

Zoulfa El Jerroudi und Jürgen Ziegler

# Interaktives Vergleichen und Zusammenführungen von Ontologien

## Interactive Ontology-Mapping and Merging

Ontologie-Mapping\_Ontologie-Matching\_Ontologie-Merging\_Wissensintegration

**Zusammenfassung.** In dieser Arbeit werden semi-automatische Verfahren vorgestellt, die den Domänenexperten beim *interaktiven Vergleichen* und *Zusammenführen* von Ontologien unterstützen sollen. Das Ziel beim Vergleichen von zwei Ontologien besteht darin, eine Abbildung zwischen der Quell- und Zielontologie zu finden, die dem Nutzerverständnis einer semantischen Entsprechung am nächsten kommt. Für die Initialisierung des Vergleichsprozesses startet der *iMERGING-Editor* mit der linguistischen Ähnlichkeit von zwei Konzepten, danach werden die strukturellen Eigenschaften der Ontologie berücksichtigt und im dritten Schritt wird die Ähnlichkeit aufgrund zusätzlich vorhandener Informationen wie die Verknüpfung mit Dokumenten untersucht. Nutzereingaben, wie das Ablehnen oder Ändern eines Mappings, werden im Vergleichsprozess interaktiv berücksichtigt, so dass Domänenwissen des Nutzers mit einbezogen werden kann, um semantische Entsprechungen zwischen zwei Elementen der Wissensstruktur zu finden.

**Summary.** In this paper we present a semi-automatic method, which helps the user in the process of interactive ontology mapping and interactive ontology merging. The main goal is to produce a mapping that tries to approximate the understanding of what the users consider to be a good match. We start by using linguistic similarity matches between concepts for initiating the matching process, and then use the underlying ontological structures and auxiliary information to inform a set of heuristics for identifying further matches between ontologies. User inputs like rejecting or changing mappings are taken into account, so that the plausibility of the correspondences can increase.

## 1. Einleitung

In der Ontologieentwicklung werden die Aktivitäten des *Vergleichens* und *Zusammenführens* benötigt, wenn die zu entwickelnde Ontologie aus mehreren existierenden Ontologien zusammengestellt werden soll, wenn Unterschiede zwischen verschiedenen Ontologie-Versionen erkannt werden müssen oder wenn mehrere Gruppen mit spezialisiertem Fachwissen kollaborativ an der gleichen Ontologie arbeiten, aber Teilstrukturen unterschiedlich verfeinern. Beim Vergleichen und Zusammenführen von Ontologien aus unterschiedlichen Quellen führen verschiedene Terminologien und semantische Unterschiede in der Repräsen-

tation der Daten häufig zu Konflikten. Obwohl diese oftmals nur durch Domänenexperten endgültig aufgelöst werden können, muss die Arbeit der Experten durch geeignete Werkzeuge unterstützt werden. In dieser Arbeit werden semi-automatische Verfahren vorgestellt, die den Domänenexperten beim *interaktiven Vergleichen* und *Zusammenführen* von Wissensstrukturen unterstützen sollen. Der hier vorgestellte Ontologie-Editor *iMERGING* (*interactive Merging*) besitzt dazu verschiedene Vergleichssichten (siehe Bild 2), die gleiche und ähnliche Konzepte bzgl. einer Vergleichsontologie anzeigen. Verschiedene Vergleichsalgorithmen können vom Benutzer ausgewählt werden. Neben einem grundlegenden Vergleich auf Basis des Konzeptnamens werden Konzepte auch bzgl. struk-

tureller Eigenschaften wie gemeinsame *Subkonzepte* und *Properties* verwendet. Nutzereingaben, wie das Ablehnen oder Ändern eines Mappings, werden im Vergleichsprozess interaktiv berücksichtigt. Basierend auf den vorangehenden Vergleichen können zwei Vergleichsontologien zu einer gemeinsamen Konsensontologie zusammengeführt werden (siehe Bild 4).

Für das Vergleichen von Ontologien wurden bisher schon einige Verfahren entwickelt, wie FOAM (Ehring et al. 2005), QOM (Ehrig und Staab 2004), PROMPT (Noy und Musen 2003), GLUE (Doan et al. 2003) und FCA-Merge (Stumme und Maedche 2001). Eine gute Übersicht über Arbeiten zu Ontologie Mapping geben Noy (2004) und Kalfoglou (2005).

In allen erwähnten Verfahren wird auf linguistische Informationen, in einigen auch auf strukturelle Informationen der Ontologie zurückgegriffen, jedoch ohne die interaktiven Aspekte des Vergleichens und der Zusammenführung zu berücksichtigen (Ausnahme: PROMPT). Betrachtet man die große Anzahl von Arbeiten, die das Vergleichen von Ontologien behandeln, so findet man nur sehr wenige, die sich damit beschäftigen, wie man den Nutzer interaktiv einbinden kann. Falconer und Noy (2006) haben Aufgaben definiert, die bei einem Ontologie-Vergleich den Nutzer unterstützen sollen. Der in dieser Arbeit vorgestellte Ontologie-Editor berücksichtigt davon folgende wichtige Anforderungen:

- Sowohl die Quell- als auch die Zielontologie sollen explorierbar sein.
- Gefundene Abbildungen (Mappings) sollen übersichtlich dargestellt werden.
- Mapping-Kandidaten sollten im Kontext der ursprünglichen Ontologien angezeigt werden.
- In interessante Teilbereiche der Ontologie sollte hineingezoomt werden, z. B. sollte für die gefundenen Mapping-Kandidaten, die Nachbarschaft exploriert und Detailinformationen eingesehen werden können, um das Mapping nachzuvollziehen.
- Ergebnisse sollen gesichert werden können.
- Der Nutzer solle die Möglichkeit haben, selber Abbildungspaare hinzuzufügen.
- Der Nutzer soll über die aktuellen Vergleichsschritte informiert werden.

