

Atlas poznawczy: w stronę fundamentów wiedzy w neurokognitywistyce¹¹

Russell A. Poldrack^{*I}, Aniket Kittur^{II}, Donald Kalar^{III}, Eric Miller^{IV}, Christian Seppa^{IV}, Yolanda Gil^V, D. Stott Parker^{VI}, Fred W. Sabb^{VII} i Robert M. Bilder^{VII}

Przełożył Przemysław Nowakowski

Abstrakt

Neuronauka poznawcza dąży do odwzorowania procesów umysłowych na funkcjach mózgowych, co nasuwa pytanie o to, jakie „procesy umysłowe” istnieją i jak są powiązane z zadaniami wykorzystywanymi do manipulowania nimi i ich mierzenia. Temat ten poruszyliśmy nieformalnie we wcześniejszej pracy, ale sądzimy, że kumulatywny rozwój neuronauki poznawczej wymaga bardziej systematycznego podejścia do reprezentowania bytów umysłowych, które są odwzorowywane na funkcje mózgu oraz zadania wykorzystywane do manipulowania i mierzenia procesów umysłowych. Opisujemy nowy projekt otwartej współpracy, nazwany Atlaselem Poznawczym [Cognitive Atlas], który dąży do dostarczenia bazy wiedzy dla neuronauki poznawczej (dostępny online pod adresem: <http://www.cognitiveatlas.org>). Nakreśliśmy potencjał tego projektu jako prowadzącego do nowych odkryć dotyczących tak umysłu, jak i mózgu.

¹¹ Przekład artykułu pt. „The cognitive atlas: toward a knowledge foundation for cognitive neuroscience” opublikowanego w *Frontiers in Neuroinformatics*, September 2011, Volume 5, Article 17, source: <https://www.frontiersin.org/articles/10.3389/fninf.2011.00017/full>
DOI: 10.3389/fninf.2011.00017 Przekład opublikowano na mocy niewyłącznej licencji udzielo-
nej Frontiers Media SA przez autorów.

* Kontakt: Russell A. Poldrack, Imaging Research Center and Departments of Psychology and Neurobiology, University of Texas, 3925-B W. Braker Lane, Austin, TX 78759, USA. e-mail: poldrack@gmail.com

I Imaging Research Center and Departments of Psychology and Neurobiology, University of Texas, Austin, TX, USA

II Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA

III National Aeronautics and Space Administration, Ames Research Center, Mountain View, CA, USA

IV Interactive Design, Squishymedia, Portland, OR, USA

V Information Sciences Institute, University of Southern California, Marina Del Rey, CA, USA

VI Department of Computer Science, University of California Los Angeles, Los Angeles, CA, USA

VII Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA

Słowa kluczowe: ontologia; informatyka; neuroobrazowanie; nauki poznawcze

„Toniemy w informacjach, łąkąc wiedzy” – Rutheford B. Rogers

Neuronauka poznawcza staje przed coraz poważniejszym wyzwaniem: Jak możemy zintegrować wiedzę z gwałtownie rosnącą ilością badań o rozmaitych metodologiach, aby opisać sposób implementacji procesów umysłowych w mózgu? Stworzenie neuroobrazowych baz danych zawierających dane z wielu badań dostarczyło podstaw do skutecznych metaanaliz (Laird i inni 2005). Jednakże infrastruktura semantyczna do charakteryzowania psychologicznych aspektów tych badań została daleko w tyle za techniczną infrastrukturą do gromadzenia w bazach danych i analizowania wyników obrazowania. Sądzymy, że kumulatywny rozwój neuronauki poznawczej wymaga infrastruktury semantycznej i że problem ten należy rozwiązać przez rozwój baz wiedzy na temat procesów umysłowych (Price i Friston 2005; Bilder i inni 2009). Szkicujemy tu nowy projekt nazwany *Atlasem Poznawczym (Cognitive Atlas, CA)*¹², który ma na celu rozwinięcie takich ram przez oparte na współpracy społeczne budowanie wiedzy.

W czym tkwi problem?

Neuronaukowcy poznawczy chcą odpowiedzieć na takie pytania jak: „Jakie są neuronalne substraty pamięci roboczej?” Według bazy PubMed do lutego 2011 roku opublikowano 2613 artykułów naukowych wymieniających „pamięć roboczą” razem z funkcjonalnym rezonansem magnetycznym (fMRI), pozytonową tomografią emisyjną (PET), EEG/ERP, albo analizą lezji. Pomimo tak znaczącej liczby publikacji naukowych nadal trudno jest zintegrować te badania tak, aby zrozumieć pojęcie „pamięć robocza” i to jak się ono odnosi do funkcji mózgu. Jest tak z dwóch głównych powodów: niejednoznacznej terminologii oraz mylenia procesów poznawczych z zadaniami używanymi do ich mierzenia.

¹² Przez „poznawczy” chcemy odnosić się do procesów umysłowych bardzo szeroko, wliczając w to także, historycznie oddzielane od poznania, dziedziny jak emocje, motywacje.

Niejednoznaczna terminologia

W neuronauce poznawczej mają miejsce istotne niejednoznaczności w używaniu terminologii. Z jednej strony wiele terminów używanych jest do oznaczania rozmaitych, potencjalnie odrębnych procesów. Przykładowo termin „pamięć robocza” posiada szereg odrębnych definicji w literaturze neuronaukowej, jako:

1. utrzymywanie bieżącej informacji w pamięci, w rozumieniu Goldman-Rakic (1995), mierzone u naczelnych innych niż ludzie przy użyciu zadań takich jak opóźnienie odpowiedzi okulomotorycznej;
2. manipulowanie informacją utrzymywaną w pamięci, w rozumieniu Baddley’a (1992), mierzone u ludzi przy użyciu takich zadań jak powtarzanie ciągów literowo-liczbowych;
3. pamięć dla zmiennych czasowo aspektów zadania (ogólnie równoważna pojęciu pamięci epizodycznej), jak ujmuje ją Olton i współpracownicy (1979), mierzona u gryzoni przy użyciu zadania koncentrycznie ramiennego labiryntu ze zmieniającym się umiejscowieniem jedzenia.

Zatem wyszukiwanie „pamięci roboczej” może wyszukać artykuły istotne dla wielu różnych, specyficznych procesów psychologicznych. Z drugiej strony wiele procesów jest opisywanych w literaturze przy użyciu kilku różnych terminów. Przykładowo pierwsze z wymienionych powyżej znaczeń „pamięci roboczej” jest zazwyczaj opisywane przy użyciu innych terminów, wliczając „pamięć krótkotrwałą” czy „aktywne utrzymywanie”. Z powyższych powodów wyszukiwania, które będą zawierały wyłącznie „pamięć robocza”, nie pozwolą na wyszukanie artykułów, w których użyto innych terminów, chyba że wyszukiwanie zostanie rozszerzone tak, by zawierało pozostałe terminy. Jednak takie rozszerzenie zapytania może dostarczyć niedopuszczalnie wiele nieistotnych dokumentów.

Zadania kontra konstrukty

W pracach z neuronauki poznawczej od dawna występuje tendencja do zrównywania zadań z konstrukcjami umysłowymi. Przykładowo „zadanie rozpoznania jednostek Sternberga” jest zazwyczaj wspomniane jako „zadanie pamięci roboczej Sternberga”, co sugeruje, to że mierzy on określony konstrukt umysłowy („pamięć robocza”). To pomieszanie zadań i konstruktywów powoduje wiele trudności. Po pierwsze pomiar konstruktywów psychologicznych wymaga porównania pomiędzy określonymi warunkami zadania (Sternberg 1969), a zatem, chociaż kontrast między pewnymi warunkami w ramach zadania Sternberga (np. wysokie obciążenie kontra

niskie obciążenie) może w istocie być powiązany z konstruktem pamięci roboczej, inny może nie być z nim połączony (np. dopasowanie próby kontra niedopasowanie próby). Po drugie każde powiązanie między zadaniami i konstrukcjami odzwierciedla szczególnie teorię tego, jak zadanie jest wykonywane, a zatem, zrównując zadania z konstrukcjami, czynimy założenia teoretyczne, które mogą nie być podzielane przez wspólnotę (a co więcej, założenia podzielane przez wspólnotę mogą być niepoprawne). Przykładowo zadanie kolor-słowo Stroopa jest czasami wspomniane jako „zadanie hamowania” (Donohoe i inni 2006). Jednak rola procesu aktywnego hamowania w wytwarzaniu efektu Stroopa, została podważona przez wielu badaczy (Cohen i inni 1990). Podobnie chociaż zadanie *N*-wstecz jest zazwyczaj przywoływane jako „*N*-wstecz zadanie pamięci roboczej”, przedstawiono poważne wątpliwości, czy faktycznie mierzy on konstrukt „pamięć robocza” (Kane i inni 2007). Zrównanie zadań i procesów tworzy poważne zamieszanie dotyczące tego, co jest właściwie mierzone w badaniach neurokognitywistycznych.

