft ist aber noch nicht notariell bestätigt
- Um Interessenten und Verkäufer zusammen zu bringen verwendet die Weber AG die jeweiligen Angaben
  - Ort (Stadtteil)
  - Preis
  - Größe in m<sup>2</sup>
- Dieses Matching soll automatisiert erfolgen und den Mitarbeitern mitgeteilt werden.
- Für spätere Auswertung aus dem Datawarehouse interessiert sich die Weber AG für folgende Werte (KPI's\*):
  - Art des Objektes
  - Standort des Objektes
  - Ort des Kunden
  - Preis
  - Zeiten (Entscheidung zum Kauf, Laufzeit Verkauf ...)
  - Merkmale Objekt
  - Mitarbeiter
  - Kontakte (Anzahl, Art der Kontakte, Zeitverlauf)

\* KPI ist die Abkürzung für Key Performance Indicator.  
Der Begriff bezeichnet Kennzahlen, mit denen die Leistung von Aktivitäten in Unternehmen ermittelt werden kann. Welche KPIs betrachtet werden sollten, um Erfolg oder Misserfolg zu messen, hängt vom Unternehmen, der jeweiligen Maßnahme und deren Zielen ab.

Folgende Informationen werden benötigt:

## 3.2 Lead Management

### 3.2.1 Felder bei Leads Interessenten:

- Anrede
- Vorname
- Nachname
- Straße
- Hausnummer
- PLZ
- Ort
- E-Mail
- Telefonnummer
- Art des Objektes (Haus, Mehrfamilienhaus, Eigentumswohnung, etc.)
- Preisvorstellung
- Größe im qm
- Ort wo das Objekt gesucht wird
- Anzahl Zimmer

### 3.2.2 Felder bei Leads Verkäufer:

- Firma
- Anrede / Ansprechpartner
- Vorname
- Nachname
- Straße
- Hausnummer
- PLZ
- Ort
- E-Mail
- Telefonnummer
- Art des Objektes (Haus, Mehrfamilienhaus, Eigentumswohnung, etc.)
- Preisvorstellung
- Größe im qm
- Ort wo das Objekt gesucht wird
- Anzahl Zimmer
- Exposee vorhanden (optional)

### 3.2.3 Felder im bestehenden Formular Kaufinteressent:

- Anrede (Herr, Frau, Firma)
- Titel
- Vorname
- Nachname
- Straße
- PLZ
- Ort
- Telefon geschäftlich
- Telefon privat
- Telefon mobil
- E-Mail
- Sonstige Bemerkungen
- Ansprechpartner
- Kontaktwunsch
- Gewünschte Kontaktart (tel., email, etc.)

### 3.2.4 Felder zu Kaufinteressent Wunsch:

- Art des Objektes
- Ort
- Lage
- Preisminimum
- Preismaximum
- Größe Objekt in qm

- Anzahl Zimmer
- Sonstige Bemerkungen

### 3.2.5 Felder im bestehenden Internetformular Verkäufer:

- Anrede (Herr, Frau, Firma)
- Titel
- Vorname
- Nachname
- Strasse
- PLZ
- Ort
- Telefon geschäftlich
- Telefon privat
- Telefon mobil
- E-Mail
- Sonstige Bemerkungen
- Ansprechpartner
- Kontaktwunsch (Gewünschte Kontaktart (tel., email, etc.))
- Infos zum Objekt
- Art des Objektes
- Ort
- Lage
- Preis
- Größe Objekt in qm
- Anzahl Zimmer
- Sonstige Bemerkungen
- Exposee (optional)

### 3.2.6 Kommunikationskanäle:

- Telefon, Mobiltelefon
- E-Mail
- Fax
- WhatsApp
- SMS
- MMS
- Internetformular

### 3.2.7 Mögliche Aktionen bei Aufnahme Kundenanforderungen:

- Exposee senden
- TelKo mit Käufer
- TelKo mit Verkäufer
- Besichtigungstermin vereinbaren
- einen weiteren Besichtigungstermin vereinbaren
- Notartermin

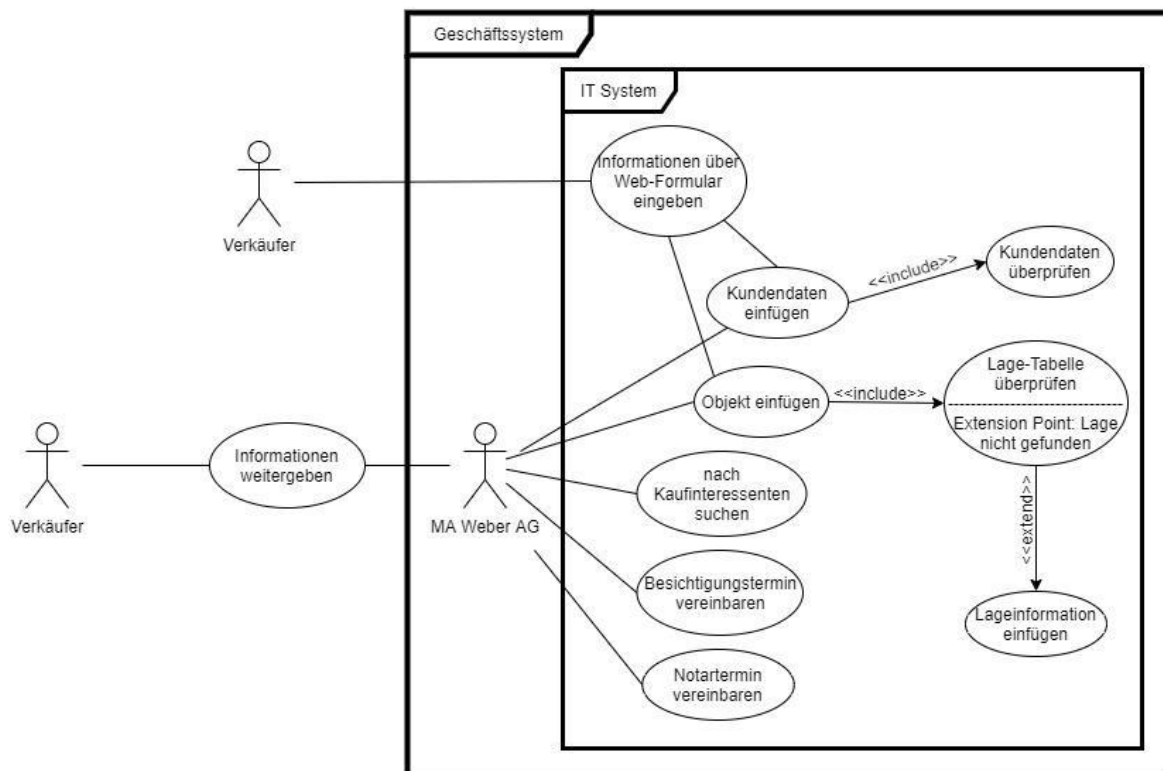


- Fragen beantworten

### 3.3 Use Cases

- ❖ Neues Objekt über alle Kanäle
- ❖ Neuer Interessent über alle Kanäle
- ❖ Matching über Ort (Ortsteil), Preis, Größe -> Exposees automatisch an Interessenten verschicken
- ❖ Altdatenbestand einlesen
- ❖ Lead-Dateien einlesen
- ❖ Kontakte mit Kunden (Erstkontakt, Expose, Besichtigungstermin, Notartermin)
- ❖ Verkauf des Objektes

