



Fakultät Informatik, Mathematik und  
Naturwissenschaften  
Studiengang Informatik Master

**Abstrakt zum Oberseminar  
Datenbanksysteme: Aktuelle Trends**

**Data Lakes**

**Autor:** Raphael Drechsler  
**Matrikelnr.:** 69872  
**Abgabedatum:** ??.??..2018

## 1 MOTIVATION – TAUCHEN SIE EIN!

### DRT:

Begriff ist 2010 entstanden und in den letzten Jahren gehyped (Quellen 1,2,3)  
 Anschließend gab es Kritik bis hin zu einem Artikel über Fake News? (Quelle 4)  
 Zum Zeitpunkt der Recherche kein deutsch-Sprachiger Artikel auf Wiki  
 Motivation besteht darin diese Unklarheiten zu beleuchten, klären was ein Data-Lake ist und sich mit der Frage "Fake-News" auseinanderzusetzen.

## 2 DEFINITIONSFRAGE "DATA LAKE"

### DRT:

Kein Akademischer Ursprung. Dixon hat's gemacht.

### 2.1 Definition nach Dixon

Quellen: 5,6

### 2.2 Pentaho-Architektur 2010

Quellen: 6

### 2.3 Begriffs-Chaos

Quellen: 7

## 3 WIE FUNKTIONIERT EIN DATA LAKE?

### 3.1 Aufbau

QUELLEN : 8, 9

### 3.2 Workflow

Allgemein und die Rolle des Data Scientist QUELLEN : 8, 9

**STORAGE** QUELLEN : 8, ?

**INGESTION** QUELLEN : 8, 10, 11

**PROCESS** QUELLEN : 8, 12

**CONSUMPTION** QUELLEN : 8, 12

**MONITORING** QUELLEN : 8, ?

**DATA GOVERNANCE** QUELLEN : 8, 12

## 4 DATA SWAMPS: KRITIK AM DATA LAKE

Mögliche Darstellungen die es so gibt:

- Sumpf
- Finnland
- Flohmarkt

Gartners wesentliche Punkte

Battle: Gartner vs. Dixon und

QUELLEN : 3, 12, 13

DL revisited → ungenaue Definition ein Problem zu kritisieren oder nur ein Kritikpunkt?

Quellen: 14,15,

In Praxis dann Sean Martin zitieren.

Trend des Vorsichtiger werdens und die Flut kommen sehen. Paradigmenwechsel.

Quelle: 1

## 5 FAKE-NEWS! EXISTIEREN DATA LAKES ÜBERHAUPT?

Blogeintrag nur nennen und erste Zeile zitieren.

Quelle: 4

Nach Recherche lassen sich da schon ein paar Firmen finden, die Data Lake Lösungen anbieten. → auflisten

Nach weiterer Recherche auch success-stories auffindbar. Also irgendwie schon.

Die wesentliche Frage ist allerdings die Definitionsfrage. Selber Schluss im Blogeintrag. Lösung die dem Paradigma grundlegend folgen gibt es. Jetzt im Auge des Betrachters ob man das Kind beim Namen nennt oder nicht.

## LITERATUR/QUELLEN