Pojawia się tu następujący problem: pojedyncze zadanie często łączone jest w literaturze z wielorakimi konstrukcjami. Na przykład Sabb i współpracownicy (2008) użyli narzędzi do eksploracji literatury, by zbadać publikacje związane z konstruktem „kontrola poznawcza”. Odkryli oni, że konstrukt ten połączony jest z wieloma innymi konstrukcjami (wliczając w to: „pamięć robocza”, „hamowanie reakcji”, „selekcja reakcji” i „zmiana zadanie/zestaw”), a także to, że nie było w literaturze ani jednego zadania, które byłoby połączone wyłącznie z konstruktem „kontrola poznawcza”, każde z zadań połączonych było jeszcze z co najmniej jednym, odmiennym konstruktem. Ponadto znajdujące się w literaturze połączenia pomiędzy zadaniami a konstrukcjami zmieniały się z biegiem czasu. Ten brak spójności w piśmiennictwie w traktowaniu zadań i pojęć utrudnia wyciąganie sensownych wniosków z istniejącej literatury i ogranicza łączną wartość reprezentowanej w niej wiedzy.

Ku ontologii dla poznania

Tym, czego pilnie potrzebujemy, są zasoby informatyczne, które rozwiązałyby wymienione wyżej problemy. Takie zasoby dostarczałyby zdolności do identyfikowania konkretnych użyc terminów, aby umożliwić przeglądanie powiązanych pojęć oraz identyfikację istotnych świadectw w literaturze powiązanej z tymi pojęciami. To umożliwiłoby inteligentne gromadzenie wyników badań, które może przyczynić się do przewyciężenia przeciążania informacyjnego, które dotyka badaczy. Proponujemy, że temu wyzwaniu najprościej będzie sprostać dzięki rozwojowi i

powszechnej implementacji ontologii w neuronauce poznawczej. W filozofii „ontologia” odnosi się do badania istnienia czy bytu. Jednak w bioinformatyce termin ten jest w coraz większym stopniu używany w znaczeniu zdefiniowanym przez Grubera (1993) jako „jawna specyfikacja konceptualizacji” lub jako ustrukturyzowana baza wiedzy przeznaczona do wspierania dzielenia się wiedzą jak również zautomatyzowanych wnioskowań o tej wiedzy. Ontologie dostarczają również podstaw dla efektywnego gromadzenia wiedzy w biologii molekularnej i genomice (Bard, Rhee 2004). Jednym z najbardziej znanych przykładów jest Gene Ontology¹³ (GO; Ashburner i inni 2000). Ontologia ta dostarcza spójnych deskryptorów dla wytworów genów, wliczając składniki komórkowe (np. „rybosom”), procesy biologiczne (np. „przewodzenie sygnału”) i funkcje molekularne (np. „aktywność katalityczną”). GO stanowi podstawę, na której adnotuje się zbiory danych w zależności od ich funkcji, co nie dopuszcza do dość powszechnego, obecnego u wielu badaczy, problemu używania różnych nazw na tę samą strukturę czy proces biologiczny występujący u różnych organizmów. Zapewnia ona również możliwość przemierzania ontologii, by odkrywać regularności na większą skalę przez rozszerzanie wyszukiwania tak, aby włączało do ontologii terminy podrzędne. Istnieją coraz skuteczniejsze narzędzia budowane wokół takich ontologii jak GO. Mając zbiór danych (taki jak wzorzec ekspresji genu), narzędzia te dostarczają szeroki zakres funkcji takich jak porównywanie zbiorów danych genetycznych na podstawie podobieństwa wzorców anotacji w GO (Ruths i inni 2009) i wydobywanie nowych faktów biologicznych z tekstu artykułów (Muller i inni 2004). Ontologie były także rozwijane w wielu innych obszarach neuronauk (Martone i inni 2004); najbardziej istotne dla neuronauki poznawczej są dobrze rozwinięte ontologie struktury mózgu (Bowden, Dubach 2003)

Duża część badań w naukach poznawczych rozwinęła szczegółowe dziedzinowo-specyficzne teorie procesów umysłowych, ale bardzo mało jest prac, które systematycznie charakteryzują, jak te procesy są zdefiniowane i jak wspólnie pasują do większej struktury. Najprawdopodobniej częściowo odzwierciedla to funkcjonalistyczny charakter współczesnej psychologii, która powstała jako reakcja na dziewiętnastowieczne podejścia strukturalistyczne (np. takie, jakie widzimy w tzw. „psychologii władz poznawczych” wykorzystywanej przez frenologów; Boring, 1950). Były pewne próby formułowania zakrojonych na szeroką skalę „zunifikowanych teorii poznania”, takie jak ACT-R Andersona (Anderson i inni 2004) i SOAR Newella (Laird i inni 1987), ale te podejścia były pierwotnie skupione raczej na

¹³ <http://www.geneontology.org>

rozwijaniu ogólnych, unifikujących zasad obliczeniowych, a nie na systematycznej charakterystyce szerokiego zakresu procesów poznawczych.

Inne rozległe słowniki, jak Medical Subject Headings (MeSH), dostarczają pewnych treści istotnych dla neuronauki poznawczej, lecz są bardzo ograniczone. Przykładowo hierarchia dla „poznania” w MeSH zawiera wyłącznie następujące pojęcia: przytomność (Awareness), dysonans poznawczy (Cognitive Dissonance), rozumienie (Comprehension), świadomość (Consciousness), wyobraźnia (Imagination) i intuicja (Intuition). Terminy te nie mają żadnego znaczącego związku z aktualnymi ramami pojęciowymi nauk poznawczych. Dodatkowo terminy z MeSH są mieszaniną procesów umysłowych (np. „rozumienie”), zjawisk doświadczalnych (np. „iluzje”) i procedur doświadczalnych (np. „uczenie się labiryntu”) wraz z takimi przestarzałymi terminami jak „programowanie neurolingwistyczne” (które w najlepszym razie można uznać za pseudonaukę). W związku z tym, że MeSH jest słownikiem używanym do indeksowania artykułów i rozszerzania zapytań w PubMedie, wydaje się, że przeszukiwanie tej literatury może zostać znacznie udoskonalone dzięki użyciu słowników, które lepiej odzwierciedlają aktualne myślenie.

Rozwój formalnych ontologii poznania mierzy się z wyzwaniami odmiennymi niż inne obszary biologii jak neuroanatomia czy funkcje komórkowe: w tej dziedzinie istnieje znikomy konsensus w kwestii podstawowych jednostek funkcji umysłowych. W tym przypadku mamy dwa rozwiązania. Pierwsze byłoby pójściem do przodu i rozwinięciem pojedynczej ontologii opartej na zgodzie osiągniętej między niewielką grupą jednostek. Zaletą tego rozwiązania byłoby dostarczenie ontologii zatwierdzonej dzięki zgodzie jej projektantów, ale byłaby ona bezużyteczna dla każdego, kto nie podziela zobowiązań ontologicznych niniejszej grupy. Alternatywne podejście, którego się trzymamy w tej pracy, to pozwolenie na konflikty i ich wychwytywanie, aby reprezentować pełen zakres ujęć obecnych tej dziedzinie. Nasze podejście do tego zagadnienia jest inspirowane sukcesem opartych na współpracy społecznych projektów budowania wiedzy, jak Wikipedia, które umożliwiają dyskusję i wyrażanie rozbieżnych poglądów w celu rozwijania szerszego konsensusu, a także na możliwości elastycznej modyfikacji podczas wyłaniania się nowej wiedzy.

Atlas poznawczy

W celu zaspokojenia potrzeby formalnej podstawy dla wiedzy, która objęłaby szeroki zakres struktur pojęciowych w naukach poznawczych, rozwinęliśmy CA (dostępny online, pod adresem: <http://www.cognitiveat->

las.org). System ten jest nieustannie rozwijany, a nowe funkcje będą dodawane w przyszłości. Bieżący system oferuje podstawowe funkcje do specyfikacji wiedzy na temat procesów i zadań poznawczych. System ten został zaprojektowany z zamiarem uczynienia interakcji z bazą wiedzy tak prostą, jak to możliwe, bez wymagania, by użytkownik miał doświadczenie w ontologii czy rozwijaniu baz wiedzy (Miller i inni 2010). Dodatkowo w celu umożliwienia programowego dostępu do bazy danych system używa standardowych mechanizmów (takich jak: protokół SPARQL, język zapytań RDF, SPARQL), które pozwalają innym stronom czy bazom danych automatycznie wykorzystywać jego zawartość.

Ważną, przewodnią zasadą przy projektowaniu CA był rozróżnienie między zadaniami a procesami umysłowymi. Procesy umysłowe nie są dostępne bezpośrednio, ale do ich mierzenia i manipulowania nimi można użyć zadań psychologicznych, a zachowanie lub aktywność mózgu obserwowana w trakcie tych zadań interpretować jako odzwierciedlającą ukryte konstrukty umysłowe. Status ontologiczny zadań nie podlega wątpliwości (tj. niemal każdy zgodzi się co do tego, czym jest „test Stroopa”), ale relacja pomiędzy tymi zadaniami a ukrytymi konstruktami umysłowymi znajduje się w centrum wielu debat w naukach poznawczych. Z tego powodu sugerujemy, że istotne jest dokonanie wyraźnych rozróżnień pomiędzy procesami umysłowymi a zadaniami psychologicznymi oraz rozwinięcie odrębnych ontologii dla tych dwóch dziedzin (otrzymując dwie odrębne, ale wzajemnie połączone ontologie, które stanowią graf dwudzielny).