#### 3.3.1 Neues Objekt über alle Kanäle

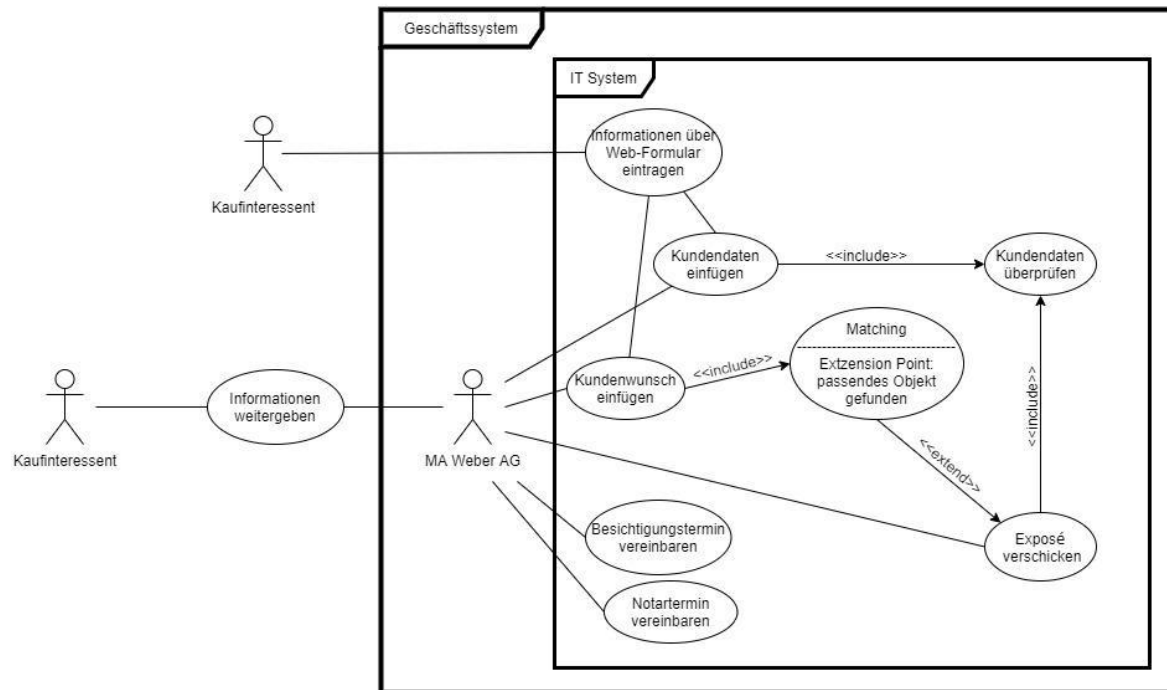


Ein Verkäufer übermittelt seine Daten und die des zu verkaufenden Objektes über einen der angebotenen Kommunikationskanäle aus 4.2.6. Objekt wird erst nach Prüfung durch MA der Weber AG freigeschaltet.

Name	Objekt einfügen
Ziel	Neues Objekt in der Objekt-Tabelle einfügen
Vorbedingung	Daten des neuen Objekts müssen vorliegen
Nachbedingung Erfolg	Neues Objekt wurde hinzugefügt
Nachbedingung Fehlschlag	Neues Objekt konnte nicht hinzugefügt werden

Akteure	Mitarbeiter/IT-System
Auslösendes Ereignis	Kunde möchte ein neues Objekt verkaufen
Beschreibung	1. Info über neues Objekt 2. Prüfen, ob alle Daten vorhanden sind 3. Prüfen, ob die Lage in der Lage-Tabelle vorhanden ist 4. Daten zu dem neuen Objekt in die Objekt-Tabelle einfügen
Erweiterungen	3a. Lageinformation einfügen
Alternativen	-

### 3.3.2 Use Case neuer Interessent über alle Kanäle



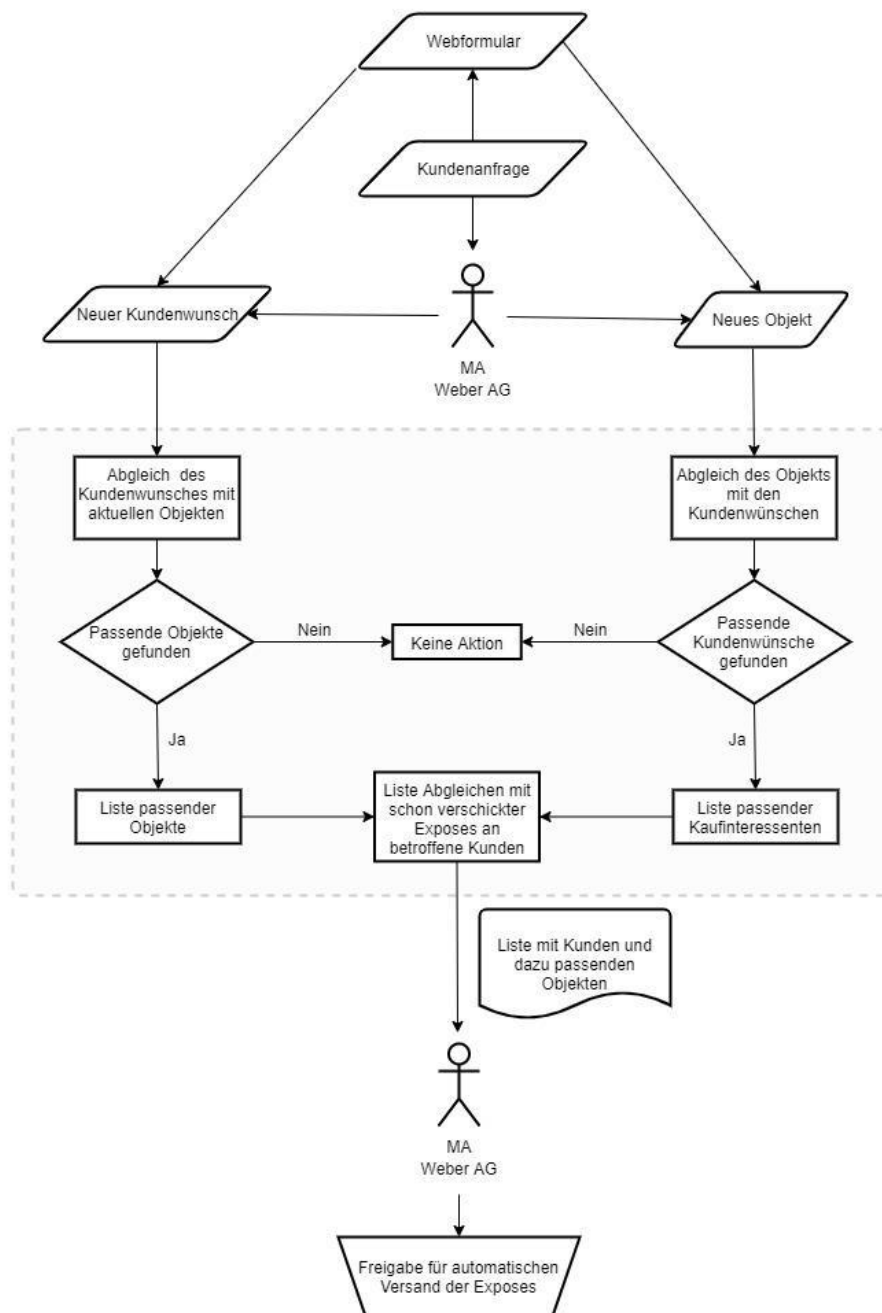
Ein Interessent übermittelt seine Daten über einen der angebotenen Kommunikationskanäle aus 4.2.6. Ein Kundenwunsch kann, muss aber nicht angegeben werden. Der Interessent ist direkt im System aktiv. Bei Eintrag über Web wird automatisch der Kommentar mit „Automatisch über Web“ befüllt.

Name	Kundendaten einfügen
Ziel	Daten eines neuen Kunden in der Kunden-Tabelle einfügen
Vorbedingung	Daten des neuen Kunden müssen vorliegen
Nachbedingung Erfolg	Neuer Kunde wurde hinzugefügt
Nachbedingung Fehlschlag	Neuer Kunde konnte nicht hinzugefügt werden
Akteure	Mitarbeiter/IT-System
Auslösendes Ereignis	Neuer Kunde möchte die Dienste von Weber AG Real Estate nutzen
Beschreibung	1. Info über neuen Kunden 2. Prüfen, ob alle Daten vorhanden sind 3. Prüfen, ob der Kunde bereits im System vorhanden ist 4. Daten des neuen Kunden in die Kunden-Tabelle einfügen
Erweiterungen	-

