

THATCamp Warszawa

6 XI 2013

Środa, 6 XI 2013, 17:00-22:00,

[MiTo art.café.books](#) ([ul. Waryńskiego 28](#))

[Wstęp wolny]

Organizatorzy: Centrum Humanistyki Cyfrowej IBL PAN, MiTo

Patronat medialny: Człowiek i technologie

PROGRAM

17:00-17:15	Wprowadzenie
17:00	Piotr Seweryn Rosół (MiTo) <i>Uroczyste otwarcie obrad</i>
17:06	Maciej Maryl (CHC IBL PAN) <i>Po co literaturoznawcom komputery? O pożytkach z humanistyki cyfrowej</i>
17:15-19:35	Prezentacje warsztatowe
17:15	Włodzimierz Bolecki (IBL PAN) <i>Nowe metody prezentacji wyników badań naukowych na przykładzie witryny projektu „Sensualność w kulturze polskiej”</i>
17:35	Jerzy Tyszkiewicz (UW) <i>Zmienne koleje losu programu edytorskiego „Magik”</i>
18:15	Jan Rybicki (UJ) i Maciej Eder (UP) <i>Stylistyka komputerowa: sukcesy i wyzwania</i>
18:55	Maciej Piasecki (PWr) <ul style="list-style-type: none">• CLARIN-PL - od programów dla wtajemniczonych do przystępnych informatycznych narzędzi badawczych do analizy zapisów języka• Słowosieć - wielka sieć znaczeń na potrzeby tropienia znaczeń w sieci słów.
19:40-20:30	Krótkie prezentacje (10 min.) cz. I – Bazy
	Piotr Wciślik (OSA, CEU) <i>Zabawy z MARC w archiwum i poza nim.</i>
	Małgorzata Marciniak i Agnieszka Mykowiecka (IPI PAN) <i>Korpus tekstów o sztuce</i>
	Maciej Strybel (ICM UW) <i>CEJSH- bibliograficzna baza czasopism humanistycznych i społecznych.</i>
	Joanna Raczkowska (ICM UW) <i>„Otwórz książkę”. Dziel się wiedzą, twórz naukę</i>
20:35-21:25	Krótkie prezentacje (10 min.) cz. II – Książka
	Łukasz Karpiński (UW) <i>Parametryczny opis tekstu - program Pantext 2.0</i>
	Jakub Jagiełło (UAM) <i>Symulacje jako narzędzie prezentacji idei humanistycznych</i>
	Gwido Zlatkes (UC Riverside) <i>Dwie przyszłości książki</i>
	Urszula Pawlicka (UWM) <i>Antologia polskiej literatury cyfrowej - work in progress</i>
21:30-22:20	Krótkie prezentacje (10 min.) cz. III – Kultura literacka
	Aneta Bassa (UW) <i>Krytyka literacka we Francji a internet</i>
	Jarosław Kopeć (BN) <i>Cyfrowy humanista bada czytelnictwo</i>
	Marcin Wilkowski (Historia i Media) <i>Otwarty skrypt akademicki - potencjał i problemy</i>
	Michał „rysiek” Woźniak (FWiOO) <i>Kultura tylko do odczytu</i>

ABSTRAKTY

Maciej Maryl

Centrum Humanistyki Cyfrowej Instytutu Badań Literackich Polskiej Akademii Nauk

Po co literaturoznawcom komputery? O korzyściach z humanistyki cyfrowej

Wprowadzenie w problematykę spotkania. Prezentacja głównych nurtów humanistyki cyfrowej w kontekście badań literackich i lingwistycznych. Być może uda się nawet przekonująco odpowiedzieć na pytanie postawione w tytule. [\[Powrót\]](#)

Włodzimierz Bolecki

Instytut Badań Literackich Polskiej Akademii Nauk

Nowe metody prezentacji wyników badań naukowych na przykładzie witryny projektu „Sensualność w kulturze polskiej”

Przedstawienie nowoczesnej platformy cyfrowej, pełniącej funkcję tematycznej, specjalistycznej, encyklopedii multimedialnej, dostępnej w formie serwisu internetowego. „[Sensualność w kulturze polskiej](#)” to projekt badawczy skupiony na roli zmysłów w języku, piśmiennictwie i sztuce od średniowiecza do współczesności. Sensualność rozumiana jest tu jako historycznie zmienny zbiór form reprezentacji ludzkich zmysłów. [\[Powrót\]](#)