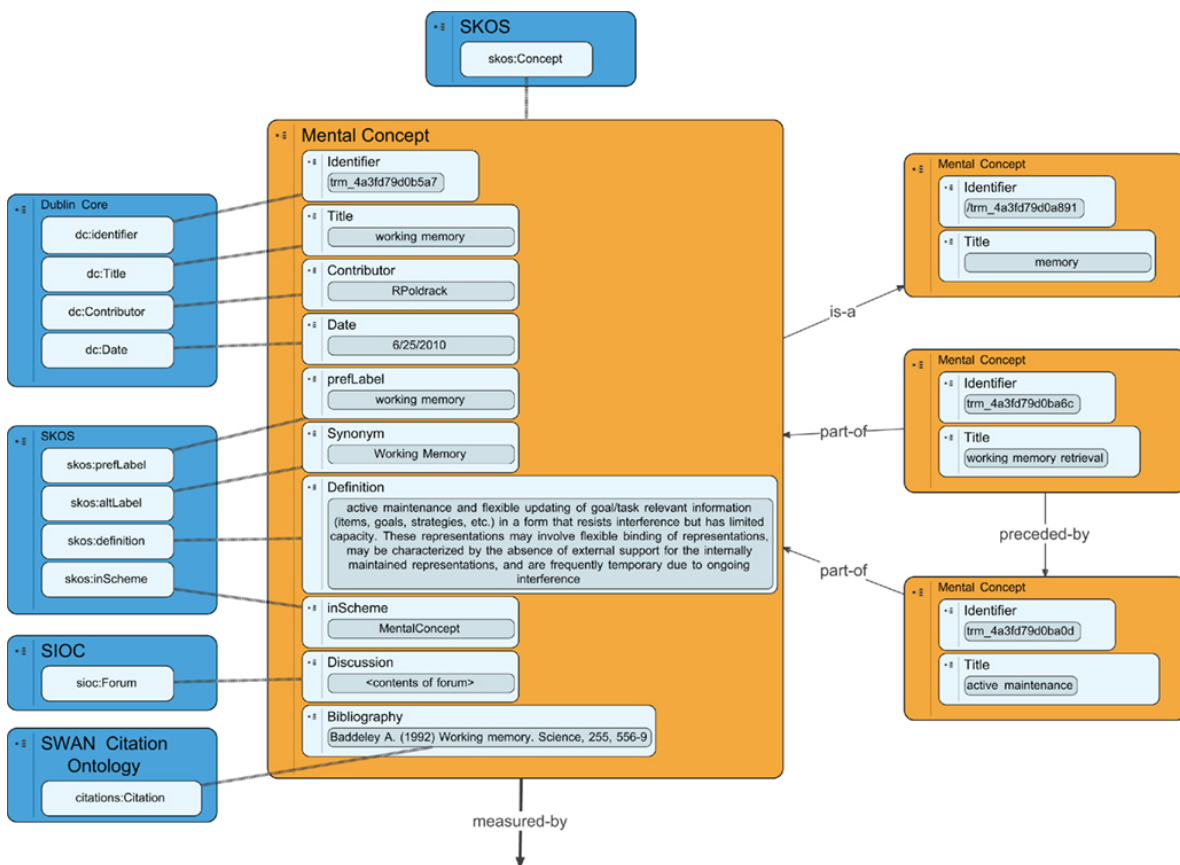
Struktura bazy wiedzy

Rozwój schematów dla bazy wiedzy CA wymaga analizy rodzajów struktur wiedzy, które są używane w naukach poznawczych. Wstępny słownik zawierający ponad 800 terminów został zidentyfikowany manualnie, dzięki analizie dużego zbioru publikacji z nauk poznawczych i neuronauki poznawczej (jego twórcami są trzej autorzy: Russell A. Poldrack, Robert M. Bilder, Fred W. Sabb). Były te zostały podzielone na dwie rozległe klasy: *pojęcia umysłowe* i *zadania umysłowe*.

Pojęcia umysłowe

Pojęcia umysłowe to ukryte, nieobserwowalne konstrukty postulowane w teoriach psychologicznych. Choć pojęcia te ostatecznie realizowane są przez tkankę mózgową, to byt pojęcia umysłowego w bazie wiedzy odnosi się do niejawnych konstruktów (np. na poziomie obliczeniowym lub al-

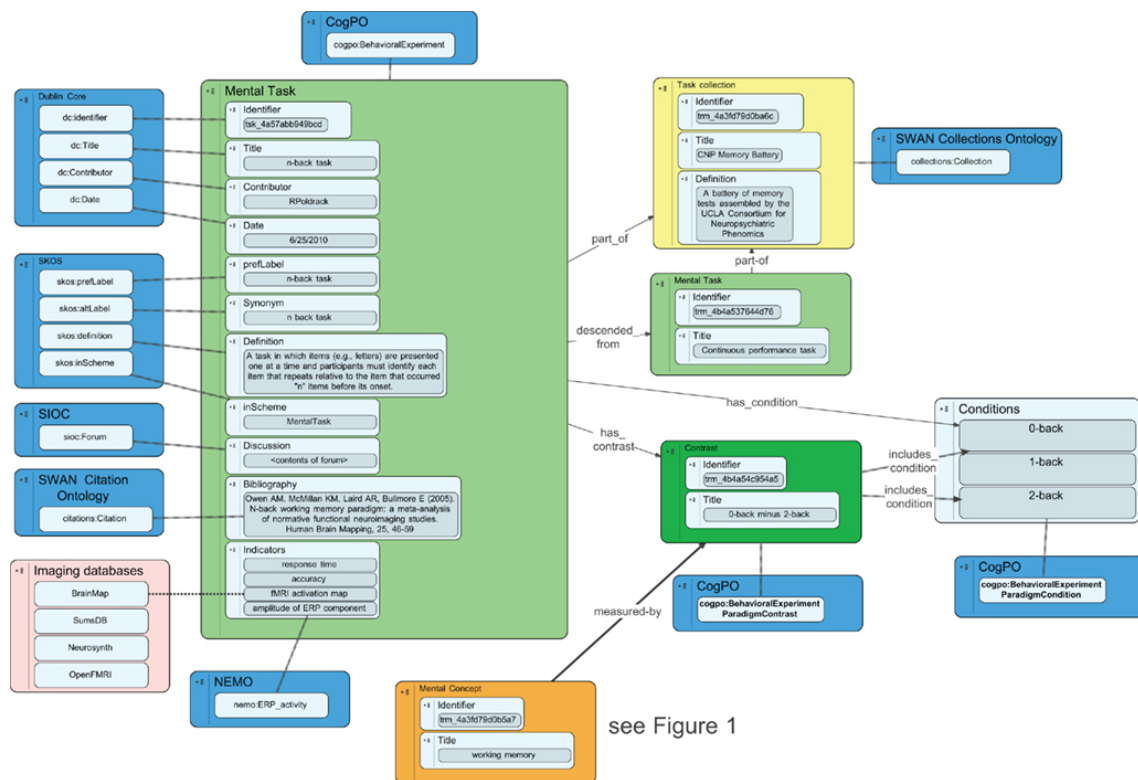
gorytmicznym u Marra), a nie do ich fizycznej realizacji. Niektóre potencjalne pojęcia umysłowe zawierają (ale nie są ograniczone do) umysłowych procesów i reprezentacji. *Reprezentacje umysłowe* są bytami umysłowymi, które są powiązane z pewnymi bytami fizycznymi (np. umysłowy obraz sceny wzrokowej jest powiązany lub jest izomorficzny względem pewnego układu przedmiotów w świecie fizycznym). *Procesy umysłowe* są bytami, które operują na reprezentacjach umysłowych lub przekształcają je (np. procesy, które szukają konkretnego przedmiotu w umysłowej reprezentacji sceny wzrokowej). W celu dostosowania możliwie jak najszerszej grupy teorii poznania (również niereprezentacyjnych jak teoria Edelmanna, 1989), baza wiedzy nie wymaga, by pojęcia umysłowe określano tymi podklasami, i w istocie jest ona neutralna w kwestii powyższego rozróżnienia, nie wymagając, ale pozwalając na to, by „reprezentacje” poznawcze istniały. Pojęcia umysłowe w bazie wiedzy CA modelowane są przy użyciu klasy *Pojęcie*, pochodzącej z prostego systemu organizacji wiedzy (*simple knowledge organization system*, SKOS; Bechhofer, Miles 2009), który opisuje podstawową strukturę bytów pojęciowych. Ogólny zarys schematu bazy danych dla pojęć umysłowych jest prezentowanych na **Ilustracji 1**.



Ilustracja 1. Przegląd schematu bazy danych do reprezentowania pojęć umysłowych w CA.¹⁴ Błękitne boksy odzwierciedlają ontologie zewnętrzne, linie przerywane klasy dziedziczenia, podczas gdy linie ciągłe relacje ontologiczne.

