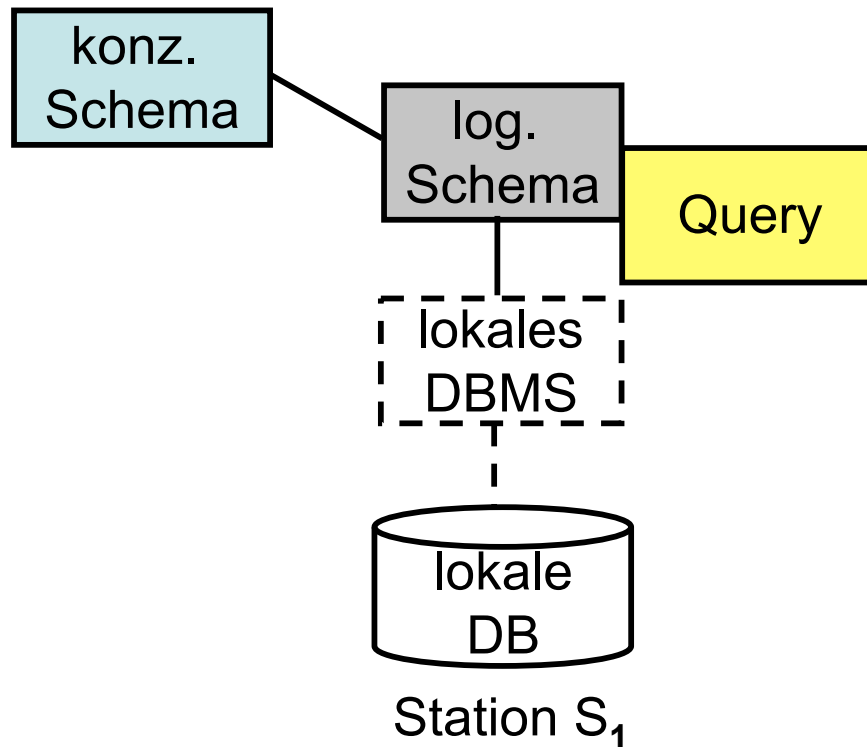


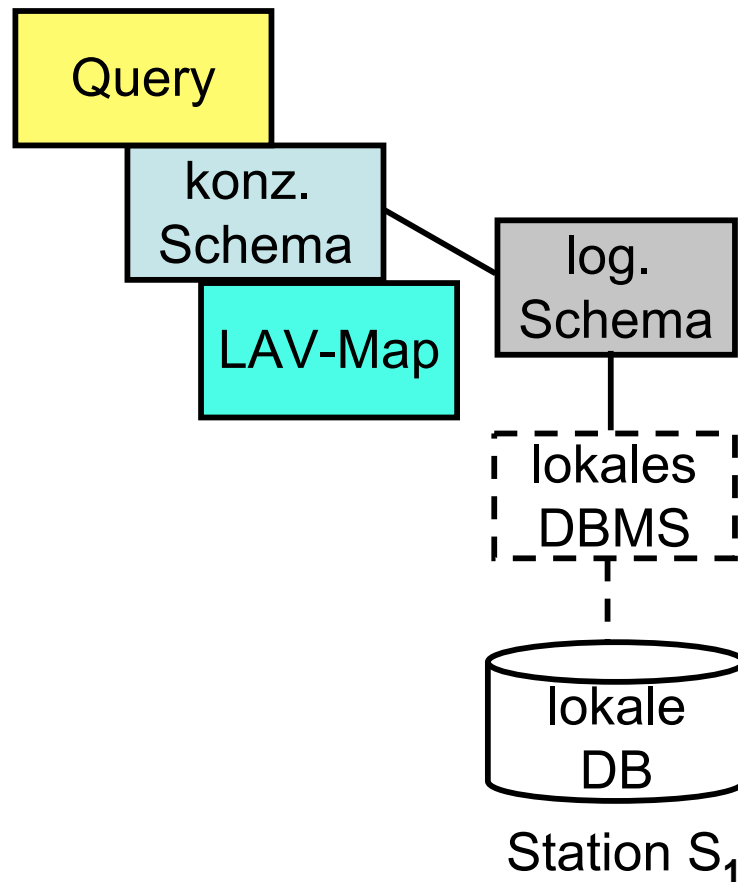
Konzeptuelle Informationsintegration

R. Möller

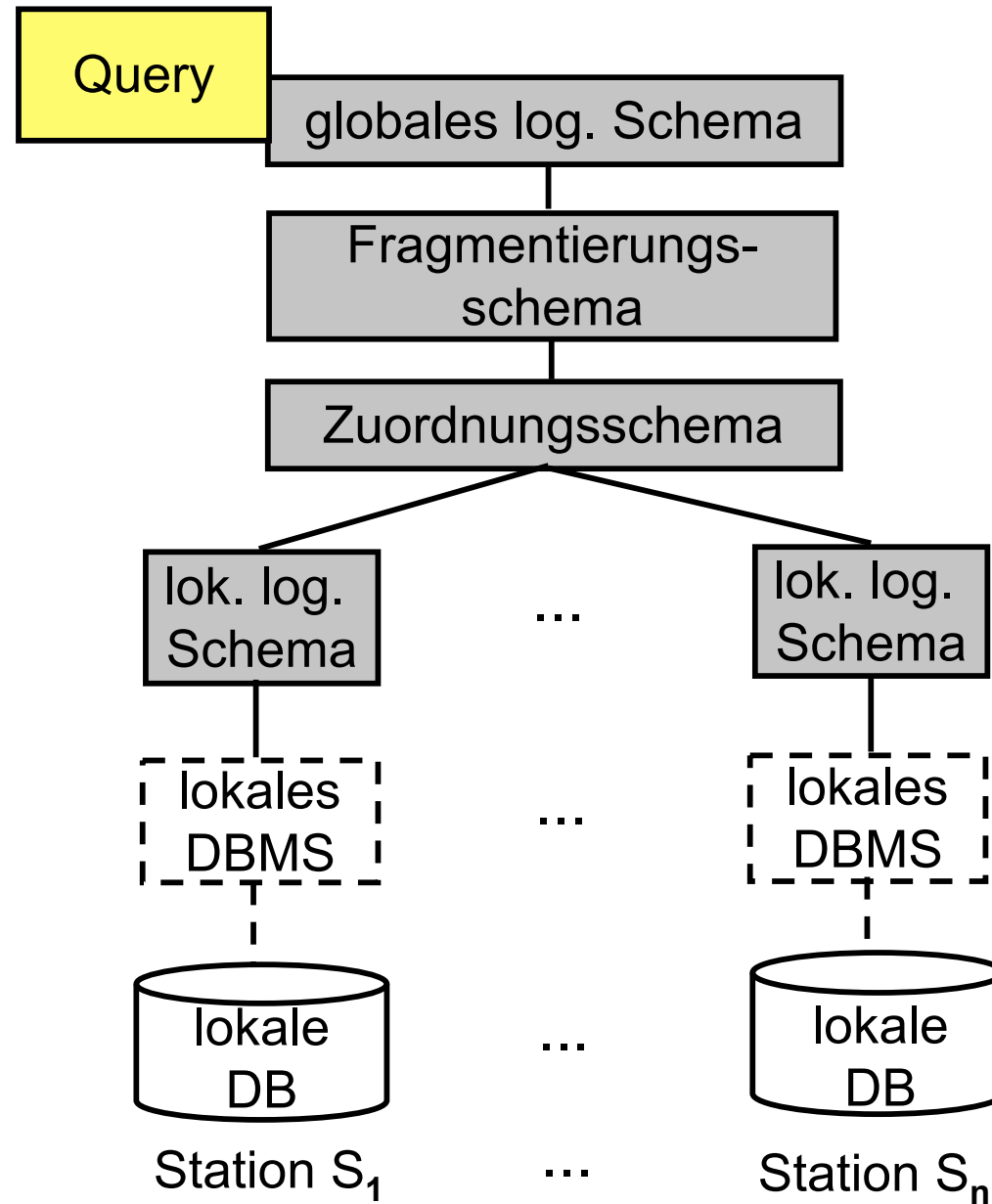
Szenario: Anfragen bzgl. Log. Schema



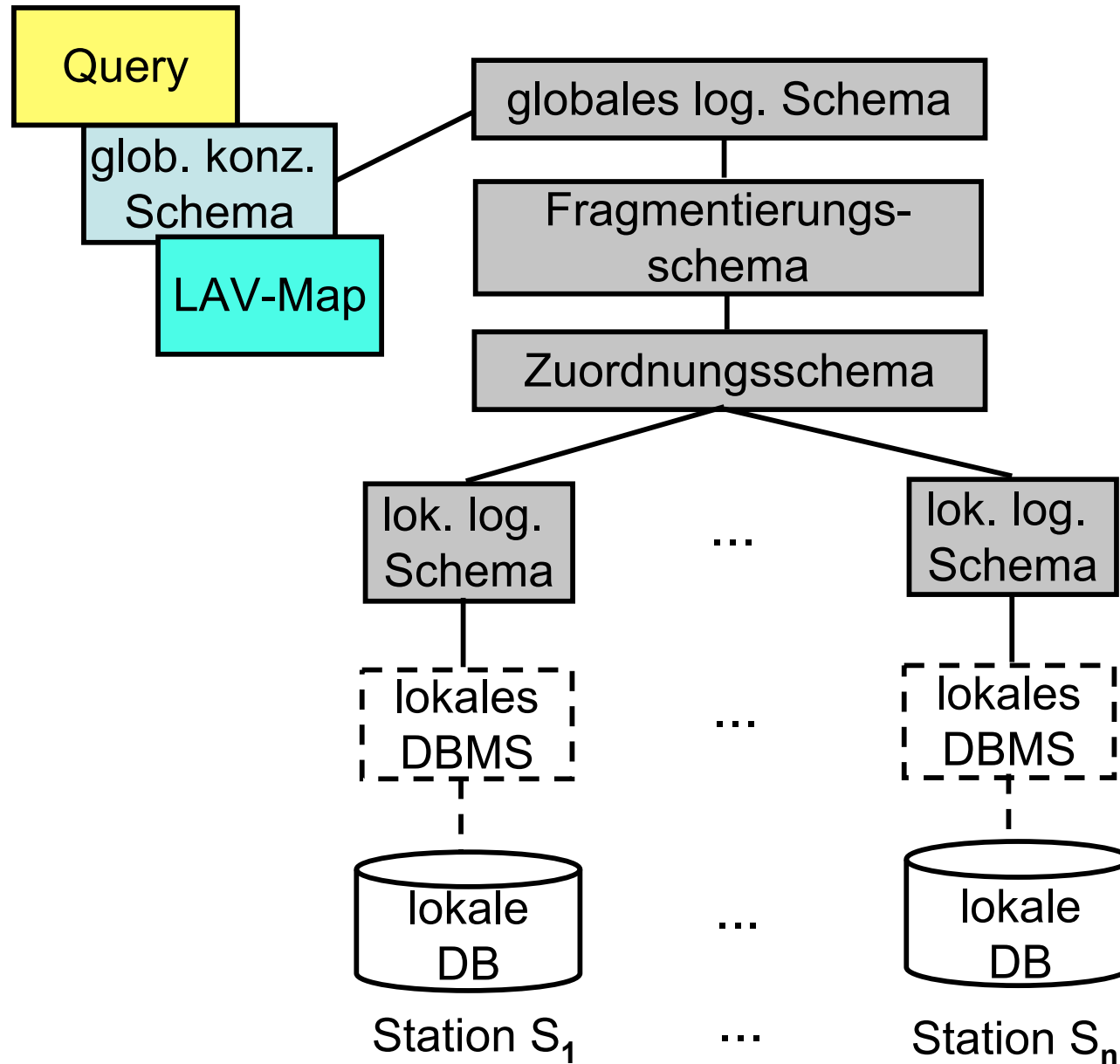
Szenario: Anfragen bzgl. Konz. Schema



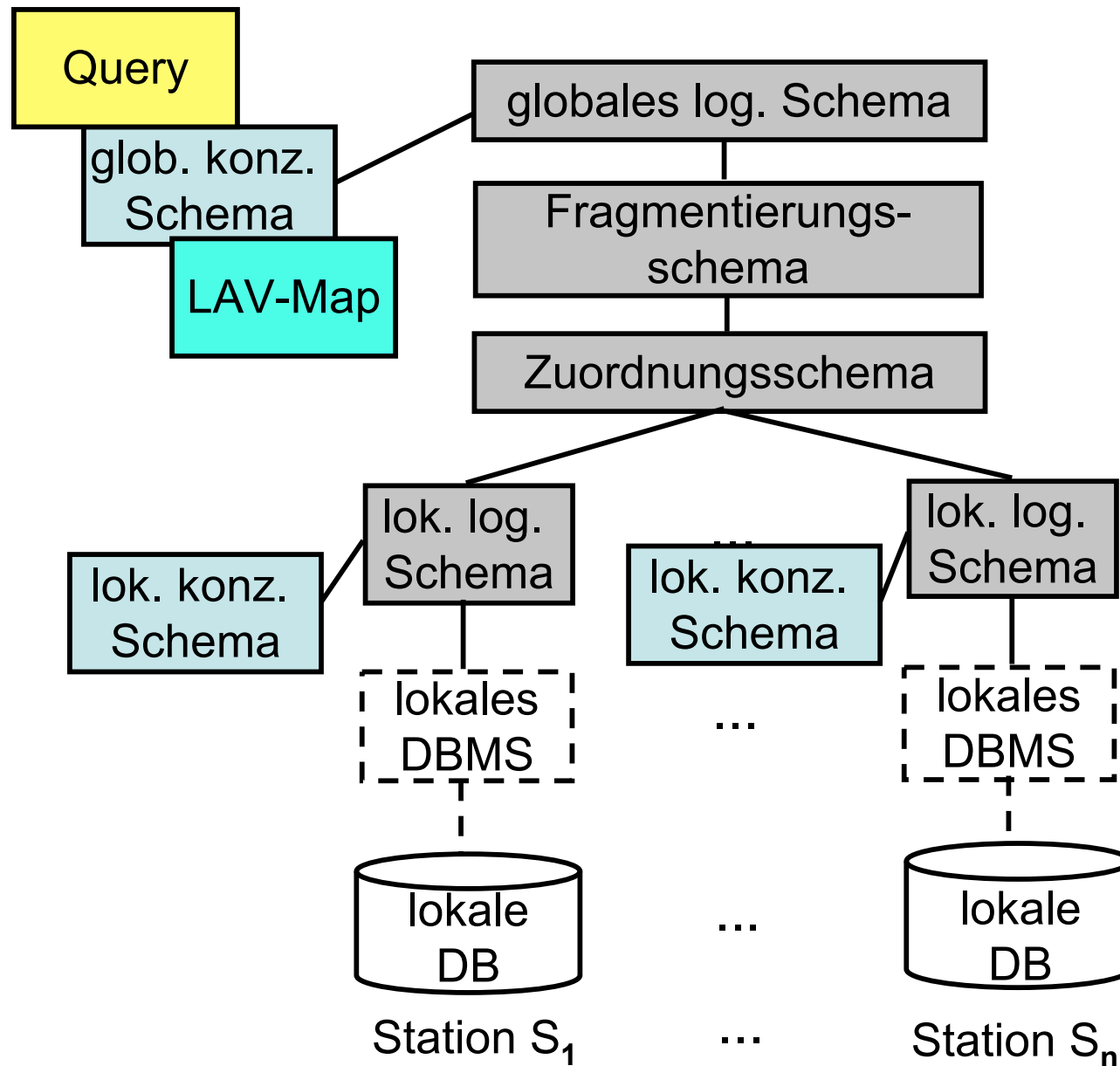
Szenario: Verteilte Datenbanken



Szenario: Informationsintegration



Ziel: Konzeptuelle Informationsintegration



Unterschiedliche Domänen

- Die Grundmenge der unterschiedlichen Informationsquellen müssen nicht gleich sein