Jerzy Tyszkiewicz

Uniwersytet Warszawski

Zmienne koleje losu programu edytorskiego „Magik”

„Magik” to specyficzny edytor tekstów, którego podstawową funkcją nie jest pisanie i formatowanie, tylko porównywanie tekstów napisanych w innych edytorach. Program został napisany przez studentów Wydziału Matematyki, Informatyki i Mechaniki UW pod kierunkiem prelegenta, który opowie zarówno o samym programie, jak i jego dalszych losach. [\[Powrót\]](#)

Jan Rybicki i Maciej Eder

Uniwersytet Jagielloński, Uniwersytet Pedagogiczny w Krakowie

Stylistyka komputerowa: sukcesy i wyzwania

Na ogół wystarczy obliczyć częstości kilkudziesięciu najczęstszych słów w kilku czy kilkunastu badanych tekstach, np. powieściach, by skutecznie wnioskować o ich autorstwie. Taka literacka daktyloskopia już od dawna wspomaga atrybucję autorską; ale czy wykrywane metodami statystycznymi podobieństwa między tekstami różnych pisarzy mogą stać się nowym narzędziem analizy, kategoryzacji i periodyzacji literackiej? [\[Powrót\]](#)

Maciej Piasecki

Politechnika Wroclawska

CLARIN-PL - od programów dla wtajemniczonych do przystępnych informatycznych narzędzi badawczych do analizy zapisów języka

[CLARIN](#) (Common Language Resources & Technology Infrastructure) to ogólnoeuropejska infrastruktura naukowa, która umożliwi badaczom z dziedziny nauk humanistycznych i społecznych wygodną pracę z bardzo dużymi zbiorami tekstów. Celem CLARIN pokonanie szeregu barier w dostępie do tzw. technologii językowej przez naukowców, którzy nie mają specjalistycznej wiedzy w dziedzinie informatyki lub inżynierii języka naturalnego. CLARIN-PL to polska część wielkiej europejskiej sieci, ściśle zintegrowana, ale szczególnie ukierunkowana na język polski, zarówno w piśmie jak i mowie. W trakcie prezentacji zostaną przedstawione zarówno plany jak pierwsze dostępne narzędzia. [\[Powrót\]](#)

Maciej Piasecki

Politechnika Wroclawska

Słowsiec - wielka sieć znaczeń na potrzeby tropienia znaczeń w sieci słów.

[Słowsiec](#) (ang. plWordNet) to unikalny słownik zapewniający sformalizowany opis znaczeń ogromnej liczby słów języka polskiego. Każde znaczenie jest opisane poprzez jego relacje semantyczne z wieloma innymi znaczeniami innych słów. W trakcie prezentacji przedstawione zostaną podstawowe zasady budowy Słowsieci, sposobu posługiwania się nią i możliwości jej wykorzystania w analizie tekstów i mowy. Omówimy też krótko trwające prace nad dalszą rozbudową Słowsieci mające na celu przekształcenie jej w wyczerpujących opis polskiego systemu relacji leksykalno-semantycznych. [\[Powrót\]](#)

Piotr Wciślik

Open Society Archives at the Central European University

Zabawy z MARC w archiwum i poza nim

W wystąpieniu opowiem o powstawaniu cyfrowej kolekcji bibuły w Open Society Archives, a zwłaszcza o naszych eksperymentach z rekordami bibliograficznymi Biblioteki Narodowej. W tym kontekście poruszę również bardziej ogólną kwestię potencjału przetwarzania danych bibliograficznych w badaniach humanistycznych. [\[Powrót\]](#)

Małgorzata Marciniak i Agnieszka Mykowiecka

Instytut Podstaw Informatyki Polskiej Akademii Nauk

Korpus tekstów o sztuce

W prezentacji chcemy odpowiedzieć na pytanie w jakim celu zajmujemy się gromadzeniem tekstów z dziedziny kultury i sztuki. Dlaczego Narodowy Korpus Języka Polskiego nam nie wystarcza jeśli chcemy na przykład umieć efektywniej korygować artykuły dziedzinowe. Czy wystarcza nam słownik ogólnego języka? Jeśli nie, to jak nauczyć się specyficznego słownictwa i zwrotów? Dlaczego chcemy umieć rozpoznawać w tekstach cytaty w obcym języku i na ile powinniśmy umieć identyfikować nazwy własne, w tym tytuły tak charakterystyczne dla tekstów o sztuce. [\[Powrót\]](#)

Maciej Strybel

**Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego,
Uniwersytet Warszawski**

CEJSH- bibliograficzna baza czasopism humanistycznych i społecznych.