Zadania umysłowe

Zadaniem umysłowym jest określona aktywność mająca na celu zaangażowanie lub manipulowanie funkcją umysłową w celu uzyskania wglądu w znajdujące się u jej podstaw procesy umysłowe. Struktura reprezentacji zadań umysłowych w CA opiera się na Cognitive Paradigm Ontology (CogPO¹⁵; Turner, Laird 2012), ontologii zawierającej podstawowe klasy Behavioral Experimental Paradigm, opisującego zadania umysłowe. Ogólny zarys schematu dla bazy danych dla zadań umysłowych przedstawiono na **Ilustracji 2**.



see Figure 1

¹⁴Treść definicji zamieszczonej na schemacie: „Definicja: aktywne utrzymywanie i elastyczna aktualizacja informacji ważnych dla celu/zadania (jednostek, celów, strategii itd.), w postaci, która ma ograniczone zdolności, ale jest odporna na interferencje. Reprezentacje te mogą wymagać łączenia reprezentacji, charakteryzować się brakiem zewnętrznego wsparcia dla reprezentacji utrzymywanych wewnętrznie, a ze względu na nieustanne interferencje często są tymczasowe” (przyp. Tłum).

¹⁵www.cogpo.org

Ilustracja 2. Przegląd schematu bazy danych do reprezentowania zadań umysłowych w CA. Błękitne boksy odzwierciedlają ontologie zewnętrzne, linie przerywane klasy dziedziczenia, podczas gdy linie ciągłe relacje ontologiczne. Linia przerywana łączy mapy aktywacji z bazami danych z obrazowania ma odzwierciedlić relację empiryczną, ponieważ, aktualnie te bazy danych aktualnie nie ujawniają ontologii formalnych.

Zadania często, ale nie zawsze, wywołują jawne reakcje (jak działania ruchowe), np. badania oparte na neuroobrazowaniu mogą mierzyć reakcje neuronalne wywołane przez poszczególne formy stymulacji (takie jak oglądanie filmu) bez jakiegokolwiek jawnego zachowania. Każde poszczególne zadanie ma wiele różnych cech, które muszą być odróżnione i/lub mierzone.

Warunki doświadczalne

Warunki doświadczalne stanowią podzbiory eksperymentu określające istotne manipulacje eksperymentalne. Przykładowo w zadaniu słowo-kolor Stroopa, występują na ogół trzy warunki (próby: zgodne, niezgodne i neutralne). Jednak mogą być również tak rozszerzone, by włączyć manipulacje parametrami, które są określone zgodnie z klasą Behavioral Experimental Paradigm Conditions w CogPO.

Wskaźniki

Wskaźniki to specyficzna ilościowa bądź jakościowa zmienna, która jest rejestrowana dla celów analizy. Mogą to być zmienne behawioralne (takie jak czas czy dokładność reakcji, czy inne miary wykonania zadania) lub zmienne fizjologiczne (w tym genetyka, psychofizjologia, efekty lezji lub dane z neuroobrazowania). W bieżącej implementacji CA koncentrujemy się głównie na wskaźnikach behawioralnych, ale zmierzamy do tego, by system można było stosować do wszelkich wskaźników mierzonych w kontekście funkcji umysłowych, włączając w to pomiary fizjologiczne, genetykę i dane z obrazowania.

Kontrasty

Choć bezwzględne miary zachowania czasami mogą być znaczące (np. wyniki w standaryzowanych testach), to zazwyczaj łączymy porównanie wskaźników pomiędzy różnymi warunkami doświadczanymi z poszczególnymi procesami umysłowymi, przy pomocy logiki subtrakcji lub innego układu doświadczalnego. W CA kontrast definiujemy jako wszelką funkcję

na warunkach doświadczalnych (zapożyczamy tę definicję z pojęcia kontrastów liniowych w ogólnym modelu liniowym, zgodnie z powszechnym użyciem w literaturze dotyczącej neuroobrazowania). Najprostszym kontrastem jest wartość wskaźnika konkretnego warunku. W pewnych przypadkach będzie to miało znaczenie w wartościach bezwzględnych, podczas gdy w pozostałych (np. z danymi neuroobrazowania) będzie odzwierciedlało porównanie między prostszymi warunkami podstawowymi, takimi jak spoczynek i fiksacje wzrokowe. Bardziej złożone kontrasty zawierają liniowe lub nieliniowe funkcje wskaźników w różnych warunkach eksperymentalnych. Przykładowo, w zadaniu Stroopa, oprócz takich kontrastów, jak (niespójne-spójne) i (niespójne-neutralne) wskazujących dobrze znany „efekt Stroopa”, będą to dodatkowe proste kontrasty dla każdego z trzech warunków. Są one definiowane zgodnie z klasą Behavioral Experimental Paradigm Contrasts w CogPO.

Relacje

Chociaż dobrze zdefiniowany słownik ma zasadnicze znaczenie dla naszej bazy wiedzy, to najbardziej interesują nas relacje pomiędzy terminami ze słownika, ponieważ wyrażają założenia teoretyczne koncepcji poznawczych. CA zawiera różne, liczne rodzaje relacji (choć nie wszystkie zostały zaimplementowane w aktualnej wersji). Gdzie było to tylko możliwe, ponownie używaliśmy już istniejących ontologii.

Podstawowe relacje ontologiczne

Standardowy zbiór relacji ontologicznych został skodyfikowany w Open Biological Ontologies (OBO) Relational Ontology (Smith i inni 2005) zapewniającej wytyczne dotyczące spójnego użycia konkretnych ontologicznych relacji w różnych ontologiach. Przyjęliśmy kilka podstawowych relacji z tych ontologii:

- *jest (pewnym)* (np. „pamięć deklaratywna jest rodzajem pamięci”)
- *jest częścią* (np. „przywoływanie z pamięci jest częścią pamięci deklaratywnej”)
- *jest przekształcaniem* (np. „pamięć skonsolidowana jest przekształceniem pamięci zakodowanej”)
- *poprzedzony przez* (np. „konsolidacja pamięci jest poprzedzona przez kodowanie pamięci”)

Wykluczaliśmy zdefiniowanych w OBO Relational Ontology zbiór relacji przestrzennych, ponieważ dotyczą one wyłącznie bytów, które mają lokalizację przestrzenną. Wykluczaliśmy również zbiór relacji uczestniczenia (ma_uczestnika, ma_sprawcę), operatory te wyrażają relacje pomiędzy procesami i kontynuantami (rzeczami), a w kontekście niejawnych bytów umysłowych, nie jest jasne, jak te konstrukty należy konceptualizować.

Relacje pomiędzy procesami a zadaniami

Oprócz podstawowych relacji zawartych w OBO Relational Ontology definiujemy również relację *mierzony przez*, która oznacza relację pomiędzy procesem poznawczym a określonym kontrastem z zadania (np. „przetwarzanie niezgodności w zadaniu Stroopa mierzone przez kontrast [spójne-niespójne]”). Celem tego jest odzwierciedlanie pierwotnego kształtu twierdzeń teoretycznych formułowanych przez psychologów poznawczych, mianowicie tego, że pewne manipulacje w zadaniu wpływają na określony proces umysłowy.

Relacje między zadaniami

Kolejną nową relacją wprowadzaną w CA jest relacja *pochodzi od*, która reprezentuje historyczne i/lub pojęciowe relacje pomiędzy zadaniami. Jasne jest, że w szerokim wyborze zadań (np. „zadanie kolor-słowo Stroopa”) będzie wiele potencjalnych zmian, które mogą mieć implikacje funkcjonalne, a te rozwijają się z czasem. By uchwycić te relacje, używamy pojęcia „filogeneza zadań” (Bilder i inni 2009), która traktuje te zadania zgodnie z drzewem rodzinnym, na którym dziedziczą one poszczególne cechy od zadań wcześniejszych. Dlatego relacja *pochodzi od* odzwierciedla coś w rodzaju biologicznej relacji dziedziczenia. W tym przypadku „specjacja” jest określana przez to, czy uzyskane dane są współmierne dla metanalizy. Jeżeli nie są, to jedno z zadań zostanie uznane jako wywodzące się z innego zadania, a nie jako nieco inna odmiana tego samego zadania.

Literatura dotycząca relacji

Wszystkie byty w bazie wiedzy CA (wliczając pojęcia, zadania i relacje) mogą być połączone z cytowaniami literatury przy użyciu interfejsów wbudowanych w bazę literatury PubMed. Baza wiedzy pozwala również na anotację relacji do poszczególnych cytowań, przy użyciu relacji zdefiniowanych w Citation Typing Ontology (CiTO; Shotton, 2010). Ontologia ta wspiera anotację, która określa, czy poszczególne cytowanie wspiera lub

obala określone twierdzenie, jak również wiele innych aspektów cytowań. Aktualnie mamy zaimplementowaną jedynie jedną relację literaturową, która jest równa relacji z CiTO: „cytujeDlaInformacji”.