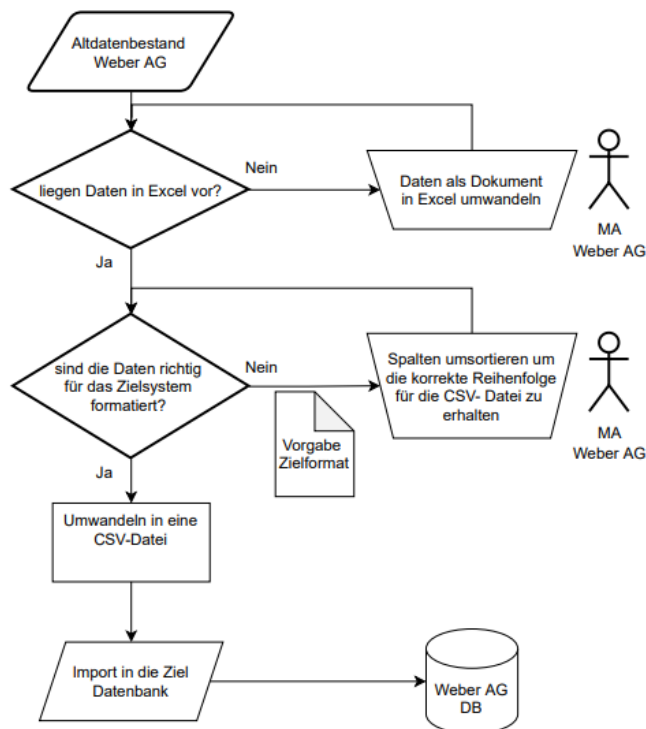
Alternativen	3a. Kunde ist schon vorhanden 3a1. Prüfen, ob alle Daten übereinstimmen 3a2. Daten aktualisieren
--------------	--

Name	Exposé verschicken
Ziel	Ein Exposé über ein für den Kunden interessantes Objekt an den Kunden weiterleiten
Vorbedingung	Ein für den Kunden interessantes Objekt wurde gefunden
Nachbedingung Erfolg	Das Exposé wurde an den Kunden verschickt
Nachbedingung Fehlschlag	Ein Exposé konnte nicht verschickt werden
Akteure	Mitarbeiter
Auslösendes Ereignis	Ein zu dem Kundenwunsch passendes Objekt wurde gefunden
Beschreibung	1. Das Exposé zu dem gefundenen Objekt wird abgerufen 2. Die E-Mail-Adresse des Kunden wird abgerufen 3. Exposé wird verschickt
Erweiterungen	-
Alternativen	2a. Keine E-Mail-Adresse vorhanden 2a1. Prüfen, welchen Kontaktart der Kunde bevorzugt 2a2. Das Exposé über den bevorzugten Weg weitergeben

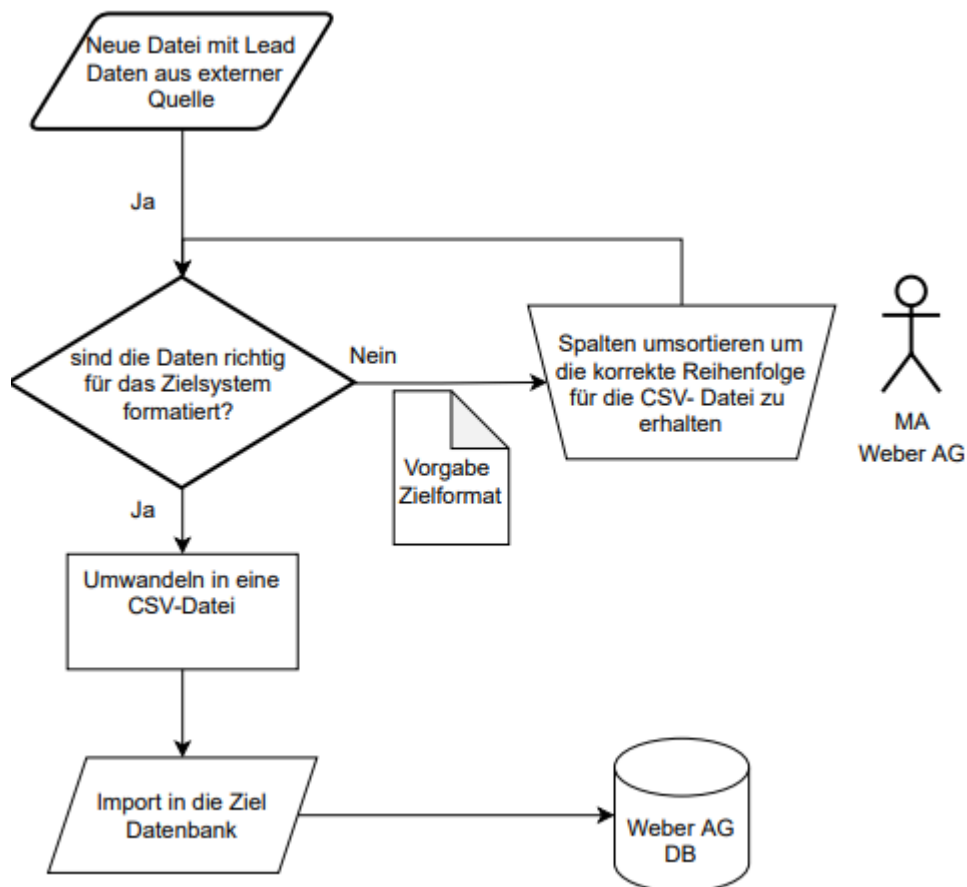
### 3.3.3 Aktivitätsdiagramm Matching Objects



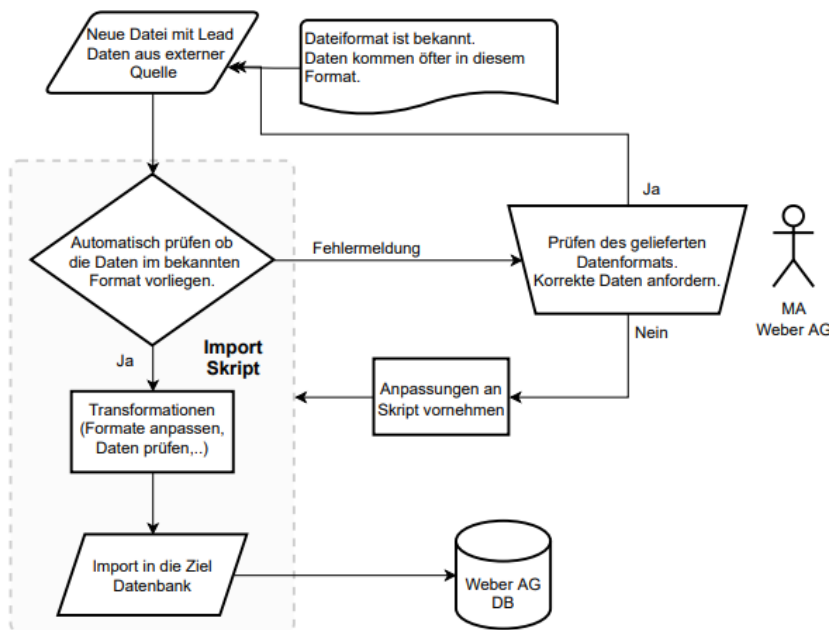
### 3.3.4 Aktivitätsdiagramm Altdaten einlesen



### 3.3.5 Aktivitätsdiagramm Lead Daten manuell laden



### 3.3.6 Aktivitätsdiagramm Lead Daten automatisch laden



## 4 Offene Fragen 2 (Kundentermin 07.04.2021)

- ☐ Muss ein neuer Kaufinteressent auch direkt einen Wunsch mit angeben? (Mandatory) **NEIN**
- ☐ DWH: Sollen die Analyseanforderungen später noch erweitert werden.
- ☐ Ist es wichtig dem Kunden ein Merkmal der Kundenart (Verkäufer, Interessent oder beides zu geben) **Aufteilen in 2 Entitäten oder IS A**
- ☐ Wenn sich ein Kontakt auf mehrere Objekte bezieht, soll dies im DWH mit abgebildet werden? (z.B. Kontakt mit Interessent wegen mehrerer Wunschobjekte) **Mehrere Kontakt aufnehmen**
- ☐ Daten aus den Web Formularen, wohin sollen diese geladen werden? ( Firma, Adresse..)
  - Nicht konsistent ☐ gefordert aber im Web nicht vorhanden ☐ Anpassung Web-Auftritt?
- ☐ Nichtfunktionale Anforderungen
- ☐ Anforderungen an Frontend?
- ☐ Berechtigungskonzept nötig?
- ☐ Sollen alle Kundenaktionen wirklich immer über den MA der Weber AG laufen?
  - Automatismus das über WEB eingegebene Daten nur durch MA freigegeben werden?
- ☐ Muss die Kundennummer bei Altdatenübernahme und/oder neuen Leads mit gespeichert werden?
- ☐