Lediglich die Verifikation der vorgestellten Mappings wird bisher noch nicht durchgeführt. D. h. das hier vorgestellte System zeigt ähnliche Kandidaten aus den unterschiedlichen Ontologien an, berücksichtigt jedoch nicht evtl. Konflikte, die bei der Zusammenführung dieser Kandidaten auftreten könnten. Im Folgenden wird ein kurzes Szenario beschrieben, um die prinzipielle Verwendung des Ontologie-Editors *iMERGING* zu verdeutlichen.

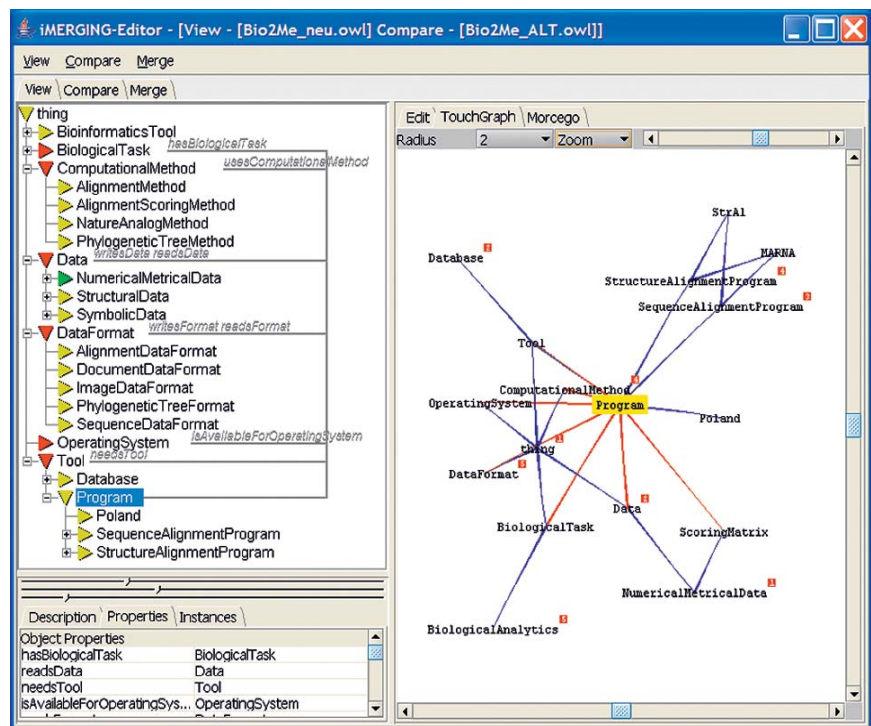


Bild 1: Visualisierung der Bio2Me-Ontologie im iMERGING-Editor

## 2. Szenario zur integrierten Bearbeitung von Ontologien

Innerhalb der Forschungs-Gemeinschaft der Bioinformatik werden zahlreiche bio-medizinische Datenbanken und andere Software-Tools eingesetzt, die für Datenanalysen und Experimente benötigt werden. In dem Verbundprojekt ONTOVERSE<sup>1</sup> wurde die Ontologie *Bio2Me* erstellt, um die vielfältigen Applikationen zu ordnen und so die Suche nach geeigneten Tools zu beschleunigen. Eine wichtige Zielsetzung der Ontologie ist eine intuitive Klassifikation, so dass eine eindeutige Einordnung neuer Werkzeuge von mehreren Ontologie-Designern vorgenommen werden kann. Der Einsatz des Ontologie-Editors *iMERGING* lässt sich mit dem folgenden Szenario veranschaulichen:

Die Arbeitsgruppe A hat nun eine neue Version der *Bio2Me*-Ontologie veröffentlicht und die Forschungsgemeinschaft mit Hilfe der ONTOVERSE-Plattform darüber informiert. Die Arbeitsgruppe B möchte die neue Version näher betrachten und lädt die neue Ontologie im *iMERGING*-Editor ein (siehe Bild 1). Zur Exploration der Ontologie bietet der Editor verschiedene Sichten.

Die Arbeitsgruppe B interessiert sich hauptsächlich für das Konzept *Program*. Wird *Program* in der taxonomischen Sicht auf der linken Seite der Abbildung selektiert, erscheinen zusätzliche Beziehungen zu anderen Konzepten (Propertielines).

Im unteren Teil der Abbildung werden Detailinformationen zum selektierten Konzept präsentiert (siehe Bild 1). Auf der rechten Seite wird der entsprechende Ausschnitt der Ontologie als Teilgraph visualisiert (die Größe des Ausschnitts kann eingestellt werden).

