HSLU Lucerne University of Applied Sciences and Arts

Conflation für POI



Business

FH Zentralschweiz

AACSB ACCREDITED

Was sind POIs?

POI = Point of Interest (Interessenpunkt)

Georeferenzierte Punkte mit spezifischen Informationen

Beispiele: Restaurants, Tankstellen, Sehenswürdigkeiten, Geschäfte

Enthalten typischerweise:

- Geographische Koordinaten (Längen- und Breitengrad)
- Kategorie
- Name
- Attributdaten
  - Öffnungszeiten
  - Kontaktinformationen
  - Bewertungen
  - Zugänglichkeit

Das Problem: Mehrere Datenquellen

Verschiedene Anbieter erfassen POI-Daten:

- Kommerzielle Kartenanbieter (Google Maps, HERE, TomTom)
- Open-Source-Plattformen (OpenStreetMap)
- Tourismusorganisationen (Züich Tourismus, Basel Tourismus, Graubünden Ferien teilweise Schweiz Tourismus)
- Diverse behördliche Abteilungen

Jede Quelle hat anderen

- Aktualisierungsrhythmus
- Detailierungsgrad und Inhalte (Beschreibung, Bilder etc.)
- Datenstandard
- Qualitätssicherungsverfahren
- Abdeckung

Was ist POI Conflation?

Conflation = Prozess der Zusammenführung mehrerer Datensätze zu einem qualitativ hochwertigeren Ergebnis

Bei POIs bedeutet dies:

- Identifizierung identischer realer Objekte in verschiedenen Datensätzen
- Zusammenführung redundanter Einträge
- Ergänzung fehlender Informationen
- Auflösung von Widersprüchen
- Schaffung eines konsistenten, umfassenden Datensatzes

Nebeneffekt kann sein: Rückmeldung von systematischen Inkonsistenzen an die Datenlieferante

Der Conflation-Prozess

Datenvorverarbeitung durch Standardisierung von Formaten und Attributen

Matching (Zuordnung) von POI aus verschiedenen Datenbanken durch Identifizierung potenzieller Übereinstimmungen zwischen Datensätzen unter Verwendung räumlicher und semantischer Ähnlichkeiten (z.B. Geo-Koordination, Kategorie, Name)

Fusionierung

- Zusammenführung übereinstimmender Einträge
- Auswahl der zuverlässigsten Informationen

Qualitätssicherung

- Überprüfung der Ergebnisse
- Behandlung von Ausnahmen

Herausforderungen bei der POI-Conflation

Positionsunterschiede: Unterschiedliche GPS-Koordinaten für dasselbe Objekt

Namensabweichungen: Unterschiedliche Schreibweisen, Abkürzungen, Sprachen

Zeitliche Unterschiede: Verschiedene Erfassungszeitpunkte

Strukturelle Unterschiede: Verschiedene Detailgrade und Attributmodelle

Mehrdeutigkeiten: Mehrere ähnliche POIs in unmittelbarer Nähe

Praktisches Beispiel: Restaurant-Daten

Ausgangssituation:

- **Datenquelle A (OpenStreetMap):**
  - Name: "Ristorante Firenze"
  - Koordinaten: 48.1364, 11.5729
  - Kategorie: Restaurant
  - Küche: Italienisch
  - Öffnungszeiten: N/A
- **Datenquelle B (Tourismusorganisation)**
  - Name: "Restaurant Firenze"
  - Koordinaten: 48.1366 N, 11.5730 E
  - Kategorie: Restaurant
  - Subkategorie Küche: N/A
  - Öffnungszeiten: Mo-Sa, 11:30-22:00

Fusionierter POI-Datensatz:

- Name: "Ristorante Milano" (bevorzugt vollständigeren Namen)
- Koordinaten: 48.1365, 11.5730 (Mittelwert oder Präzisere)
- Kategorie: Restaurant
- Küche: Italienisch (ergänzt aus Quelle A)
- Öffnungszeiten: Mo-Sa, 11:30-22:00 (ergänzt aus Quelle B)