## 4.1 Ergebnis

- Ein Kaufinteressent kann auch ohne Kundenwunsch angelegt werden.
- Es wurden folgende Anforderungen an die DWH Analysen bestätigt:
  - Werte (KPI's\*):
    - Art des Objektes
    - Standort des Objektes
    - Ort des Kunden
    - Preis
    - Zeiten (Entscheidung zum Kauf, Laufzeit Verkauf ...)
    - Merkmale Objekt
    - Mitarbeiter
    - Kontakte (Anzahl, Art der Kontakte, Zeitverlauf)
- Der endgültige Kaufpreis eines Objektes ist in der Entität Kauf abgelegt. Beim Konzept für DWH beachten da dies ein Schlüsselwert ist.
- Entität Kunde über „is a“ aufteilen oder als 2 Entitäten (Interessent, Verkäufer) ausbilden. Redundanz von Daten (Kunde ist Interessent und Verkäufer) können vernachlässigt werden.
- Nach Fertigstellen des Datenbankmodells einen Änderungsvorschlag für die Webformulare erstellen.
- Das Feld Firma bei Interessent und Verkäufer vorsehen.
- Weder bei den Altdaten, noch bei den neuen Leads gibt es eine alte Kundennummer die in der DB gespeichert werden soll.
- Anforderungen an die Benutzeroberfläche:
  - Deutsch
  - Schnelle Bereitstellung von Informationen aus der Datenbank
  - Unterstützung bei der Eingabe (Vorbelegung, Drop-Down,..)
- Datenbank befindet sich Inhouse
- System muss skalierbar sein, Planung mindestens auf 20.000 Anfragen pro Tag.
- Lieferdatum Konzept: 08.04.2021
- Es gibt Benutzer und 1 Power User (Administrator IT).
- Derzeit kommen ca. 900 Anfragen über die Webanwendung.
- Eingehende Interessenten werden automatisch im System angelegt und als Aktiv gekennzeichnet. Zusätzlich erfolgt eine Markierung als „Automatisch über Web“.
- Eingehende Verkäufer werden automatisch im System angelegt und als „Aktiv“ gekennzeichnet. Zusätzlich erfolgt eine Markierung als „Automatisch über Web“. Das dazugehörige Objekt wird ebenfalls automatisch angelegt aber als Status „Inaktiv Web“ und Aktuell „Nein“ markiert. Eine Freigabe erfolgt, nach Prüfung, durch MA der Weber AG.
- DWH Architektur (Wie kommen Daten in das DWH, ETL, Warum wurde sich für ein Schema entschieden, Vorschlag für Umsetzung, mögliche Erweiterungen)

- Bei Import wird die Lage über den angegebenen Ort ermittelt. Ggf. müssen Anpassungen an der Entität Ort erfolgen.

## 5 Das Modell der operativen Datenbank

### 5.1 Das ERM-Modell

Die Abbildung 1 zeigt das WERM-Modell des geplanten Datenbanks. Die gewünschten Daten zu den Objekten und Kundenwünschen sind jeweils in der gleichnamigen Entität (blaue Rechtecke) gespeichert. Die Exposés werden getrennt abgelegt und mit dem Datenbank verlinkt. Die Entität „Lage“ enthält die Namen der Stadtteile, in denen die Objekte liegen. In einem Kundenwunsch können mehrere dieser Stadtteile angegeben werden. Die Lage ist ein zentrales Kriterium für den Abgleich zwischen Kundenanfragen und Angebot. Zu jedem Datensatz wird die Quelle angegeben (Web-Formular, Lead, etc.), woher er stammt.

Die Kunden sind entweder Käufer (haben mindestens einen Kundenwunsch) oder Verkäufer (haben mindestens ein Objekt angeboten). Die gemeinsamen Attribute von Käufer und Verkäufer werden in der Tabelle „Kunde“ gespeichert. Jeder Kontakt mit einem Kunden wird in der Tabelle „Kontakt“ festgehalten. Dabei wird gespeichert, über welchen Channel der Kontakt erfolgte, mit welchem Mitarbeiter hatte der Kunde Kontakt, um welchen Objekt ging es dabei und welche Aktionen bzgl. des Käufers bzw. bzgl. des Verkäufers wurden getätigt (z.B. Exposé verschicken, Besichtigungstermin vereinbaren, etc.). Auch Besichtigungstermine an sich werden als Kontakt eingetragen. Der Kauf wird in einer eigenen Entität festgehalten.



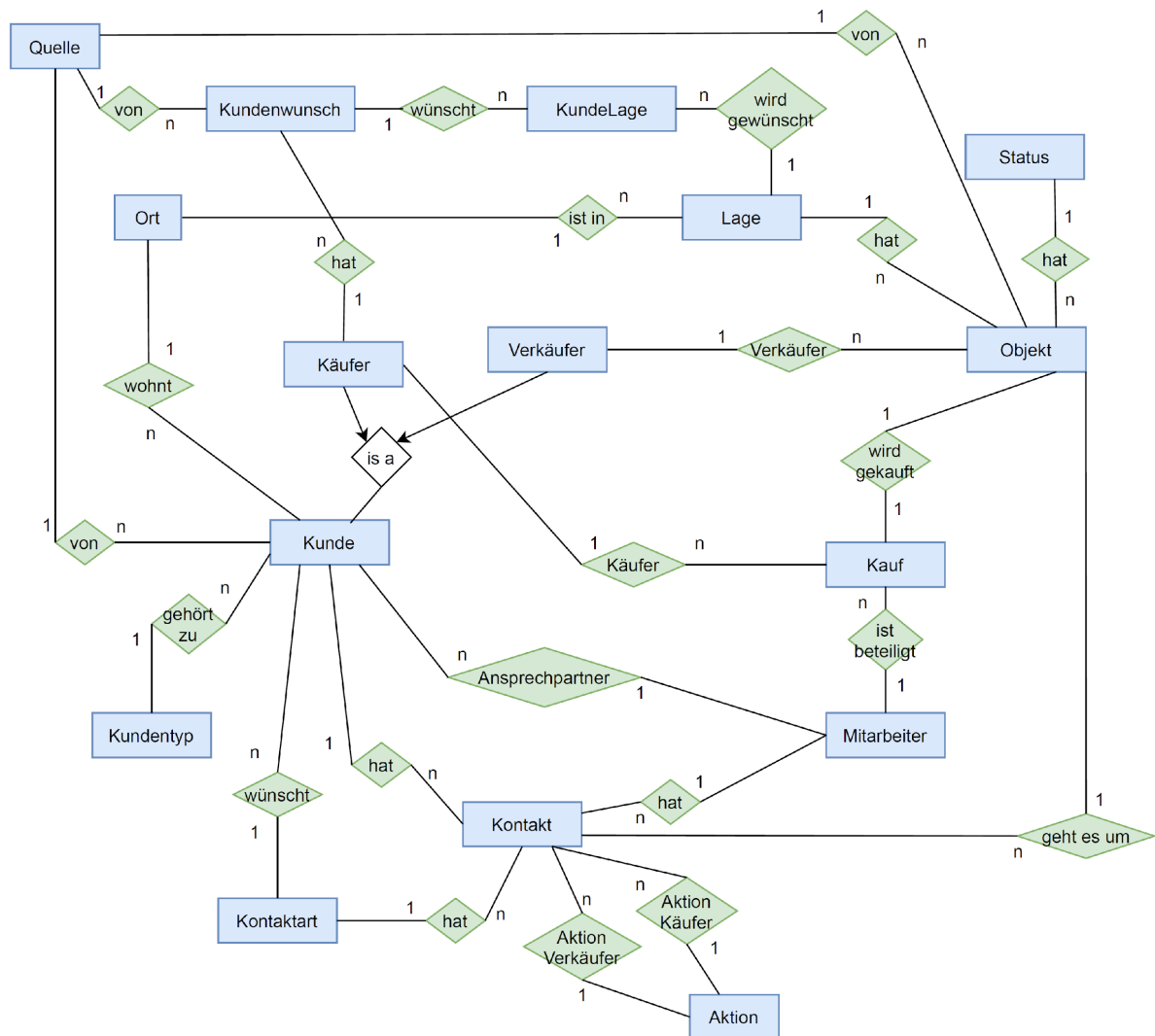


Abbildung 1: ERM Modell für das operative Datenbank der Firma Weber AG, Abteilung Real Estate

## 5.2 Tabellendefinition

In den folgenden Tabellen sind die jeweiligen Attribute der Tabellen mit einer kurzen Beschreibung des Inhalts dokumentiert. Die grau hinterlegten Attribute (Schlüssel) sowie die Spalten „Datentyp“ und „Einschränkungen“ haben nur eine technische Bedeutung.