Die Arbeitsgruppe B möchte nun das Konzept *Program* im Vergleich zur alten Version betrachten. Dazu wird die alte Version ebenfalls in den *iMERGING*-Editor geladen, um die beiden Versionen einander gegenüberzustellen.

Der *iMERGING*-Editor ermöglicht den paarweisen Vergleich aller Konzepte der beiden Ontologien. Um eine generelle Übersicht darüber zu erhalten, in welchen Teilen der Ontologie Veränderungen aufgetreten sind, eignet sich die *Matrix-Sicht* (siehe Bild 2, rechts). Dabei werden die zu vergleichenden Ontologien auf zwei Achsen gegenübergestellt. Eine hohe Übereinstimmung der beiden Ontologien wird mit grünen Verbindungspunkten signalisiert, Teilbereiche in denen Unterschiede auftauchen, werden mit roten Verbindungspunkten signalisiert.

<sup>1</sup> Das ONTOVERSE-Projekt wird vom Bundesministerium für Bildung und Forschung gefördert (Projekt-Nr. 01C5975).

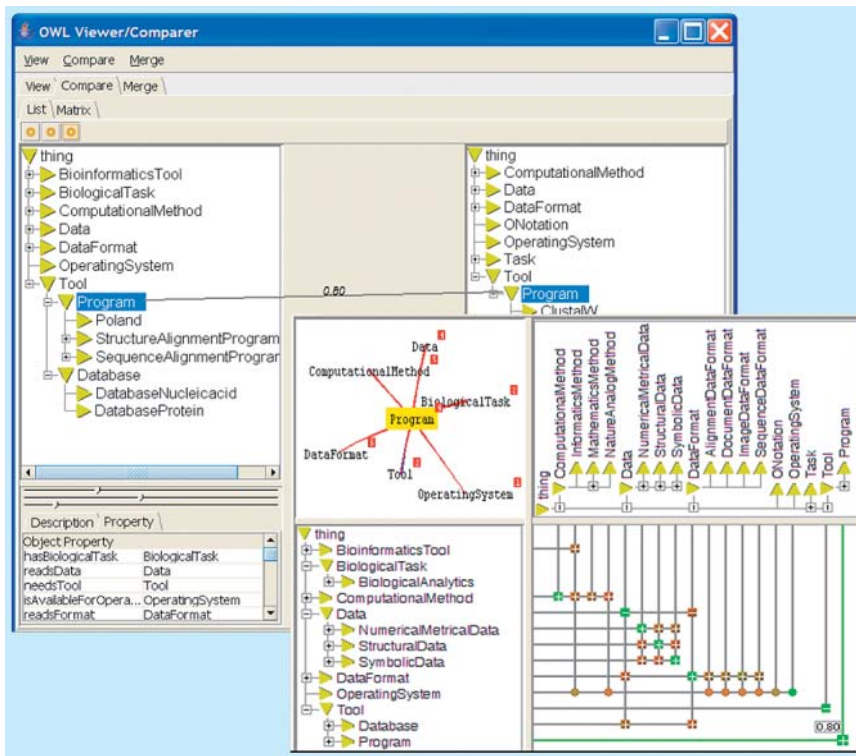


Bild 2: Vergleich der zwei Versionen von Bio2Me

siert. Ein kleines Plus in der Matrix deutet an, dass sich in dem Unterbaum weitere ähnliche Konzepte verstecken. Wenn ein Konzept selektiert wird, wird ein genauer Zahlenwert für das Ähnlichkeitsmaß angezeigt, sowie die Nachbarschaft des Knotens in einem Teilgraph visualisiert.

Für eine genauere Betrachtung der Mapping-Kandidaten eignet sich eher die Listen-Sicht (Bild 2, links). In dieser Sicht

werden die Mapping-Kandidaten im Kontext der Ontologie angezeigt, wobei Detailinformationen, wie *Eigenschaften*, *Instanzen* und *Kommentare* im unteren Teil des Fensters erscheinen. Die Arbeitsgruppe B stellt nun fest, dass die Programme in der neuen Version der Ontologie klassifiziert wurden (in *StructureAlignmentProgram* und *SequenceAlignmentProgram*). Trotz der syntaktischen Über-

einstimmung der beiden Begriffe (aus den versch. Versionen) wird aufgrund der strukturellen Unterschiede (Konzepte haben unterschiedliche Subklassen) ein Ähnlichkeitswert von 0,8 berechnet (siehe Bild 2).

Die Arbeitsgruppe B bemerkt, dass das Programm *Poland* nicht in die neue Klassifizierung eingeführt wurde. Dies und andere Änderungen widersprechen den Auffassungen dieser Arbeitsgruppe, so dass sie sich entschließt, ausgehend von der ursprünglichen Ontologie eine weitere Version der *Bio2Me* zu erstellen (siehe Bild 3).