- Universum besteht aus möglicherweise überlappenden Grundmengen
- Formalisierung notwendig (siehe Tafel)

Interschema Assertions

$$C_i \stackrel{\cdot}{=}_{int} C_j$$

$$C_i \sqsubseteq_{int} C_j$$

$$C_i \stackrel{\cdot}{=}_{ext} C_j$$

$$C_i \sqsubseteq_{ext} C_j$$

Conceptual level. Ctd.

- Inter-schema relationships

$$\begin{array}{l} L_i \sqsubseteq_{ext} L_j \\ L_i \sqsubseteq_{int} L_j \end{array}$$

- **The first assertion: L_i is extensionally included in L_j , which means that every object that satisfies the expression L_i in source i also satisfies the expression L_j in source j**

$$\text{Student}_1 \sqsubseteq_{ext} \text{Student}_2$$

- **The second assertion: the concept denoted by the expression L_i in source i is a subconcept of the one denoted by the expression L_j in source j , which means that every object in source i satisfying L_i also satisfies L_j in source j , provided that it does appear in source j**

$$\text{Student}_1 \sqsubseteq_{int} \text{Person}_2$$

Umwandlung int in ext (1)

- Für die Transformation müssen die Domänen der einzelnen zu integrierenden Systeme durch spezielle Konzepte für Top (\top) im integrierten Schema benannt werden können. Die Vereinigung der n einzelnen Domänen ist dann die Domäne des zentralen Systems:

$$\top_0 \doteq_{ext} \top_1 \sqcup \dots \sqcup \top_n$$

- Extensionale Aussagen können ohne Veränderungen verwendet werden, da sie den normalen Intraschemaaussagen entsprechen. Die intensionalen Aussagen werden dann unter Berücksichtigung der Domänen auf extensionale Aussagen zurückgeführt. Aus $C_i \sqsubseteq_{int} C_j$ wird