Kunde			
Attribut	Beschreibung	Datentyp	Einschränkungen
KundeID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Ansprechpartner	Verweis auf zuständigen Mitarbeiter von RealEstate	INTEGER	NOT NULL
Kundentyp	Verweis auf die Tabelle Kundentyp	INTEGER	NOT NULL
Kontaktart	Auf welchem Weg möchte der Kunde kontaktiert werden; Verweis auf Tabelle Kontaktart	INTEGER	NOT NULL
Ort	Verweis auf die Tabelle Ort	INTEGER	NOT NULL
Quelle	woher stammen die Daten? Verweis zur Tabelle Quelle	INTEGER	NOT NULL
Anrede	Anrede (des Ansprechpartners)	TEXT	

Titel	Titel (des Ansprechpartners)	TEXT	
Vorname	Vorname (des Ansprechpartners)	TEXT	NOT NULL
Nachname	Nachname (des Ansprechpartners)	TEXT	NOT NULL
Firma	Firmenname, falls der Kunde eine Firma repräsentiert	TEXT	
Straße		TEXT	
Hausnummer		TEXT	
PLZ	Postleitzahl	INTEGER (5)	
Tel_geschaeft	Telefonnummer geschäftlich	INTEGER	
Tel_privat	Telefonnummer privat	INTEGER	
Tel_mobil	Handynummer	INTEGER	
Fax		INTEGER	
eMail		TEXT	
Kommentar	Sonstige Bemerkungen	TEXT	

Mitarbeiter			
Attribut	Beschreibung	Datentyp	Einschränkungen
MitarbeiterID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Anrede	Anrede des Mitarbeiters	TEXT	
Titel	Titel des Mitarbeiters	TEXT	
Vorname	Vorname des Mitarbeiters	TEXT	NOT NULL
Nachname	Nachname des Mitarbeiters	TEXT	NOT NULL
Tel_geschaeft	Telefonnummer geschäftlich	INTEGER	
Tel_privat	Telefonnummer privat	INTEGER	
Tel_mobil	Handynummer	INTEGER	
Fax		INTEGER	
eMail		TEXT	
Aktuell	ist der Mitarbeiter zur Zeit verfügbar	BOOLEAN	NOT NULL
Kommentar	Sonstige Bemerkungen	TEXT	

Kundentyp			
Attribut	Beschreibung	Datentyp	Einschränkungen
TypID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Beschreibung	Privat, Makler, Bauträger, etc.	TEXT	UNIQUE, NOT NULL

Kontaktart			
Attribut	Beschreibung	Datentyp	Einschränkungen
ChannelID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Beschreibung	Telefon (geschäftlich, privat), eMail, Fax, etc.	TEXT	UNIQUE, NOT NULL

Ort			
Attribut	Beschreibung	Datentyp	Einschränkungen
OrtID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Bezeichnung	Name des Ortes	TEXT	UNIQUE, NOT NULL

Quelle			
Attribut	Beschreibung	Datentyp	Einschränkungen
QuelleID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Beschreibung	Lead, Web-Formular, Eintrag durch Mitarbeiter	TEXT	UNIQUE, NOT NULL

Käufer			
Attribut	Beschreibung	Datentyp	Einschränkungen
KaeuferID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
KundelD	Verweis aus Tabelle Kunde	INTEGER	UNIQUE, NOT NULL
Aktuell	ist der Käufer zur Zeit verfügbar	BOOLEAN	NOT NULL

Verkäufer			
Attribut	Beschreibung	Datentyp	Einschränkungen
VerkaeufelD	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
KundelD	Verweis aus Tabelle Kunde	INTEGER	UNIQUE, NOT NULL

Kundenwunsch			
Attribut	Beschreibung	Datentyp	Einschränkungen
WunschID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Kaeufer	Verweis zu den Käufer, der den Gesuch abgegeben hat	INTEGER	NOT NULL
Quelle	woher stammen die Daten? Verweis zur Tabelle Quelle	INTEGER	NOT NULL
Haustyp	Typ des Hauses (Einfamilienhaus, Wohnung,...)	TEXT	
PreisMin	Untere Preisgrenze für die Suche	REAL	
PreisMax	Obere Preisgrenze für die Suche	REAL	NOT NULL
Groesse	Wohnfläche in m <sup>2</sup>	REAL	NOT NULL
Zimmer	Anzahl der Zimmer	INTEGER	
Bemerkungen	Sonstige Wünsche, Kommentare	TEXT	

KundeLage			
Attribut	Beschreibung	Datentyp	Einschränkungen
Kundenwunsch	Verweis zur Tabelle Kundenwunsch	INTEGER	NOT NULL
Lage	Verweis zur gewünschten Lage	INTEGER	NOT NULL

Lage			
Attribut	Beschreibung	Datentyp	Einschränkungen
LageID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
OrtID	Ort, in dem sich der Stadtteil befindet; Verweis zur Tabelle Ort	INTEGER	NOT NULL
Bezeichnung	wo befindet sich ein Objekt (Stadtteil)	TEXT	UNIQUE, NOT NULL

Objekt			
--------	--	--	--

Attribut	Beschreibung	Datentyp	Einschränkungen
ObjektID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Lage	Verweis auf Tabelle Lage	INTEGER	NOT NULL
Ort	Verweis auf Tabelle Ort	INTEGER	NOT NULL
Verkaeuer	Verweis zu den Verkäufer, der das Objekt verkauft	INTEGER	NOT NULL
Status	Verweis auf die Tabelle Status	INTEGER	NOT NULL
Quelle	woher stammen die Daten? Verweis zur Tabelle Quelle	INTEGER	NOT NULL
Straße		TEXT	NOT NULL
Hausnummer		TEXT	
PLZ	Postleitzahl	INTEGER (5)	NOT NULL
Haustyp	Typ des Hauses (Einfamilienhaus, Wohnung,...)	TEXT	NOT NULL
Preis	Preis des Objektes	REAL	NOT NULL
Groesse	Wohnfläche in m <sup>2</sup>	REAL	NOT NULL
Zimmer	Anzahl der Zimmer	INTEGER	
Bemerkungen	Sonstige Merkmale, Kommentare	TEXT	
Aktuell	ist das Objekt zur Zeit verfügbar	BOOLEAN	NOT NULL
Expose	Link zum Dokument	TEXT	

Status			
Attribut	Beschreibung	Datentyp	Einschränkungen
StatusID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Status	offen/reserviert/verkauft	TEXT	UNIQUE, NOT NULL

Aktion			
Attribut	Beschreibung	Datentyp	Einschränkungen
AktionID	Primärschlüssel	INTEGER	UNIQUE, NOT NULL
Aktion	Expose verschickt, Besichtigung, etc.	TEXT	UNIQUE, NOT NULL

Kontakt			
Attribut	Beschreibung	Datentyp	Einschränkungen
Kunde	Verweis zu den Kunden, mit dem Kontakt stattfand	INTEGER	NOT NULL
Mitarbeiter	Verweis zum Mitarbeiter, der Kontakt hatte	INTEGER	NOT NULL
Kontaktart	Auf welchem Weg erfolgte der Kontakt; Verweis zur Tabelle Kontaktart	INTEGER	NOT NULL
Objekt	Um welchen Objekt ging es bei dem Kontakt; Verweis auf die Tabelle Objekt	INTEGER	NOT NULL
AktionKaeufer	Art des Kontaktes bzgl. des Käufers; Verweis auf Tabelle Aktion	INTEGER	NOT NULL

AktionVerkaeufer	Art des Kontaktes bzgl. des verkäufers; Verweis auf Tabelle Aktion	INTEGER	NOT NULL
Datum	Datum des Kontaktes	DATE	NOT NULL
Erfolg	war die Kontaktaufnahme erfolgreich?	BOOLEAN	NOT NULL
Bemerkungen	sonstige Bemerkungen	TEXT	