Nachdem die Forschergruppe B ihre Änderungen vorgenommen und nun auch eine neue Ontologie-Version erstellt hat, entschließt sie sich dazu, mit Hilfe der Ontoverse-Plattform, einer Web-Plattform zur kollaborativen Ontologiebearbeitung, mit der Arbeitsgruppe A zu kooperieren und eine gemeinsame Ontologie zu veröffentlichen. Dazu treffen sich die beteiligten Forscher online, um die beiden Versionen zusammenzuführen (vgl. Bild 4). Um dies zu unterstützen, werden die beiden Ausgangsontologien farblich unterschiedlich markiert, z.B. blau und rot.

In die Ergebnisontologie *G* werden zunächst Konzepte aus den beiden Ontologien automatisch übernommen, für die es keine Abbildung in der Vergleichsontologie gibt. Konzepte, für die es eine Entsprechung in der Vergleichsontologie gibt, müssen zusammengeführt werden. Durch die farbliche Unterscheidung, kann der Nutzer jedoch den Ursprung der Konzepte erkennen. Der Domänenexperte kann die vorgeschlagenen Zusammenführungen *annehmen*, *ablehnen* oder *ändern*, indem er die Konzepte mit der Maus selektiert.

Die vom Domänenexperten akzeptierten Vorschläge und Änderungen nehmen in der Zielontologie eine andere Farbe an (z. B. grün). Nachdem sich die beiden Arbeitsgruppen auf eine gemeinsame Konsensontologie geeinigt haben, wird diese als akzeptierte Version veröffentlicht.

### 3. Eine integrierte Vorgehensweise für den Vergleich von Ontologien

Das Ziel beim Vergleich zweier Ontologien besteht darin, eine Abbildung zwi-

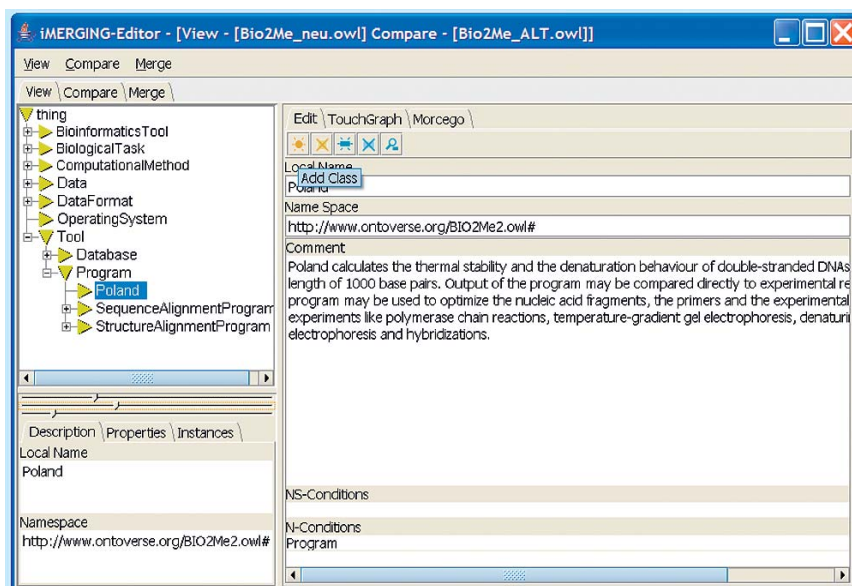


Bild 3: Hinzufügen einer neuen Konzepts

schen der Quell- und Zielontologie zu finden, die dem Nutzerverständnis einer semantischen Entsprechung am nächsten kommt. Hierfür definieren wir eine Abbildung als eine *Ähnlichkeitsrelation map*:  $(O_1 \rightarrow O_2)$  über die Elemente einer Wissensstruktur. Das Ergebnis einer Vergleichsoperation zwischen zwei Wissensstrukturen ist eine Menge von Abbildungselementen, die angeben, dass ein Element  $x$  aus  $O_1$  auf ein Element  $y$  aus  $O_2$  abgebildet werden kann ( $map(x) = y$ ). An jedem Paar von Abbildungselementen ist ein *Ähnlichkeitswert*  $sim(x, y)$  angehängt, der die Plausibilität der semantischen Entsprechung angibt.

Um eine hohe und auch genaue Abbildungsrate zwischen einer Vielzahl unterschiedlicher Wissensstrukturen zu erzielen, genügt es nicht nur, ein Vergleichskriterium, wie zum Beispiel Namensähnlichkeit, zu berücksichtigen. Aus diesem Grund werden in der vorgestellten Methode die Ergebnisse verschiedener Vergleichsalgorithmen miteinander kombiniert. Dies ermöglicht einen hohen Flexibilitätsgrad, da der Nutzer auch die Möglichkeit hat, die einzelnen Vergleichsalgorithmen manuell zu starten. Für zwei Konzepte aus verschiedenen Ontologien, können nun unterschiedliche Ähnlichkeitswerte vorliegen:

- Namensähnlichkeit:

Bei diesem Vergleich werden die Konzeptnamen syntaktisch miteinander verglichen. Dabei wurden im iMERGING-Editor zwei Methoden umgesetzt. Mit der Methode *EditDistance* (Levenshtein 1965) wird berechnet, wie viele Editieroperationen (Löschen, Einfügen eines Buchstaben) nötig sind, um einen String in den anderen umzuwandeln. Die Anzahl der Editieroperationen wird in Verhältnis gesetzt zu der Stringlänge, so dass kurze Wörter das Ergebnis nicht verfälschen können. Alternativ kann auch die Methode *N-Gramme* (Damashek 1995) verwendet werden. Diese Methode teilt die beiden Namen in Substrings der Länge  $n$  auf und ermittelt die Anzahl der Substrings der Länge  $n$  für die zu vergleichenden Wörter. Danach wird die Anzahl der *übereinstimmenden Substrings der Länge n* ermittelt und in Verhältnis gesetzt zur Gesamtzahl der Substrings mit der Länge  $n$  in den zu vergleichenden Wörtern. Z. B

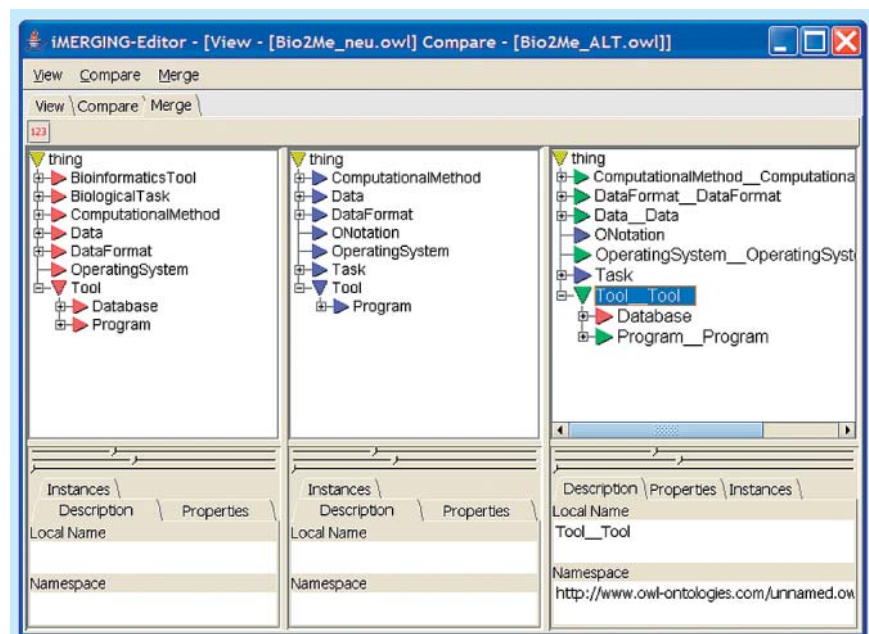


Bild 4: Zusammenführen zweier Ontologien

haben die Wörter *Daten* und *Datenbank* die Trigramme  $\{Dat, ate, ten\}$  und  $\{Dat, ate, ten, enb, nba, ban, ank\}$ . Die beiden Wörter haben nach der Methode Trigramme einen Ähnlichkeitswert von 0.55, weil sie sich zwar drei Trigramme  $\{Dat, ate, ten\}$  teilen, jedoch die Wortlänge (durch die Anzahl der Trigramme) von *Datenbank* mit ins Gewicht fällt.

- strukturelle Ähnlichkeit:

Bei diesem Vorgehen werden Beziehungen zwischen Konzepten anhand von strukturellen Informationen ermittelt. Um die Ähnlichkeit von zwei Konzepten zu bewerten, können verschiedene Nachbarschaftsbeziehungen berücksichtigt werden. Man kann die *Eltern*, *Kinder* oder auch *Eigenschaften (properties)* eines Konzepts betrachten. Die Ähnlichkeit zwischen zwei Konzepten kann z. B. von der Ähnlichkeit der Kinder (Unterknoten) abgeleitet werden. Dabei werden für die zu vergleichenden Konzepte die Unterknoten verglichen (basierend auf Namensähnlichkeiten). Sind die Unterknoten ähnlich, so wird davon ausgegangen, dass die zu vergleichenden Konzepte auch ähnlich sind, auch wenn die Namensähnlichkeit für die beiden Vergleichskonzepte nicht gegeben ist. Eine andere Methode ist, die Elternknoten bis zur Wurzel zu betrachten. Dabei werden ausgehend von den zu vergleichenden Konzepten

die Knoten auf dem Pfad bis zur Wurzel verglichen. Ist hier die Übereinstimmung hoch, kann man davon ausgehen, dass die zu vergleichenden Konzepte ähnlich sind.

- Ähnlichkeit basierend auf zusätzlichen Informationen:

Bei dieser Methode wird versucht, den semantischen Kontext von zwei Konzepten zu berücksichtigen. Oft werden Ontologien zur Annotation von Text- oder HTML-Dokumenten verwendet, um ein besseres Retrieval-Ergebnis zu erhalten. In diesem Fall betrachten wir die Dokumentensammlungen als semantischen Kontext. Die Annahme hierbei ist, wenn die Dokumente, die mit einem Konzept  $x$  (aus der Ontologie  $O_1$ ) verknüpft sind, ähnlich sind zu den Dokumenten, die mit dem Konzept  $y$  (aus  $O_2$ ) verknüpft sind, dann müssen auch  $x$  und  $y$  ähnlich sein. (Diese Methode muss noch umgesetzt werden).