Relacje do danych obserwacyjnych

CA bezpośrednio nie przechowuje danych; zamiast tego, by wspierać anotację relacji pomiędzy zadaniami i zaobserwowanymi danymi, system daje możliwość powiązania kontrastów dla poszczególnych zadań z bytami lub danymi, które są przechowywane w zewnętrznych bazach danych. Obejmuje to obszary mózgu (reprezentowane w bazach danych struktur mózgu, takich jak Foundational Model of Anatomy lub Brainmap), geny i warianty genetyczne (reprezentowane w dbSNP i EntrezGene) oraz funkcje komórkowe (reprezentowane w GO).

Infrastruktura techniczna

Baza wiedzy CA jest natywnie przechowywana w standardowej relacyjnej bazie danych MySQL. Zdecydowaliśmy się na to, a nie na natywne przechowywanie wiedzy w języku ontologii lub Resource Description Format (RDF) triplestore, by zmaksymalizować elastyczność przechowywania wiedzy. Zamiast tego rozwinęliśmy potok do generowania ontologii Web Ontology Language (OWL) z bazy danych. Pozwala to nam ujawnić wiedzę ontologiczną w standardowym formacie, utrzymując elastyczność przechowywania informacji, które nie mogą być łatwo reprezentowane w formacie języka ontologii. OWL'owa reprezentacja CA jest dostępna przez NCBO BioPortal¹⁶, a kod w Pythonie użyty do generowania reprezentacji OWL ze zrzutów z bazy danych jest dostępny pod adresem <https://github.com/poldrack/cogat>.

W celu maksymalizacji zdolności do bezpośredniej i automatycznej interakcji z innymi projektami, CA jest zbudowany wokół zbioru technologii (Miller i inni 2010) Sieci Semantycznej (Berners-Lee i inni 2001). Po pierwsze, reprezentacja każdego pojęcia w bazie wiedzy jest dostępna w RDF, który jest formatem reprezentującym zasoby semantyczne w sposób zrozumiały dla maszyn. Po drugie, strona CA udostępnia usługę sieciową znaną jako SPARQL endpoint, która umożliwia bezpośrednie zapytania bazy wiedzy przez ludzi lub inne systemy komputerowe i oddaje z bazy wiedzy wyniki w formacie opartym na standardach, który zachowują relacje pojęciowe i

¹⁶ <http://bioportal.bioontology.org/>

wartościowe informacje kontekstowe. Łącznie usługi te dostarczają innym projektom możliwości bezpośredniego i bardziej wydajnego dostępu do aktualnego stanu bazy wiedzy. Umożliwia to znacznie lepszą interoperacyjność pomiędzy systemami reprezentującymi różne rodzaje informacji i wspiera automatyzację interakcji opartej na wspólnych standardach interoperacyjności i dzielenia wiedzy.

Poza konkretnymi zyskami dla neuronauki poznawczej infrastruktura rozwinięta jako części CA powinna również stanowić budulec dla innych projektów, których celem jest tworzenie bazy wiedzy opartych na współpracy. Biorąc pod uwagę, że wspólnota informatyków ciągle dąży do standardów dla interoperacyjności między projektami, mamy nadzieję, że nasza demonstracja wydajności technologii Sieci Semantycznej dostarczy dodatkowego bodźca dla jej używania w takich projektach.

Interfejsy oparte na sieci wykorzystują HTML 5, CSS 3, i standardowego JavaScript, by generować właściwości interaktywne. Opiera się on na bibliotekach jQuery i bardzo wielu wtyczkach usług jQuery. Serwerowa część oprogramowania używa standardowych stosów LAMP (Linux, Apache, MySQL i PHP). Graphviz jest używany do generowania wizualizacji RDF, a Arc2/SemSol do parsowania bibliotek RDF i SPARQL. Klient i serwer komunikują się przy pomocy technik AJAX, przekazując obiekty JSON ze strony użytkownika końcowego do serwera i z powrotem, jak również za pomocą standardowych zapytań POST i GET. Wsparcie dla PDF jest dostarczane przez biblioteki wkhtmltopdf. Infamous zasila bąbelkowe wizualizacje RSS przez Feedburnera, który generuje kanały RSS, podczas gdy PubBrain zasila odniesienia do obrazowania fMRI. PubMed Entrez API jest używany do cytowań bibliograficznych i przeszukiwania abstraktów.

Zyski i wyzwania wspólnego budowania wiedzy

Internet umożliwił wykorzystywanie wiedzy ludzi z całego globu na niespotykaną wcześniej skalę. Setki tysięcy ludzi pracowało razem nad pisanie oprogramowania, które uruchamia Internet, pisze największą w historii encyklopedię, odkrywa nowe gwiazdy i galaktyki, realizuje wiele innych celów niemożliwych do osiągnięcia przez ludzi czy komputery działające w pojedynkę. Takie systemy stanowią zarówno istniejące dowody, jak i modele projektowe kolaboracyjnego budowania wiedzy w nauce. Przykładowo, tysiące jednostek, każda dysponująca inną wiedzą i punktem widzenia, mogą przyczynić się do powstania jednego hasła Wikipedii, z wynikami rywalizującymi jakością z hasłami pisanymi przez ekspertów (Giles 2005), a znacznie przewyższającymi je co do zakresu. Jednak proces, dzięki któremu pojawia się taka wysokiej jakości wiedza oparta na

współpracy, w żadnym wypadku nie jest zagwarantowana, wymaga rozwoju norm wspólnoty, jawnych reguł, technologii i znacznych wysiłków w kierunku koordynacji i rozwiązywania konfliktów (Kittur, Kraut 2008). Naszym celem w projekcie CA jest wykorzystanie wyłaniającego się rozumienia opartego na wspólnocie kolaboracyjnego budowania wiedzy, by rozwijać mający służyć naukowcom system, którzy wychwytuje wiele różnorodnych punktów widzenia i buduje konsensus między obszarami.

Wychwytywanie i rozwiązywanie sporów

Choć na pierwszy rzut oka prosta, Wikipedia jest wyrafinowaną maszyną do kolaboracyjnego budowania wiedzy, z bardzo rozwiniętym mechanizmem obniżania kosztów uczestnictwa promującym współpracę, aktualizację informacji, rozwiązywanie konfliktów i konsolidację treści. Ponad jedna trzecia pracy w Wikipedii nie jest przeznaczona na edytowanie treści, ale na koordynowanie aktywności takich, jak debaty na temat zasad i procedur, utrzymywanie aktywności takich jak usuwanie nieodpowiednich stron oraz negocjacje dotyczące treści i problemów (Kittur et al., 2007). Bardzo duży i zwiększający się nakład pracy na koordynację akcentuje zapotrzebowanie na systemy do kolaboracyjnego tworzenia wiedzy i skupienie się na wspieraniu współpracy przynajmniej w takim samym stopniu, jak tworzenie wiedzy. Wikipedia wspiera współpracę dzięki wielu różnym mechanizmom. Jednym z najprostszych, ale i najważniejszych, jest obecność na każdej stronie (oddzielonej od jej treści) sekcji „Dyskusja”, zawierającej zapis wszystkich aktualnych i minionych konwersacji na temat strony. Nowi użytkownicy mogą zobaczyć poprzednie dyskusje, by skorzystać z ich zalet, ale i uniknąć powtarzania wcześniejszych błędów. CA wykorzystuje tę funkcjonalność, aby poradzić sobie z relacjami, rejestrując dyskusje nie tylko dla pojęć, ale także dla relacji między nimi. Ponadto, dyskusje są bezpośrednio zintegrowane ze stronami pojęć i relacji, wyświetlając je i czyniąc je istotnymi zarówno dla czytelników, jak i dla twórców. Kiedy sama dyskusja nie może rozwiązać różnicy zdań, to w trakcie jej rozwoju pojęcia mogą być albo „rozwidlane”, albo łączone.

Nadzór

Chociaż całkowicie otwarte systemy, takie jak Wikipedia, sprawdzają się dla ogólnej wiedzy encyklopedycznej, niektóre z największych sukcesów w naukowym budowaniu wiedzy pochodzą z modeli nadzorowanych takich jak GO. Celem CA jest próba pogodzenia między tymi skrajnościami, z nad-

zorem sprawowanym w razie potrzeby przez główny zespół oraz kuratorów-wolontariuszy. Jednak nadzór nigdy nie jest sprawowany w próżni, a punkty widzenia i dyskusje wspólnoty będą przywoływane przy podejmowaniu decyzji związanych z nadzorem, i oczekuje się, że decyzje te będą odzwierciedlać konsensus wspólnoty. Jednocześnie przewodnią zasadą jest, że decyzje kuratorów powinny być „oparte na dowodach”, by zapewnić, że wątpliwości dotyczące terminologii, pojęć i relacji nie są determinowane ze względu na ich popularność we wspólnocie. W miarę powiększania się Atlas będzie prawdopodobnie potrzebował bardziej wyrafinowanych procedur, by radzić sobie z nadzorem (jak w Wikipedii: Forte i inni 2009), które będą rozwijane we współpracy z i przez wspólnotę użytkowników Atlas naukowców.