Kauf			
Attribut	Beschreibung	Datentyp	Einschränkungen
Kaeufer	Verweis zu den Kunden, der einen Objekt kauft	INTEGER	NOT NULL
Objekt	Verweis zu dem Objekt, um den es geht	INTEGER	UNIQUE, NOT NULL
Verkaeufer	Verweis zu den Kunden, der einen Objekt verkauft	INTEGER	NOT NULL
Mitarbeiter	Verweis zu den beteiligten Mitarbeiter	INTEGER	NOT NULL
Datum	Datum des Notartermins	DATE	NOT NULL
Abgeschlossen	Bestätigung durch Notar, dass der Kaufvertrag abgeschlossen ist	BOOLEAN	NOT NULL
Preis	Preis, den der Kunde für das Objekt bezahlt	REAL	NOT NULL
Bemerkungen	sonstige Bemerkungen	TEXT	

Vergleicht man die Tabellen mit der ERM, so fällt auf, dass die Tabelle Objekt ein Verweis auf die Tabelle Ort enthält, zwischen den Entitäten gibt es aber keine direkte Beziehung. Stattdessen wird der Ort automatisch über die Angabe der Lage ermittelt. Genauso enthält die Tabelle Kauf keine direkte Beziehung zum Verkäufer, diese wird über das Objekt ermittelt. So sollen mögliche Unstimmigkeiten bei der Dateneingabe vermieden werden.

## 6 Datawarehouse

Die Abteilung Real Estate interessiert sich ebenfalls für Business Intelligence Themen. Konkret geht es um die Planung eines Datawarehouse (DWH). Die Fragestellung, das analysiert werden soll, bezieht sich auf die Effektivität (wie oft stehen die Mitarbeiter von Weber AG im Kontakt mit den Kunden, bis es zu einem Kauf kommt?) und die Faktoren, die eine Kaufentscheidung beeinflussen (Art des Objektes, Objektmerkmale, Ort, Preis). Im Folgenden werden einige Grundlegende Überlegungen zu DWH-Architektur erläutert mit Vorschlägen zu dem geplanten Vorhaben sowie einem ersten Modell-Vorschlag.

### 6.1 DWH Architektur

Ein DWH besteht in der Regel aus mehreren Schichten, d.h. mehrere Datenbanken oder abgegrenzte Bereiche in einer Datenbank. Die Daten werden aus dem operativen System und evtl. aus zusätzlichen (externen) Quellen in die erste Schicht, der Stageing-Area geladen. Es folgen mehrere Prozesse, die zur Bereinigung und Transformation der Daten in das im DWH benötigte Format dienen. Diese Phase ist vor Allem wichtig für die Datenqualität, besonders, wenn Daten aus mehreren Quellen zusammengeführt werden sollen. Die Schicht, in der diese Prozesse stattfinden nennt man auch Cleansing-Area. Das Ergebnis sind Tabellen, die in das Kern-DWH geladen werden können. Den ganzen Prozess von der Datenextraktion aus den Quellsystemen, der Transformation und das Laden ins Kern-DWH kürzt man ab als ETL. Im Kern-DWH werden die Daten historisiert gespeichert. Für die

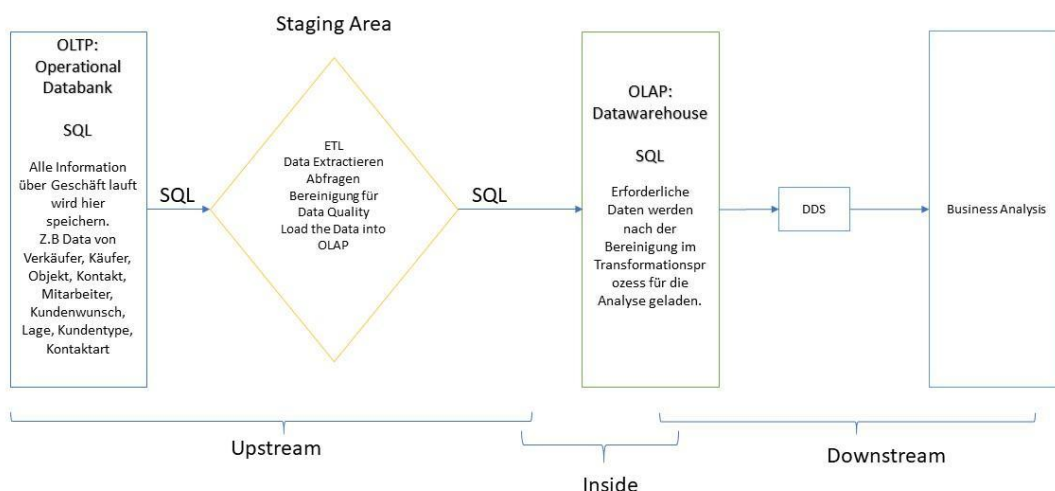
Analyse-Abfragen wird in der Regel noch eine weitere Schicht gebildet, in dem sich die Datamarts befinden. Die Datamarts enthalten den Teil der im Kern-DWH gespeicherten Daten, der für eine analytische Fragestellung nötig ist.

#### 6.1.1 DWH Architektur für Weber AG, Abteilung Real Estate

Die Abteilung hat eine konkrete Fragestellung, die mit den Daten aus dem operativem System beantwortet werden soll. Die Datenqualität wird schon beim Eingang der Daten in das operative System überprüft, so dass in dem DWH Cleansing-Area lediglich das Format der Tabellen an das DWH-Datenmodell angepasst werden soll, die ETL-Prozesse sind also nicht besonders komplex.

Da es nur eine Fragestellung gibt, ist ein Datamart-Schicht nicht zwingend notwendig. Sie kann als View auf das Kern-DWH implementiert werden. Daher haben wir für den DWH-Modell das Galaxy-Schema gewählt. Dies ist ein Dimensionales Schema, in dem die Stammdaten in Dimensionstabellen gespeichert werden, die Bewegungsdaten und Kennzahlen werden in Faktentabellen gespeichert, die die Dimensionen verknüpfen. Das Vorteil dieses Schemas ist es, dass man für die Abfragen nur wenige JOINS braucht, dadurch sind die Abfragen schneller, die Performance ist hoch. Nachteil dieses Schemas ist die denormalisierte Datenhaltung. Die Redundanz in den Dimensionstabellen führt dazu, dass Änderungen in der Datenstruktur schwieriger durchzuführen sind. Außerdem ist dieses Schema nicht Speichereffizient. Falls Weber AG vor hat die Business Intelligence Funktionen weiter auszubauen, so dass auch andere Daten zusätzlich in das DWH mit aufgenommen werden sollen (z.B. Daten aus anderen Abteilungen, externe Daten) und noch weitere Fragestellungen analysiert werden sollen, empfehlen wir das DWH schon jetzt dafür angepasst anzulegen. Bei großen Datenmengen wäre eine normalisierte Datenhaltung im Kern-DWH angebracht (Data Vault oder Snowflake-Schema), bei diesen Designs sind auch nachträgliche Änderungen einfacher Durchzuführen. Sollen zusätzliche Fragestellungen dazukommen, empfiehlt es sich schon jetzt eine Data-Mart-Schicht anzulegen (wenn auch zuerst mit nur einem Datamart).

Da der Kunde keine Erweiterung in der nächsten Zeit vorgesehen hat, ist hier das DWH als Galaxy-Schema modelliert.