In der Abschlussphase werden die individuell berechneten Ähnlichkeitswerte zu einem globalen Wert kombiniert:

$$Sim(sim_1, \dots, sim_n) = \sum_{i=1}^n w_i \cdot sim_i / \sum_{i=1}^n w_i$$

Um die einzelnen Ähnlichkeitswerte zu summieren, führen wir Gewichte  $w_i$  ein, so dass die einzelnen Ähnlichkeiten in ihrer Wichtigkeit abgestuft werden können.

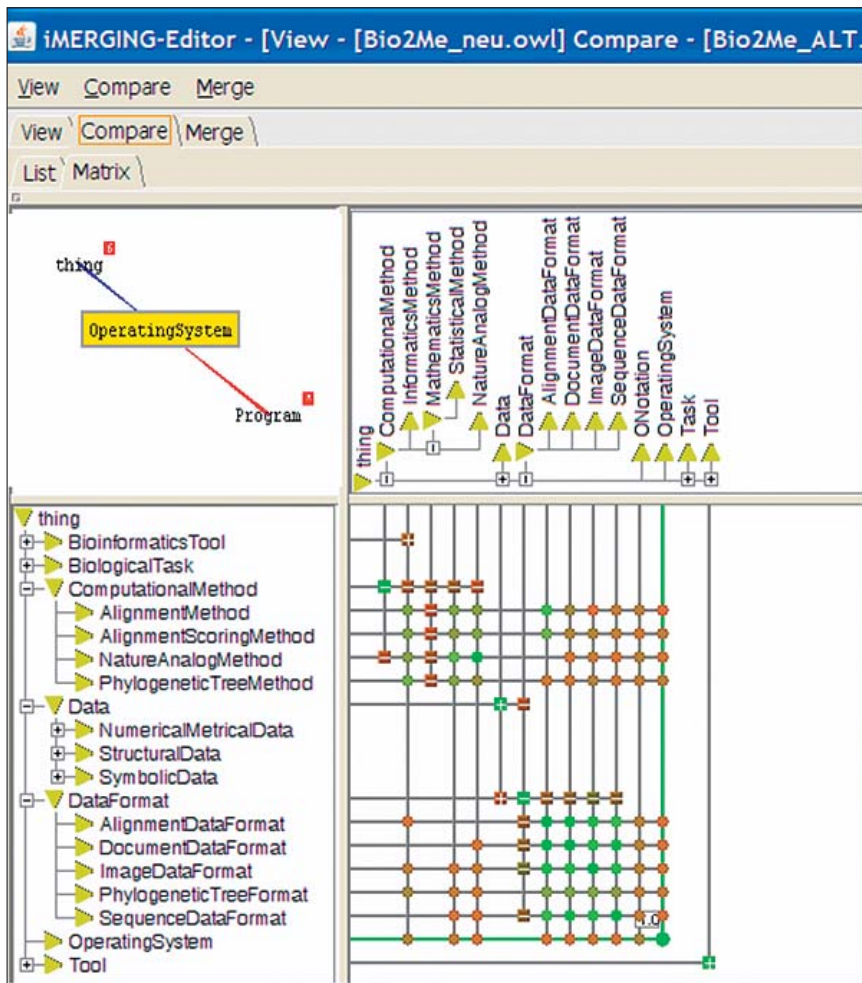


Bild 5: Matrix-Vergleich

### 3.1 Überblick durch Matrix-Vergleich

Die in Bild 2 links dargestellte Vergleichs-sicht (Listen-Sicht) ist zwar für kleine Ontologien, oder wenn die Vergleichsontologien sich nicht besonders ähnlich sind, gut geeignet, sie stößt aber auch schnell an ihre Grenzen. Die in Listen-Sicht dargestellten *Mappings* können bei sehr ähnlichen Strukturen schnell unübersichtlich werden, weil sich die Linien überkreuzen können und teilweise diagonal durchs Bild laufen, je nachdem, welche Teilbäume gerade auf- bzw. zugeklappt sind. Besonders große Wissensstrukturen erschweren hier die Visualisierung, da die Verbindungslinien in ungünstigen Fällen quer durchs Bild laufen und im nicht sichtbaren Bereich enden.

Deshalb sind auch alternative Formen der Informationsvisualisierung in Betracht zu ziehen. Ein Ansatz besteht in der Verwendung einer Matrixdarstellung für den Konzeptvergleich. Ausgangspunkt bildet das System MatrixBrowser, in dem kom-

plexe Netzstrukturen in Form einer interaktiven Matrix dargestellt und editiert werden (Ziegler und Kunz 2002).

In Bild 5 werden die zu vergleichenden Ontologien an beiden Achsen einer Matrix dargestellt, um eine systematische Übersicht über die Vergleichsergebnisse zu erhalten. Verbindungspunkte zwischen den Konzepten geben den Grad der Ähnlichkeit wieder. Durch die farbliche Unterscheidung der Verbindungspunkte (grün, gelb, rot) erhält man schnell einen Überblick über die grobe Ähnlichkeitsstruktur der Konzepte (sehr ähnlich / übereinstimmend / mittlere Übereinstimmung / keine Ähnlichkeit vorhanden). Durch Selektion einer Ähnlichkeitsrelation wird der genaue Ähnlichkeitswert angezeigt.