Mehrwert: Ein vollständigerer und genauerer Datensatz

Zusammenfassung: Mehrwert von POI-Conflation

Höhere Datenqualität: Genauere Positionen und Informationen

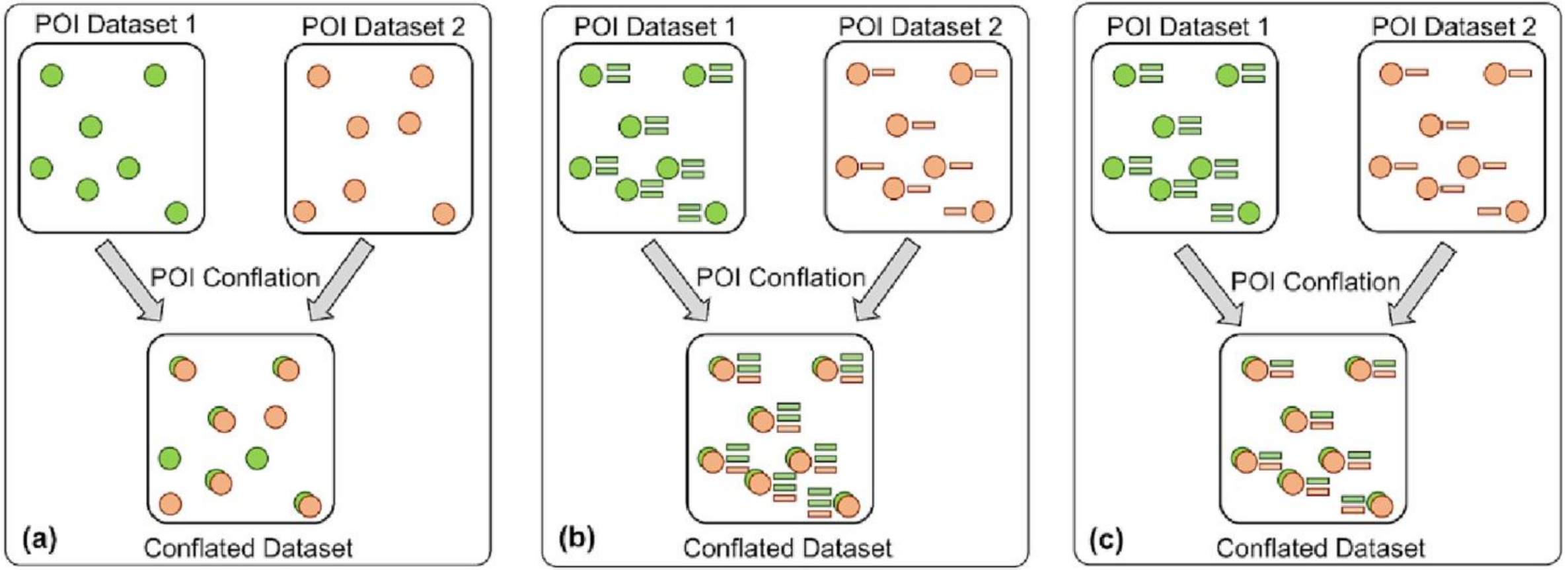
Verbesserte Vollständigkeit: Kombination komplementärer Informationen

Höhere Aktualität: Integration neuester Daten aus verschiedenen Quellen

Beseitigung von Inkonsistenzen: Einheitliche und widerspruchsfreie Datenbasis

Kosteneffizienz: Nutzung vorhandener Datenquellen statt Neuerfassung

Three common situations under which POI conflation can be helpful



(a) two datasets cover partially overlapping places
(b) two datasets provide complementary POI attributes
(c) one dataset provides more accurate attributes than those of the other dataset.

Kai Sun, Yingjie Hu, Yue Ma, Ryan Zhenqi Zhou, Yunqiang Zhu (2023): Conflating point of interest (POI) data: A systematic review of matching methods. In: Computers, Environment and Urban Systems. Vol. 103
https://www.sciencedirect.com/science/article/pii/S0198971523000406?via%3Dihub

HSLU Lucerne University of Applied Sciences and Arts

Merci

Lucerne School of Business

Prof. Dr. Andreas Liebrich
Head of International Accreditation
Deputy Head of International Office
Lecturer for Tourism

Phone direct +41 41 228 42 34
andreas.liebrich@hslu.ch

Nächste Schritte (Vorschlag)

1. Schärfung Use Case Tourismus und Erstellung Katalog, welche POI gebraucht werden