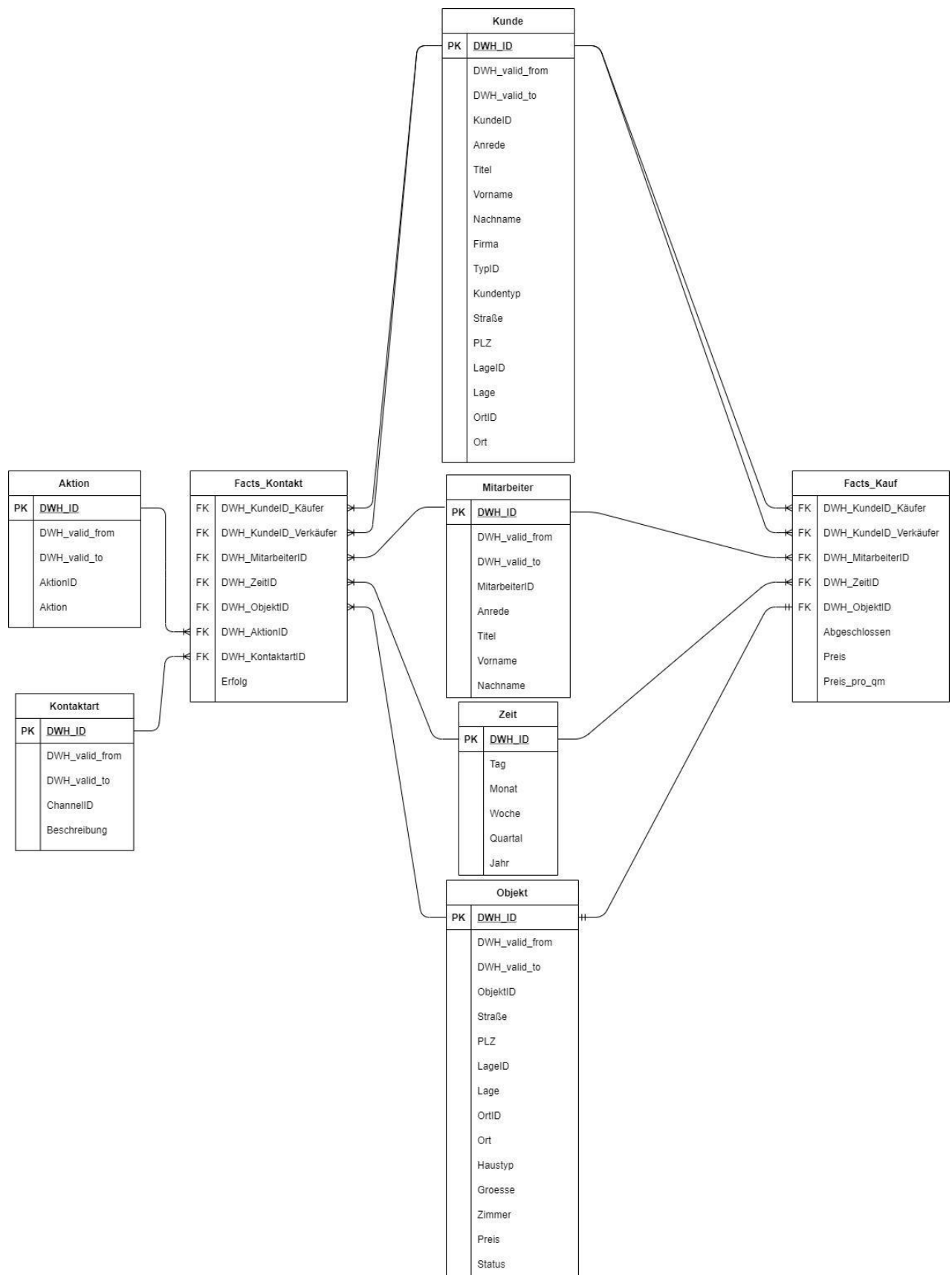


Upstream und Downstream ist der Datenflussprozess zur Analyse im neuen Immobiliengeschäft.

"Upstream" ist in Richtung der Quelle und "Downstream" ist in Richtung der Endanalyse. Upstream

wird vom Mitarbeiter durchgeführt, dessen Daten identifiziert, extrahiert oder in das Data Warehouse geladen werden, und Downstream ist näher an der Endanalyse.

## 6.2 Das Modell



Das Modell beinhaltet zwei Faktentabellen: Kontakt und Kauf. Als Dimensionen sind folgende Tabellen (mit ihren Hierarchien) vorgesehen:

- Zeit (Tag -> Woche -> Monat -> Quartal -> Jahr)
- Aktion (vgl. Aktion-Tabelle im op. DB)
- Kontaktart (vgl. Kontaktart-Tabelle im op. DB)
- Kunde (Straße -> Lage -> Ort; Kunde -> Kundentyp)
- Mitarbeiter (vgl. Mitarbeiter-Tabelle im op. DB)
- Objekt (Straße -> Lage -> Ort)

Die vom Kunden geforderten Daten sind in den entsprechenden Tabellen enthalten (Ortsinformationen zum Kunden, Ortsinformationen zum Objekt, Objektmerkmale, Preis, Zeit). In der Tabelle Kunde wurde zusätzlich die Information zu Lage (Stadtteil) eingefügt, dieses kann während des ETL-Prozesses aus der Adressinformationen ermittelt werden. Andere Daten sind in den operativen DB-Tabellen enthalten. In der Faktentabelle Kauf wurde zusätzlich das Preis pro m<sup>2</sup> als Kennzahl hinzugefügt. Für die Historisierung wurden die Felder „valid from“ und „valid to“ hinzugefügt, die für jeden Datensatz den Gültigkeitszeitraum angeben.

## 7 Import von externen Daten (Leads)

### 7.1 Vorgaben

#### 7.1.1 Intention

Um das Geschäft anzukurbeln möchte die Real Estate Fremddaten (sogenannte Leads) von Extern hinzukaufen. Es handelt sich bei den Daten um Objekte -die zum Verkauf stehen- und auch um Kaufinteressierte. Die Daten stammen von verschiedenen Anbietern. Die Datenqualität soll sich auf einem Level >90% bewegen.

### 7.2 Herausforderungen

□ Die gelieferten Daten müssen zu dem Datenmodell in der Datenbank passen.

- Datentyp
- Datenlänge
- Zeichensatz (Codepage)

□ Externe Daten sind bereits in der Datenbank vorhanden.

- Kunden oder Objekte sind bereits vorhanden.
- Welche Daten sind aktuell bzw. aktueller?
- Wie können gleiche Daten identifiziert werden?

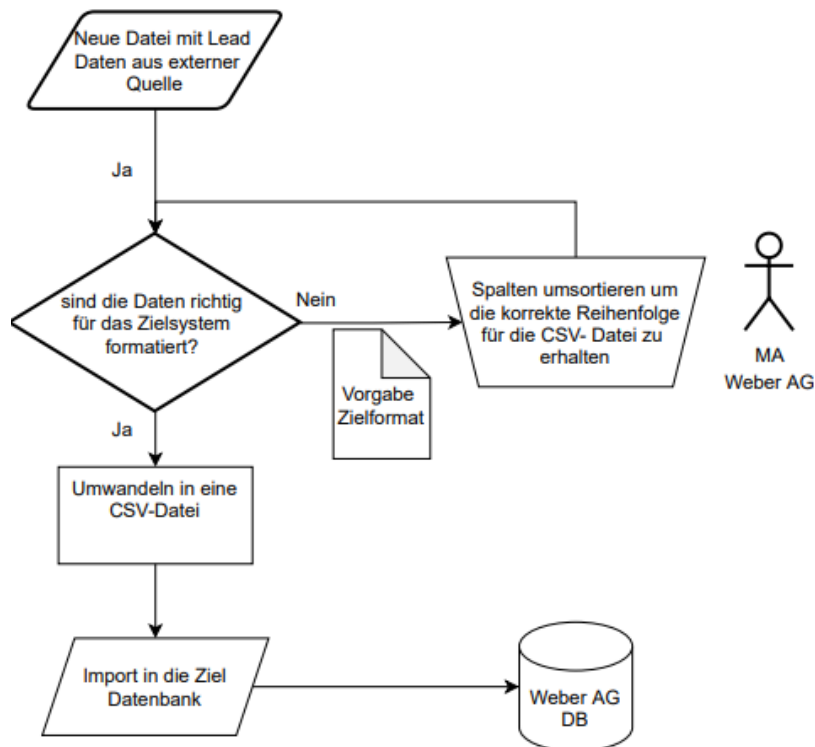
### 7.3 Schematische Darstellung

### 7.4 Beschreibung

Je nach Anwendungsfall kann eine manuelle oder eine automatische Verarbeitung vorgenommen werden.



#### 7.4.1 Manuelle Verarbeitung



Bei neuen Daten die nicht regelmäßig im gleichen Format geliefert werden kann eine manuelle Verarbeitung sinnvoll sein.

Hierbei werden die Daten im Vorfeld gesichtet und geprüft.