Durch andere Visualisierungssymbole (ein kleines Plus) in der Matrix werden Ähnlichkeiten angedeutet, die sich in zugeklappten Teilbäumen verstecken. Klickt der Nutzer das Symbol an, werden die zugehörigen Teilbäume automatisch auf-

geklappt. Der Nutzer bekommt also einen Hinweis, wo sich *ähnliche* Konzepte verstecken, ohne ihn durch aufgeklappte Strukturen durch die Vielfalt der Relationen visuell zu überfordern. Für einen ausgewählten Knoten, wird die direkte Nachbarschaft durch einen Teilgraphen visualisiert.

Nachdem sich der Nutzer einen Überblick über die zu vergleichenden ontologischen Strukturen verschafft hat, indem er sehr ähnliche bzw. sehr unterschiedliche Teilbereiche ausgemacht hat, wechselt er in die *Merging-Sicht*, um die beiden Strukturen zusammenzuführen.

### 3.2 Interaktives Zusammenführen von Wissensstrukturen

Beim Zusammenführen von zwei Ontologien  $O_1$  und  $O_2$  soll basierend auf dem vorangehenden Vergleich und den daraus resultierenden Abbildungen  $map: (O_1 \rightarrow O_2)$  eine neue Ontologie  $G$  entstehen. Das Ergebnis der Zusammenführung ist eine Ontologie, die alle überschneidungsfreien Konzepte aus  $O_1$  und  $O_2$  enthält; darüber hinaus führt sie Konzepte zusammen, die in  $map: (O_1 \rightarrow O_2)$  als redundant deklariert wurden.

Um den semi-automatischen Integrationsprozess zu unterstützen, werden die beiden Ausgangsontologien farblich unterschiedlich markiert, z. B. rot und blau (vgl. Bild 4). In die resultierende Ontologie  $G$  werden zunächst Konzepte aus der Ontologie  $O_1$  und  $O_2$  automatisch übernommen, für die es keine Abbildung in der Vergleichsontologie gibt. Konzepte für die es eine Entsprechung in der Vergleichsontologie gibt, müssen zusammengeführt werden. Dabei kann der Nutzer angeben, ab welchem Ähnlichkeitswert dies geschehen soll. Haben die beiden korrespondierenden Konzepte in den verschiedenen Ontologien unterschiedliche Unterknoten oder verschiedene Eigenschaften, werden in der resultierenden Ontologie  $G$  die Unterknoten und auch die Eigenschaften vereinigt. Ähnliche Unterknoten, mit einem Ähnlichkeitswert größer als vom Nutzer angegeben, werden rekursiv zusammengeführt. Durch die farbliche Unterscheidung kann der Nutzer jedoch den Ursprung ( $O_1$  oder  $O_2$ ) der Konzepte erkennen. Der Nutzer kann die vorgeschlagenen Zusammenführungen *annehmen*, *ablehnen* oder *ändern*, indem er die Konzepte mit der Maus selektiert. Die vom Nutzer akzep-

tierten Vorschläge und Änderungen nehmen in der Zielontologie eine andere Farbe an (z. B. grün), so dass vom Nutzer bereits bearbeitete Konzepte sichtbar werden.

Die Vereinigung von Konzepten lässt sich nicht immer auf die mengentheoretische Vereinigung zurückführen, da die Konzepte *strukturell* miteinander in Beziehung stehen. Zukünftige Arbeiten auf diesem Gebiet müssen deshalb weitergehende Vereinigungsoperationen für korrespondierende Elemente einer Ontologie definieren, um sie aus der Quellontologie in die globale Ontologie zu transformieren. Darüber hinaus kann das Ergebnis der Vereinigung zu *Konflikten* in der Ergebnisontologie führen, die im Nachgang gelöst werden müssen. Zeigt ein Konzept in der zusammengeführten Ontologie durch eine Eigenschaft (*propertie*) auf ein Konzept, das noch nicht kopiert wurde, so muss dies angezeigt werden.

## 4. Diskussion und Ausblick

Da ein vollautomatischer Vergleichs- und Integrationsprozess sich kaum realisieren lässt, ist eine benutzerfreundliche Bedienoberfläche essentiell für den praktikablen und effektiven Einsatz eines Ontologie-Editors. Diese Arbeit versucht einerseits den Vergleichsprozess soweit wie möglich automatisch durchzuführen, andererseits aber auch dem Nutzer eine *interaktive Unterstützung* zum Eingreifen in den Prozess, sowie eine *visuelle Darstellung* der vorgeschlagenen Abbildungen an die Hand zu geben.