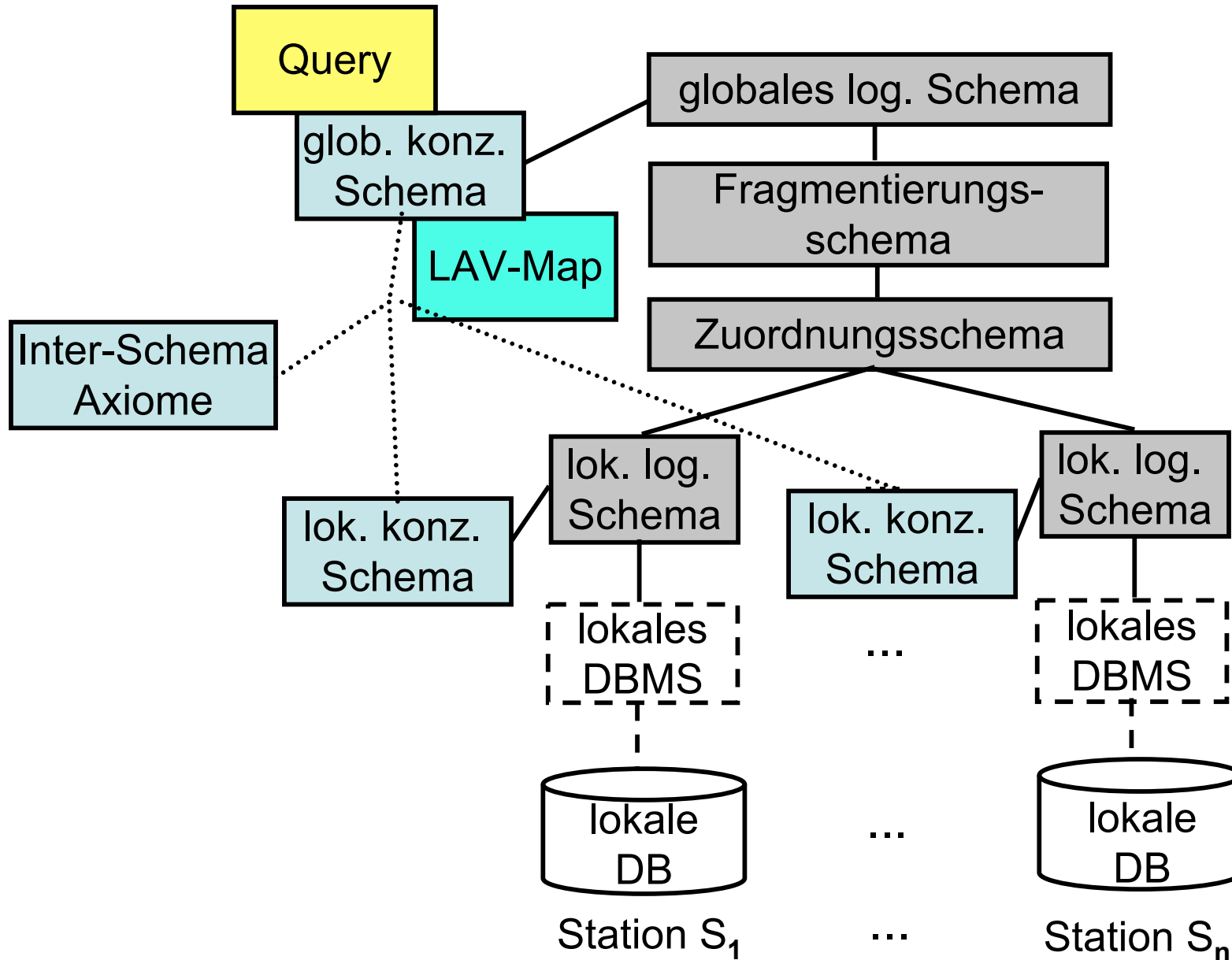
$$C_i \sqcap \top_j \sqsubseteq_{ext} C_j$$

und aus $C_i \doteq_{int} C_j$ wird

$$C_i \sqcap \top_j \doteq_{ext} C_j \sqcap \top_i.$$

- Für alle Namen A und Domänen $i : A_i \sqsubseteq \top_i$

Konzeptuelle Informationsintegration



Heterogenität der Datenmodelle

- Problem: Logische Schemata nicht notwendigerweise kompatibel
- Beispiel: Relationales vs. Semistrukturiertes Datenmodell

Offene Fragen

