

## 4. Übungsblatt

Ausgabe: 07. November 2016 · Besprechung: Ab 14. November 2016

### 1 Storing Records on Pages

1. Wie können Records auf einer Seite gespeichert werden? Beschreiben Sie dazu Aufbau und Konzept von Record Identifiers (RIDs)!
2. Wie können wir mithilfe eines Record Identifiers ein Tupel finden?
3. Wie können Felder variabler Länge gespeichert werden?

### 2 Bw-Trees

Levandoski et al. beschreiben in *"The Bw-Tree: A B-tree for New Hardware Platforms"* eine neue Index Struktur. Beantworten Sie die folgenden Fragen mit Hilfe des Papiers:

- Welche Annahmen haben sich in modernen Datenbanken geändert, die den Einsatz von Bw-Trees anstelle von B+ Bäumen begünstigen?
- Wie funktionieren *Delta Updates* und was ist ihr Vorteil?

### 3 Curse of Dimensionality

1. Welches Problem ist mit „Curse of Dimensionality“ gemeint?
2. Was bedeutet das für Datenbanken in der Praxis?
3. Schreiben Sie ein Programm, das zufällig Vektoren in einem n-dimensionalen Raum generiert. Der Wertebereich jedes Dimensionswertes soll zwischen 0 und 9 liegen. Nehmen Sie für die Werteverteilung eine Gleichverteilung an. Erzeugen Sie wiederholt Paare von zufälligen Vektoren, und berechnen Sie den Euklidischen Abstand. Wie ändern sich Mittelwert, Varianz und Standardabweichung der Distanzen zwischen den Vektoren für 3, 10, 100, und 1000 Dimensionen?