Im Rahmen dieser Arbeit wurden bisher nur informelle Benutzertests mit drei Testpersonen durchgeführt, die erste Hinweise auf die Korrektheit der getroffenen Annahmen geben sollen (*der Vergleichsprozess muss interaktiv und visuell unterstützt werden, so dass Domänenwissen des Nutzers mit einbezogen werden kann*). Dabei wurden vor allem qualitative Ergebnisse erlangt, die erste informelle Aussagen und Überprüfungen ermöglichen. Die Probanden hatten die drei folgenden Aufgaben mit dem Ontologie-Editor durchzuführen:

- sowohl die Quell- als auch die Zielontologie explorieren; das Konzept *Pro-*

*gram* selektieren und genauer betrachten;

- das gefundene Mapping für *Program* überprüfen;
- Quell- und Zielontologie zusammenführen;

Dabei fanden alle Probanden, dass der Ontologie-Editor *iMERGING* die visuelle Exploration der beiden Ontologien unterstützt und einen schnellen Überblick über den Übereinstimmungsgrad der beiden Ontologien bietet. Außerdem konnte durch die Detailsicht im unteren Bereich der Listen-Sicht nachvollzogen werden, warum sich zwei Konzepte ähnlich sind. Auch der Merging-Prozess im Ontologie-Editor wurde als nachvollziehbar und einfach zu steuern bewertet, da man Merging-Vorschläge sofort ablehnen oder akzeptieren konnte. Jedoch würden resultierende Konflikte nicht angezeigt und die Auswirkungen des Merging-Prozesses nicht ganz deutlich.

Da es sich bei dem Ontologie-Editor um eine prototypische Implementierung handelt, wurden bisher nur informelle Benutzertests durchgeführt. Es sind jedoch weitere Usability-Tests geplant, die mit einem Eye-Tracking-System die Aufmerksamkeit des Nutzers bei der visuellen Exploration der Ontologie evaluieren sollen.

Außerdem ist geplant ein *explorierendes Merging* umzusetzen. Der Nutzer soll die Möglichkeit haben, die einzelnen Merging-Schritte *nacheinander* nachzuvollziehen. Dazu wird ihm an den richtigen Stellen der Quellontologie, die vorgeschlagenen Änderungen sind farblich hervorgehoben, angezeigt, so dass er entscheiden kann, ob sie an der Stelle „passen“ oder nicht. Der Nutzer würde so im Merging-Prozess Schritt für Schritt unterstützt werden, denn die Präsentation der fertig zusammengeführten Ontologien lässt nur sehr schwer die Auswirkungen der einzelnen Schritte erkennen.

### Literatur

- Damashek, Marc: Gauging Similarity with n-Grams: Language-Independent Categorization of Text. *Science* **267** (1995) 843–848.
- Doan, A.; Madhavan, J.; Dhamankar, R.; Domingos, P.; Halevy, A.Y.: Learning to Match Ontologies on the Semantic Web. *VLDB J.* **12** (4) (2003) 303–319.

Ehrig, M.; Staab, S.: *QOM - Quick Ontology Mapping*. International Semantic Web Conference (2004) 683–697.

Ehrig, M.: *Ontology alignment: bridging the semantic gap*. New York: Springer Verlag, 2007.

Falconer, S. M.; Noy, N. F.; Storey, M.: *Towards understanding the needs of cognitive support for ontology mapping*. International Workshop on Ontology Matching (OM 2006) at ISWC 2006, Athens, GA, 2006.

Kalfoglou, Y.; Schorlemmer, M.: *Ontology Mapping: The State of the Art*. In: *Semantic Interoperability and Integration, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI)* (Hrsg. Kalfoglou, Y.; Schorlemmer, M.; Sheth, A.; Staab, S.; Uschold, M.) Germany: Dagstuhl Seminar Proceedings, 2005.

Levenshtein V.: Binary codes capable of correcting deletions, insertions, and reversals. *Problems in Information Transmission* **1** (1965) 8–17.

Noy, N.F.; Musen, M.A.: The PROMPT Suite: Interactive Tools for Ontology Merging and Mapping. *International Journal of Human-Computer Studies* **59** (2003) 983–1024.

Noy, N.F.: Semantic Integration: A Survey Of Ontology-Based Approaches. *SIGMOD Record* **33** (4) (2004) 65–70.

Stumme, G.; Maedche, A.: FCA-MERGE: Bottom-Up Merging of Ontologies. *IJCAI* (2001) 225–234.

Ziegler, J.; Kunz, C.; Botsch, V.: Matrix Browser - Visualisierung und Exploration vernetzter Informationsräume. *Mensch & Computer 2002* (Hamburg, 2.–5. Sept. 2002) GI and German Chapter ACM.



1



2

**1 Zoulfa El Jerroudi**, Dipl.- Inform., ist wissenschaftliche Mitarbeiterin am Lehrstuhl Interaktive Systeme und Interaktionsdesign an der Universität Duisburg-Essen (Arbeitsgruppe Interaktive Systeme und Interaktionsdesign). Sie beschäftigt sich mit dem interaktiven Vergleichen (Ontologie-Mapping) und dem interaktiven Zusammenführen von Ontologien (Ontologie-Merging).  
E-Mail: eljerroudi@interactivesystems.info

**2 Prof. Dr. Ing. Jürgen Ziegler** ist seit 2003 Inhaber des Lehrstuhls Interaktive Systeme und Interaktionsdesign der Universität Duisburg-Essen. Seine Forschungsschwerpunkte liegen in den Feldern Mensch-Computer-Interaktion, Informationsvisualisierung, kontextadaptive Webanwendungen.  
E-Mail: ziegler@interactivesystems.info  
www.interactivesystems.info