Użyteczność atlasu poznawczego

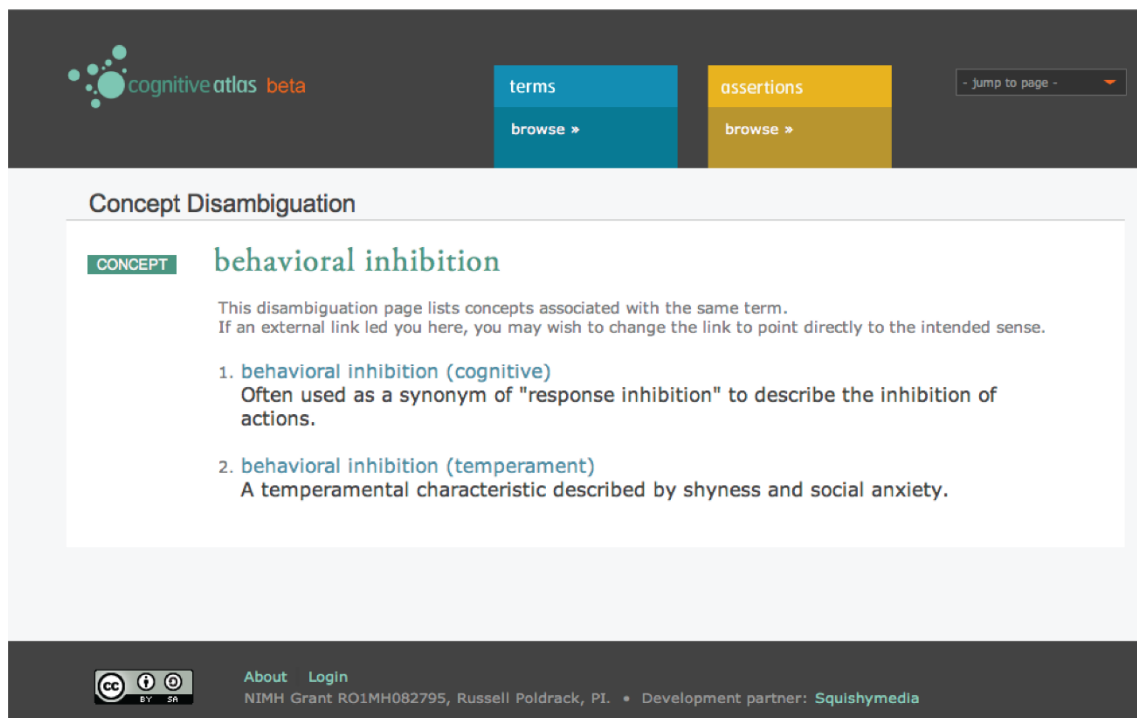
CA dostarczy formalizacji tego, czym w neuronauce poznawczej są najczęściej niejawne schematy pojęciowe, a w szczególności pokaże jasno odwzorowywanie kontrastów poszczególnych zadań nad konkretne procesy umysłowe. Przewidujemy, że CA może na wiele sposobów wpłynąć na badania w neuronauce poznawczej.

Jaśniejsza terminologia

Kontrolowany słownik CA dostarcza badaczom sposobu na precyzyjniejsze użycie terminów i pomaga zredukować wieloznaczność, gdy jeden i ten sam termin u różnych badaczy oznacza różne rzeczy. Zdarza się to zaskakująco często w literaturze neuronaukowej. Przykładem jest termin „pamięć robocza”, który, jak wcześniej pisaliśmy, ma w literaturze kilka różnych znaczeń. To może prowadzić do komplikacji w automatycznym przetwarzaniu literatury, ponieważ nie da się wiedzieć, które z tych znaczeń jest przyjmowane w danym konkretnym użyciu.

W CA każdy termin zaczyna się od pojedynczej definicji pojęcia. Badacze, którzy nie zgadzają się z tą definicją, mogą przedyskutować swoją różnicę zdań, używając wbudowanej funkcji dyskusji, podobnej do sposobu, w który rozwiązywane są konflikty w Wikipedii. Jednak jeżeli w trakcie dyskusji stanie się jasne, że jest to różnica pojęciowa nie do pogodzenia, pojęcie może być „rozwidłone”, w takim przypadku oryginalne pojęcie jest rozbijane na kilka znaczeń, z których każdy będzie miał swoją własną stronę i odrębnie uczestniczył w relacjach z zadaniami. Przykładowo, w przypadku pamięci roboczej początkowe pojęcie „pamięć robocza” będzie rozbite na oddzielne sensory, takie jak „pamięć robocza (utrzymanie)”, „pamięć robocza

(manipulacja)”, „pamięć robocza (czasowa)” z oryginalną stroną zmienioną na stronę ujednoznaczniającą dla różnych znaczeń. Przykład takiej strony dla pojęcia „hamowanie behawioralne” pokazuje **Ilustracja 3**.



The screenshot shows the Cognitive Atlas website interface. At the top left is the logo 'cognitive atlas beta'. To the right are navigation buttons for 'terms' and 'assertions', each with a 'browse »' link. A search bar with a dropdown arrow is labeled '- jump to page -'. The main content area is titled 'Concept Disambiguation' and features a 'CONCEPT' label next to the term 'behavioral inhibition'. Below this, a paragraph explains that the page lists concepts associated with the term and suggests changing external links. Two numbered entries are listed: 1. 'behavioral inhibition (cognitive)' described as a synonym for 'response inhibition', and 2. 'behavioral inhibition (temperament)' described as a characteristic of shyness and social anxiety. The footer contains a Creative Commons license icon, 'About Login' links, and funding information: 'NIMH Grant RO1MH082795, Russell Poldrack, PI. • Development partner: Squishymedia'.

Ilustracja 3. Zrzut z ekranu strony ujednoznaczniającej dla “hamowania behawioralnego”, która wskazuje na dwa różne znaczenia tego terminu.

I odwrotnie, powszechne jest również to, że różne terminy są używane do opisu tego samego procesu podstawowego. Przykładowo, terminy „pamięć deklaratywna” i „pamięć jawna” często są używane do odnoszenia się do tej samej funkcji umysłowej. W ramach CA możliwe jest określanie terminów jako synonimicznych, tak by kolejne analizy używające tej bazy wiedzy, rozpoznawały synonimiczność tych terminów. To rozwiązanie jest bardziej cenione niż łączenie ich w pojedyncze pojęcie, ponieważ zachowuje ono oryginalne terminy w bazie wiedzy, choć nadal odnotowuje je jako odnoszące się do tego samego procesu.

Udoskonalone rozszerzenie zapytań

Precyzja i wyniki wyszukiwania literatury mogą być znacznie usprawnione przez użycie wiedzy ontologicznej do prowadzenia rozszerzania zapytań przy szukaniu. Przykładowo, PubMed aktualnie rozszerza zapytania przy użyciu słownika MeSH, który (jak omawiano powyżej) nie odzwierciedla aktualnego stanu rzeczy. Przykładowo, zapytanie (sublexical route¹⁷) jest rozszerzane przez PubMed jako [sublexical[All Fields] AND („drug administration routes¹⁸”[MeSH Terms] OR („drug”[All Fields] AND „administration”[All Fields] AND „routes”[All Fields]) OR „drug administration routes”[All Fields] OR „route”[All Fields])]. Niejasność MeSH odnośnie do znaczenia „route”¹⁹ w tym kontekście prowadzi do niepoprawnego rozszerzenia zapytania, podczas gdy użycie bazy danych, która zawiera „sublexical route” jako pojęcie, dokładniej trafiałoby w konkretne wyrażenia. Dodatkowo można potencjalnie rozszerzyć zapytanie tak, by włączyć inne powiązane terminy takie jak „phonological assembly”²⁰, prowadząc do rezultatów wyszukiwania, które czynią bardziej prawdopodobnym znalezienie istotnej literatury.

Anotacja metadanych i metaanalizy

Dostępność dużych neuroobrazowych baz danych, szczególnie takich jak Brainmap.org (Laird et al., 2005) umożliwiło skuteczne metaanalizy. Jednak zdolność do przeprowadzania metaanaliz jest ograniczana przez metadane związane z danym zbiorem danych; aby ocenić który system mózgowy jest połączony z określonym procesem umysłowym, dane muszą być anotowane przy użyciu ontologii procesów umysłowych. Baza danych Brainmap aktualnie używa względnie zgrubnej ontologii procesów umysłowych, która ogranicza możliwość dokonywania dokładniejszych ocen na temat związków struktura-funkcja (Poldrack, 2006). Jednakże dostępność bardziej szczegółowej ontologii poznawczej zapewni możliwość wykonywania takich odwzorowań. Przykładowo, Poldrack ze współpracownikami (2009) używa zbioru anotowanych związków zadanie-proces, wraz z utajonymi zmiennymi zidentyfikowanymi na podstawie danych fMRI uzyskanych przy pomocy zestawu ośmiu zadań, do określania, które procesy umysłowe zostały odwzorowane na sieciach mózgowych. Dysponując większymi zbiorami danych anotowanych przy użyciu bardziej szczegółowej ontologii,

¹⁷ Szlak nieleksykalny.

¹⁸ Drogi przyjmowania leków.

¹⁹ Droga/szlak.

²⁰ Grupa fonologiczna.