CEJSH (The Central European Journal Of Social Sciences And Humanities) jest bazą danych (bibliograficzno-abstraktową i pełnotekstową), powstałą w 2004 r. jako wspólne przedsięwzięcie akademii nauk krajów Grupy Wyszehradzkiej. CEJSH współpracuje z ok. 600 czasopismami z dziedziny humanistyki i nauk społecznych, ukazującymi się w Europie Środkowej i Wschodniej. Baza współdziała z ICM UW w ramach Wirtualnej Biblioteki Nauki, promując idee Otwartego Dostępu i współtworząc największe w Polsce cyfrowe zasoby wiedzy. Od 2011 roku CEJSH znajduje się na liście baz referencyjnych dla czasopism z zakresu nauk humanistycznych i społecznych uwzględnianych przez MNiSW podczas ewaluacji krajowych czasopism naukowych. [\[Powrót\]](#)

Joanna Raczkowska

**Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego,
Uniwersytet Warszawski**

„Otwórz książkę”. Dziel się wiedzą, twórz naukę

Wystąpienie ma charakter ściśle informacyjny. Ma na celu prezentację serwisu „Otwórz książkę” (<http://otworzksiazke.pl/>), prowadzonego przez Centrum Otwartej Nauki ICM UW. To projekt ponadregionalnej, ponadinstytucjonalnej i multidyscyplinarnej kolekcji prac polskich uczonych, złożonej ze zdigitalizowanych publikacji naukowych. Zarys mapy działalności serwisu obejmuje: cele, rodzaje zbiorów i sposoby ich udostępniania (w tym narzędzia technologiczne i prawne), społeczność czytelników i autorów (w tym kryteria doboru autorów i diagnozę potrzeb użytkowników), partycypację serwisu w kształtowaniu polskiej komunikacji naukowej, perspektywy itp. Wskażemy też na to, co wyróżnia nas na tle bogatej oferty polskich cyfrowych bibliotek naukowych. [\[Powrót\]](#)

Łukasz Karpiński

Uniwersytet Warszawski

Parametryczny opis tekstu - program Pantext 2.0

Prezentacja dotyczyć będzie wyników raportów generowanych przez program komputerowy Pantext ukierunkowany na eksperymentalne metody pomiaru i analizy wybranych tekstów nacechowanych różnymi emocjami i przedstawienia ich w formie danych liczbowych lub wykresów. Docelowym efektem prac będzie podanie metody analitycznej oraz gotowych zestawień wartości liczbowych dla poszczególnych rodzajów tekstu. Analiza parametryczna stawia sobie za zadanie przekazanie translatorowi właściwych zmiennych numerycznych o określonych wartościach, co pozwoli algorytmowi translatora efektywnie dobrać odpowiedni (zestaw) ekwiwalent(ów) w innym języku. Opracowywany algorytm może wskazywać również potencjalne formy nietypowe – metafory, a także archaizmy. Prace stanowią również wstęp do budowy modeli cyfrowych definicji. [\[Powrót\]](#)

Jakub Jagiełło

Uniwersytet Adama Mickiewicza

Symulacje jako narzędzie prezentacji idei humanistycznych

Wykorzystanie narzędzi, jakie oferuje informatyka i interfejsy współczesnych systemów, do prezentacji wiedzy naukowej (w formie symulacji i różnych form hipertekstowych) nie jest niczym nowym, jednak do tej pory najlepiej sprawdzało się na gruncie nauk ścisłych, takich jak matematyka, fizyka i astronomia — w których większość teorii daje się sprowadzić do formy wykresu, formuły matematycznej lub schematu. Dlaczego jednak nie użyć takiego sposobu narracji i w innych dziedzinach? Aby pokazać, że jest to możliwe przygotowałem hipertekstowo-symulacyjne realizacje kilku teorii humanistycznych (jak np. niezdeterminowania przekładu Quine'a). [\[Powrót\]](#)