- Gibt es fehlende oder fehlerhafte Werte?
- Passen Datentyp, Datenlänge und Zeichensatz zum Zielformat.
- Liegen alle relevanten und benötigten Daten im Quellsystem vor oder können daraus abgeleitet werden? (z.B. Straße und Hausnummer in 1 Feld)
- Gibt es Duplikate im Datensatz?

Wenn Daten verwendet werden können müssen diese in ein entsprechendes Dateiformat transformiert werden um den Import in die Datenbank zu ermöglichen. Meisten wird dafür ein CSV Format verwendet. (Comma-separated values)

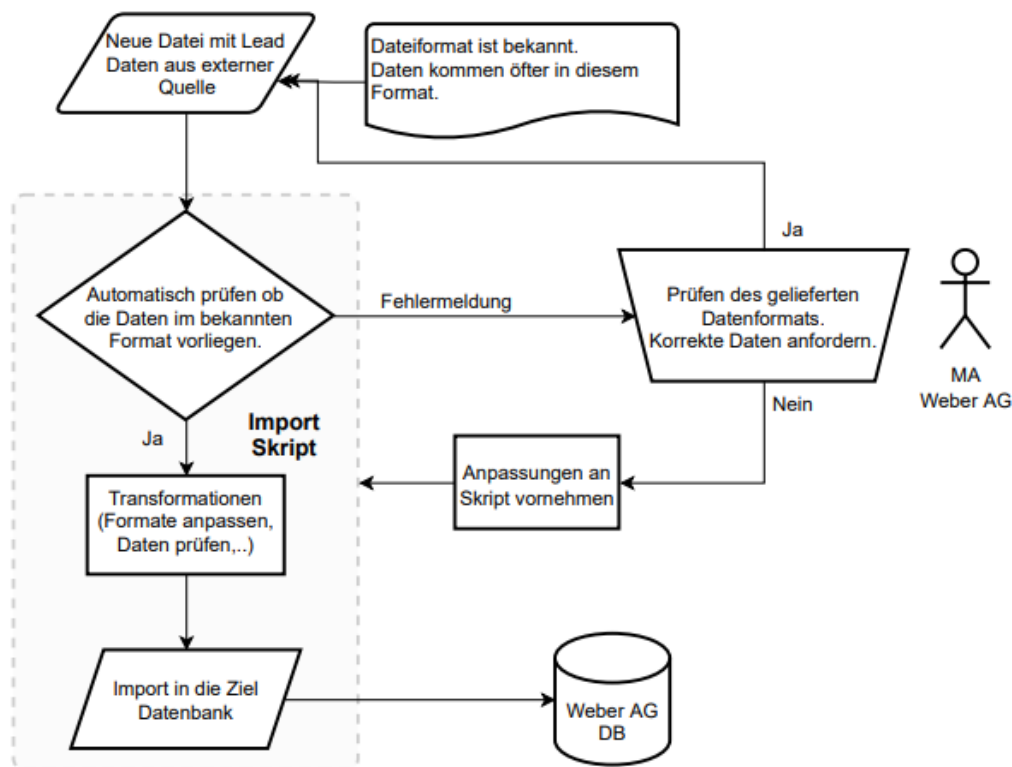
Danach kann, über eine Import Funktion der Datenbank, diese Datei eingelesen werden.

Hierbei ist darauf zu achten das es nicht zur Überschreibung von bereits existierenden Datensätzen der Datenbank kommt und wie mit solchen Daten ggf. umgegangen wird. (Definition eines Import Prozesses)

Im DB-System muss dann noch eine Prüfung der Datensätze im Bezug auf die gewünschte Verwendung für den Geschäftszweck, Inhalt/Aussage der Daten, hin geprüft werden. Dies kann nur sehr begrenzt im Vorfeld des Datenimports erfolgen.

Diese, erstmalig grob beschriebenen Schritte, sind notwendig um eine hohe Datenqualität der hinzuzufügenden Daten zu erreichen und eine ausreichende Wertschöpfung zu ermöglichen. Daten die nicht neu, vollständig und richtig sind haben keinen Wert für das Unternehmen und blockieren nur Speicherplatz.

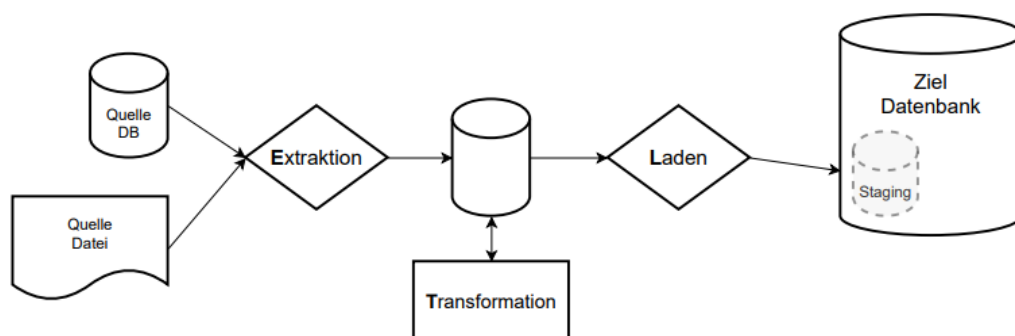
## 7.4.2 Automatisches Verarbeitung



Bei neuen Daten die regelmäßig im gleichen Format geliefert werden ist eine automatische Verarbeitung möglich.

Auch kann dieses Vorgehen mit der manuellen Verarbeitung kombiniert werden, wenn die Qualität der Quelldaten zu schlecht ist um in einem automatischen Ladeprozess geprüft zu werden.

Das Lesen, Transformieren und Importieren von Daten wird als ETL Prozess bezeichnet. (Extraction, Transformation, Load)



Dafür wird ein Softwareprogramm geschrieben welches die im Diagramm beschriebenen Schritte automatisch durchführt.

Um dieses zu ermöglichen muss die Struktur der Quelldaten dem ursprünglichen Quellformat entsprechen. Eine Änderung der Quellstruktur erfordert dann eine Anpassung des Import Skriptes! Aus diesem Grund ist es nur sinnvoll diese Verarbeitungsmethode zu verwenden wenn häufiger Daten in diesem Format eingeladen werden sollen.

Das Einladen der neuen Daten in die Datenbank erfolgt hierbei über das Skript. Um ein ungewolltes überschreiben von operativen Daten zu vermeiden sollten die Daten erst in eine Staging-Area der

Datenbank geschrieben werden um erst nach Überprüfung in das operative System geschrieben zu werden. Die Staging-Area dient als Zwischenspeicher für importierte Daten. Über Prozesse kann dann entschieden werden ob und wie die Daten ins operative System übernommen werden. Durch dieses Vorgehen kann die Qualität, und damit der Nutzen, der neuen Daten gezielt gesteuert und erhöht werden.

## 7.5 Datenqualität

Die Korrektheit, die Relevanz und die Verlässlichkeit von Daten, abhängig vom Zweck, die die Daten in einem bestimmten Zusammenhang erfüllen sollen wird als Datenqualität beschrieben.

Eine gute Datenqualität kann entscheidend für den geschäftlichen Erfolg sein da sie sich positiv auf viele Bereiche auswirkt.

- Weniger Aufwand für Import, Kontrolle und Anpassung der Daten
- Schnellerer und zuverlässiger Ablauf von Prozessen
- Zuverlässigere Datenbasis
- Verlässlichere und schnellere Business Analysen
- Vermeidung von redundanten Daten

Eine schlechte Datenqualität hingegen verursacht zusätzlichen Aufwand in vielen Bereichen.

Misstrauen gegenüber den Daten führt zu mehrfachem prüfen und weniger Akzeptanz von Reports und Analysen. Der operative Datenbestand wird z.B. durch doppelte oder veraltete Datensätze aufgebläht und langsamer gemacht.

Es ist somit sehr wichtig die Datenqualität von neuen externen Daten auf ein Level von >90% zu bringen bevor diese dem operativen System zur Verfügung gestellt werden.

Darauf muss bei der Einführung von einer Lösung zum Import von Fremddaten besonders geachtet werden. Das gleiche gilt für die Qualität der eingekauften Fremddaten, da nur maximal das importiert werden kann was zur Verfügung steht. (Waste in → Waste out!)

Zusätzlich spielt die Relevanz der gelieferten Daten für den Geschäftszweck eine wichtige Rolle und müssen ebenfalls mit in die Bewertung der Datenqualität einfließen.