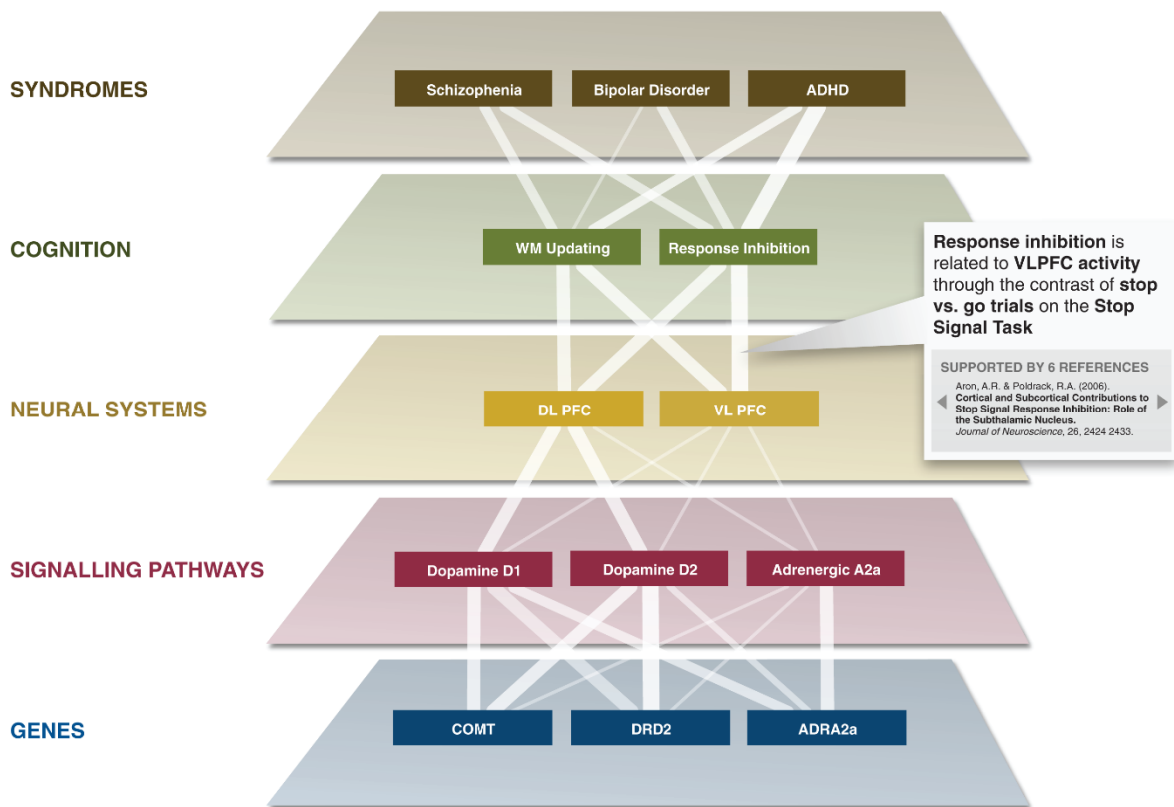
ten rodzaj analizy mógłby dostarczyć nowych wglądów w to, które rozróżnienia ontologiczne w ontologii procesów umysłowych są biologicznie realizowane, a które nie (Lenartowicz et al., 2010).

Testowanie teorii

Można także przewidywać użycie metaanaliz z CA do testowania większych teorii organizacji poznawczej. Przykładowo, w psychologii kategoryzacji trwa od dawna debata pomiędzy teoriami, które postulują jeden wspólny lub dwa odrębne procesy u podstaw kategoryzacji oraz pamięci rozpoznawczej (Poldrack and Foerde, 2008). Każda z tych teorii dokonałaby odmiennych założeń ontologicznych odnośnie do procesów poznawczych i ich relacji do zadań umysłowych. Te założenia ontologiczne potencjalnie mogłyby być przełożone na tezy na temat struktur kowariantnych w danych zgromadzonych o tych zadaniach, a różne teorie mogłyby być porównane pod względem dopasowania do tych danych, przy użyciu metod modelowania struktur kowariantnych. Chociaż system nie jest aktualnie zdolny do wspierania takiego testowania teorii, to jest to istotny cel dla tego systemu na przyszłość.

Translacyjne badanie zdrowia psychicznego

W psychiatrii coraz częściej nie używa się kategoryjnych, ale wymiarowych podejść do charakteryzowania zaburzeń psychicznych (Kraemer 2007). Trwające wysiłki takie jak projekt NIH Research Domain Criteria (Insel et al., 2010), dążą do scharakteryzowania tych podstawowych wymiarów w kategoriach ich podstaw poznawczych i neuronalnych. Społeczne, kolaboracyjne narzędzia budowania wiedzy, których dostarcza CA, dają takim projektom możliwość interaktywnego rozwijania ich baz wiedzy i anotowania bogatych zbiorów powiązań pomiędzy procesami poznawczymi a danymi na innych poziomach takich jak obwody neuronalne, szlaki sygnalizacji komórkowej, czy geny. CA odegra rolę istotnego połączenia, umożliwiającego powiązania, które muszą zostać dokonane między poziomem neuronalnym a poziomem symptomów i syndromów psychiatrycznych (zobacz: **Ilustracja 4**)



Ilustracja 4. Atlas poznawczy dostarcza ram do odniesienia funkcji i procesów biologicznych do symptomów i syndromów psychiatrycznych. Powiązania pomiędzy każdym poziomem na tym grafie odzwierciedlają odnośne związki empiryczne; siła każdego powiązania (oznaczana przez ich szerokość) jest proporcjonalna do obecnych w literaturze związków pomiędzy każdym zbiorem terminów (zdefiniowana jako współrzędna Jaccarda pomiędzy dwoma wyszukiwaniami terminu w PubMed). Każde powiązanie może również być połączone z określonymi wynikami empirycznymi, jak oznaczono w boksie ukazującym szczegółową anotację dla jednej z krawędzi.

Tłumaczenie treści ilustracji:

Syndromy	Schizofrenia, Zaburzenia afektywne dwubiegunowe, ADHD	
Poznanie	Aktualizacja pamięci roboczej, hamowanie reakcji	Hamowanie reakcji powiązane jest z aktywnością VL PFC w zadaniu „Stopu”.

Systemy nerwowe	Grzbietowa (DL PFC) i brzuszna (VL PFC) boczna kora przedczołowa	Na podstawie 6 publikacji: Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. <i>Journal of Neuroscience</i> , 26(9), 2424-2433.
Szlaki sygnalizacji komórkowej	Dopaminergiczny D1 i D2, Adrenergiczny A2a	
Geny	COMT, DRD2, ADRA2a	

Obecny status i plany na przyszłość

Rozwijanie CA rozpoczęliśmy w roku 2008. W fazie projektów analizowaliśmy strukturę wiedzy, która miała być reprezentowana i rozwijana we wstępny schemat bazy danych. We wstępnej fazie implementacji pracowaliśmy nad rozwojem zespołu do implementacji podstawowych funkcji prezentowania i edytowania bazy wiedzy (Miller i inni 2010), a także, wraz z niewielką grupą badaczy, rozpoczęliśmy wypełnianie bazy danych i ulepszanie projektu interakcji ze stroną. W aktualnej fazie kontynuujemy implementowanie nowych funkcjonalności, jak również udoskonalanie istniejącego interfejsu, szczególnie skupiając się na skalowaniu odpowiednim dla większej ilości treści. Zaczęliśmy również starania o otwarty udział innych badaczy z dziedziny. CA aktualnie ma wejścia dla 904 terminów, w tym 708 konstruktów umysłowych, 196 zadań, z definicjami obecnymi dla 795 z tych terminów. Większość z tych definicji nie będzie postrzeganych przez ekspertów w dziedzinie jako wystarczające, ale służą jako punkt wyjścia dla tych ekspertów do edytowania i udoskonalania. Na razie baza danych określa mniej niż 900 relacji. Podczas gdy pierwsza faza projektu skupiała się na definicjach pojęć, głównym celem kolejnej fazy będzie zwerbowanie szerokiego grona badaczy, którzy przekażą swoją wiedzę o wspomnianych relacjach.

Poza dodatkową treścią, przyszły rozwój strony skoncentruje się na trzech obszarach. Po pierwsze, planujemy dodać funkcje personalizujące, by umożliwić użytkownikom lepsze śledzenie istotnych informacji. Wliczamy w to śledzenie ich własnego udziału, a także bieżących zmian w interesujących tematach oraz polecanie treści i/lub publikacji istotnych dla zainteresowań użytkowników (na podstawie ich wcześniejszego wkładu). Po drugie, planujemy włączenie treści opartej na przeszukiwaniu opublikowanej literatury. Na przykład, możemy włączyć w bazę danych relacje oparte na związkach pomiędzy terminami w opublikowanej literaturze, mogące następnie służyć jako dane wejściowe do procesu manualnej

anotacji. Po trzecie, planujemy zaimplementować większą integrację z innymi bazami danych. Słownik CA jest już zintegrowany z otwartymi serwisami sieciowymi umożliwiającymi: eksplorację związków w literaturze (PubAtlas²¹); odwzorowanie powiązań znajdujących się w literaturze przy pomocy trójwymiarowego probabilistycznego atlasu struktury mózgu (PubBrain²²); kolaboracyjne wpisy ilościowych anotacji do metaanaliz wyników ze badań poznawczych (PhenoWiki²³). Podczas gdy CA reprezentuje aktualnie strukturę zadań w relatywnie zgrubny sposób, projekt CogPO rozwija bardziej szczegółową ontologię zadań (Turner and Laird, 2012; sprawdź przypis 2). Gdy stanie się to możliwe, połączymy się z nimi bezpośrednio, dostarczając dużo dokładniejszego sposobu modelowania drobnoziarnistych szczegółów zadań. Słownik CA zostanie włączony w przyszłą wersję bazy NeuroLex²⁴, a także zaimplementujemy większą integrację ontologii biologicznych takich jak GO (Ashburner et al., 2000).