Gwido Zlatkes

University of California, Riverside

Dwie przyszłości książki

Wraz z pojawieniem się książki elektronicznej, kindle i i-padów obwieszczano zmierzch książki takiej jak ją znamy, kodeksu czyli papierowego tomu. Już obecnie widać, że prorocтва te były co najmniej grubo przesadzone jeśli nie wręcz fałszywe. Ale czy całkowicie błędne – bo jeśli nie, to w jakim zakresie i pod jakimi względami? W moim wystąpieniu postaram się pokrótce nakreślić jaka przyszłość, czy raczej jakie przyszłości – elektroniczne i papierowe – zdają się czekać książkę, czyli przekaz oparty na zapisanym tekście. [\[Powrót\]](#)

Urszula Pawlicka

Uniwersytet Warmińsko-Mazurski

Antologia polskiej literatury cyfrowej - work in progress

Prezentacja powstającej, pierwszej w Polsce, antologii literatury cyfrowej pod red. Piotra Mareckiego, Urszuli Pawlickiej i Mariusza Pisarskiego. Antologia, zawierająca ok. 40 utworów, ukaże się w Korporacji Ha!art jeszcze tego roku. Podczas wystąpienia przedstawię cel antologii, budowę, kryteria wyboru projektów oraz problemy i wątpliwości wyłaniające się w trakcie jej przygotowywania. Sugestie ze strony słuchaczy mogą być pomocne w dalszych etapach pracy nad antologią. [\[Powrót\]](#)

Aneta Bassa

Uniwersytet Warszawski

Krytyka literacka we Francji a internet

W moich badaniach zajmuję się zmianami zachodzącymi we francuskiej krytyce literackiej pod wpływem nowych technologii. Prezentacja będzie porównaniem portali poświęconych książkom, działających według odrębnych modeli. Kolejno zaprezentuję Larepubliquedeslivres.com Assouline'a, który z bloga jednego krytyka przekształcił się w portal moderowany przez kolektyw. Przykładem pośrednim jest Onlalu.com założony z inicjatywy profesjonalistów sektora wydawniczego, w którym recenzje krytyków przeplatają się z opiniami amatorów. Kończąc przedstawię serwis społecznościowy Babelio.com,

<http://thatcamp.waw.pl/>

skupiający ponad 50 tys. aktywnych czytelników. Analiza modeli funkcjonowania, profilu autorów i organizacji treści pozwala określić, w jaki sposób krytyka literacka reaguje na zmiany w technologii, a także na coraz silniejszy głos amatorów. [\[Powrót\]](#)

Jarosław Kopec

Pracownia Badań Czytelnictwa, Biblioteka Narodowa

Cyfrowy humanista bada czytelnictwo

W ramach badań czytelnictwa prowadzonych przez Bibliotekę Narodową prowadzę mały projekt badawczy, którego zadaniem jest odpowiadanie na pytania związane z czytelnictwem w związku z nowymi technologiami komunikacyjnymi. Pierwszy etap badań będzie głęboko zakorzeniony w metodach cyfrowej humanistyki, a jego celem będzie wskazanie strategii zdobywania i przetwarzania informacji z wykorzystaniem Internetu. Nadziejami i obawami związanymi z tym projektem chętnie podzielę się z uczestnikami THATCampu. Liczę na konkretny feedback. [\[Powrót\]](#)

Marcin Wilkowski

Historia i Media

Otwarty skrypt akademicki - potencjał i problemy

W wystąpieniu przedstawię koncepcje udostępniania i rozbudowy otwartego podręcznika do historii cyfrowej ("Wprowadzenie do historii cyfrowej"). [\[Powrót\]](#)

Michał „rysiek” Woźniak

Fundacja Wolnego i Otwartego Oprogramowania

Kultura tylko do odczytu

Kultura zawsze była czymś wspólnym, w kulturze się uczestniczyło, kulturę się współtworzyło. Toksyczny mix nowoczesnych form komunikacji, zapośredniczenia przez środki techniczne, archaicznego prawa autorskiego i dużych pieniędzy stworzył jednak sytuację, w której uczestnictwo staje się „naruszeniem praw na dobrach niematerialnych”, a kultura staje się coraz bardziej tylko do odczytu. Czy dalej możemy stać na ramionach olbrzymów, by widzieć dalej, czy musimy na to wykupić licencję? [\[Powrót\]](#)