Na koniec ważne jest, aby jasno stwierdzić, że sam CA nie będzie zawierał żadnych danych, ale planujemy od wewnątrz CA połączyć się z empirycznymi bazami danych. W najbliższym czasie obejmie to bezpośrednie połączenie z bazami danych neuroobrazowania opartymi na współrzędnych z ujawnionym API, takimi jak projekt NeuroSynth (Yarkoni i inni 2011) lub SumsDS25, które może dostarczyć zautomatyzowanego dostępu do metaanalizy danych empirycznych łączących pojęcia z bazy danych z systemami mózgowymi. Aktualnie baza Neurosynth zawiera mapy wnioskowań wprzód i wstecz dla wszystkich terminów pojęć z CA. Co więcej, jeśli inne bazy danych staną się dostępne (np. dane dotyczące powiązań w całym genomie, mapowania lezji, behawioralne dane pacjentów itd.), to one również mogą być połączone bezpośrednio z tą bazą danych. Również należy zauważyć, że podczas gdy CA jest obecnie skoncentrowane na pojęciach z literatury dotyczącej ludzkiej psychologii, istotne będzie także objęcie pojęć z literatury dotyczącej zwierząt niebędących ludźmi po to, aby dostarczyć rzeczywiście systematycznego przeglądu literatury.

²¹ <http://www.pubatlas.org>

²² <http://www.pubbrain.org>

²³ <http://www.phenowiki.org>

²⁴ <http://www.neurolex.org>

²⁵ <http://sumsdb.wustl.edu:8081/sums/index.jsp>

Wnioski

Odwzorowanie procesów umysłowych na systemy mózgu do tej pory polegało na nieformalnych reprezentacjach procesów umysłowych i zadań używanych do manipulowania nimi. Sądzymy, że to podejście jest fundamentalnie ograniczone i że kontynuowanie postępu naukowego w neuronauce poznawczej wymaga rozwoju i adaptacji formalnych baz wiedzy, które dostarczają bardziej systematycznych objaśnień teorii poznawczych i ich związków z danymi empirycznymi. Celem CA jest dostarczenie zasobów, które odzwierciedlą spojrzenie całej społeczności, a zainteresowanych badaczy zapraszamy do współpracy przez stronę cognitiveatlas.org. Mamy nadzieję, że dzięki udziałowi badaczy z całego obszaru CA stanie się standardową ontologią funkcji umysłowych.

Podziękowania

Artykuł powstał dzięki wsparciu grantu RO1MH082795 z NIH (dla Russella A. Poldracka) i Consortium for Neuropsychiatric Phonemics [grant NIH Roamap for Medical Research UL1-DE019580, PL1MH083271 (dla Roberta M. Bildera) i RL1LM009833 (dla D. Scotta Parkera)]. Dziękujemy Rajeevowi Raizada za pomocne uwagi do szkicu tego artykułu.

Literatura

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nature Genetics*, *25*, 25–29.
- Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.
- Bard, J. B. L., & Rhee, S. Y. (2004). Ontologies in biology: design, applications and future challenges. *Nature Review Genetics*, *5*, 213–222.
- Bechhofer, S., & Miles, A. (2009). *SKOS Simple Knowledge Organization System Reference*. W3C recommendation, W3C. Retrieved from <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>
- Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American* *284*, 34–43.

- Bilder, R. M., Sabb, F. W., Parker, D. S., Kalar, D., Chu, W. W., Fox, J., Freimer, N. B., & Poldrack, R. A. (2009). Cognitive ontologies for neuropsychiatric phenomics research. *Cognitive Neuropsychiatry*, *14*, 419–450.
- Boring, E. G. (1950). *A History of Experimental Psychology*, 2nd Edition. Prentice Hall: Englewood Cliffs, NJ.
- Bowden, D. M., & Dubach, M. F. (2003). Neuronames 2002. *Neuroinformatics*, *1*, 43–59.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the stroop effect. *Psychological Review*, *97*, 332–361.
- Donohoe, G., Corvin, A., & Robertson, I. H. (2006). Evidence that specific executive functions predict symptom variance among schizophrenia patients with a predominantly negative symptom profile. *Cognitive Neuropsychiatry*, *11*, 13–32.
- Edelman, G. M. (1989). *The Remembered Present: a Biological Theory of Consciousness*. Basic Books: New York..
- Forté, A., Larco, V., & Bruckman, A. (2009). Decentralization in Wikipedia governance. *Journal of Management Information Systems*, *26*(1), 49-72.
- Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, *438*, 900–901.
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* *14*, 477–485.
- Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*, 199–220.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (rdoc): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*, 748–751.
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615–622.
- Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in wikipedia: quality through coordination, In *CSCW 2008: Proceedings of the ACM Conference on Computer-Supported Cooperative Work*, ACM Press: New York.
- Kittur, A., Suh, B., Pendleton, B. A., & Chi, E. (2007). He says, she says: conflict and coordination in wikipedia, In *CHI 2007: Proceedings of the ACM Conference on Human-factors in Computing Systems*, ACM Press: New York, NY.
- Kraemer, H. C. (2007). DSM categories and dimensions in clinical and research contexts. *International Journal of Methods in Psychiatric Research*, *16*(S1), S8–S15.

- Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). Brainmap: the social evolution of a human brain mapping database. *Neuroinformatics*, 3, 65–78.
- Laird, J., Newell, A., & Rosenbloom, P. (1987). Soar: an architecture for general intelligence. *Artificial intelligence*, 33(1), 1-64.
- Lenartowicz, A., Kalar, D., Congdon, E., & Poldrack, R. A. (2010). Towards an ontology of cognitive control. *Topics in Cognitive Science*, 2(4), 678-692.
- Martone, M. E., Gupta, A., & Ellisman, M. H. (2004). E-neuroscience: challenges and triumphs in integrating distributed data from molecules to brains. *Nature Neuroscience*, 7, 467–472.
- Miller, E., Seppa, C., Kittur, A., Sabb, F., & Poldrack, R. A. (2010). The cognitive atlas: employing interaction design processes to facilitate collaborative ontology creation. *Nature Proceedings*. Retrieved from <http://dx.doi.org/10.1038/npre.2010.4532.1>.
- Müller, H.-M., Kenny, E. E., & Sternberg, P. W. (2004). Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biology*, 2, e309.
- Olton, D. S., Becker, J. T., & Handelmann, G. E. (1979). Hippocampus, space, and memory. *Behavioral and Brain Sciences*, 2(3), 313-322.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in cognitive sciences*, 10(2), 59-63.
- Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience & Biobehavioral Reviews*, 32(2), 197-205.
- Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science*, 20(11), 1364–1372.
- Price, C., & Friston, K. (2005). Functional ontologies for cognition: the systematic definition of structure and function. *Cognitive Neuropsychology*, 22, 262–275.
- Ruths, T., Ruths, D., & Nakhleh, L. (2009). Gs2: an efficiently computable measure of go-based similarity of gene sets. *Bioinformatics*, 25, 1178–1184.
- Sabb, F. W., Bearden, C. E., Glahn, D. C., Parker, D. S., Freimer, N., & Bilder, R. M. (2008). A collaborative knowledge base for cognitive phenomics. *Molecular Psychiatry*, 13(4), 350–360.
- Shotton, D. (2010). Cito, the citation typing ontology. *Journal of biomedical semantics*, 1(Suppl. 1), S6.
- Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., & Rosse, C. (2005). Relations in biomedical ontologies. *Genome biology*, 6(5), R46.

- Sternberg, S. (1969). The discovery of processing stages: Extensions of donders' method. *Acta psychologica*, 30, 276-315.
- Turner, J. A., & Laird, A. R. (2012). The cognitive paradigm ontology: design and application. *Neuroinformatics*. 10(1), 57-66
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D., & Wager, T. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8), 665.

Abstract

Cognitive neuroscience aims to map mental processes onto brain function, which begs the question of what “mental processes” exist and how they relate to the tasks that are used to manipulate and measure them. This topic has been addressed informally in prior work, but we propose that cumulative progress in cognitive neuroscience requires a more systematic approach to representing the mental entities that are being mapped to brain function and the tasks used to manipulate and measure mental processes. We describe a new open collaborative project that aims to provide a knowledge base for cognitive neuroscience, called the *Cognitive Atlas* (accessible online at <http://www.cognitiveatlas.org>), and outline how this project has the potential to drive novel discoveries about both mind and brain.

Keywords: ontology; informatics; neuroimaging; cognitive science