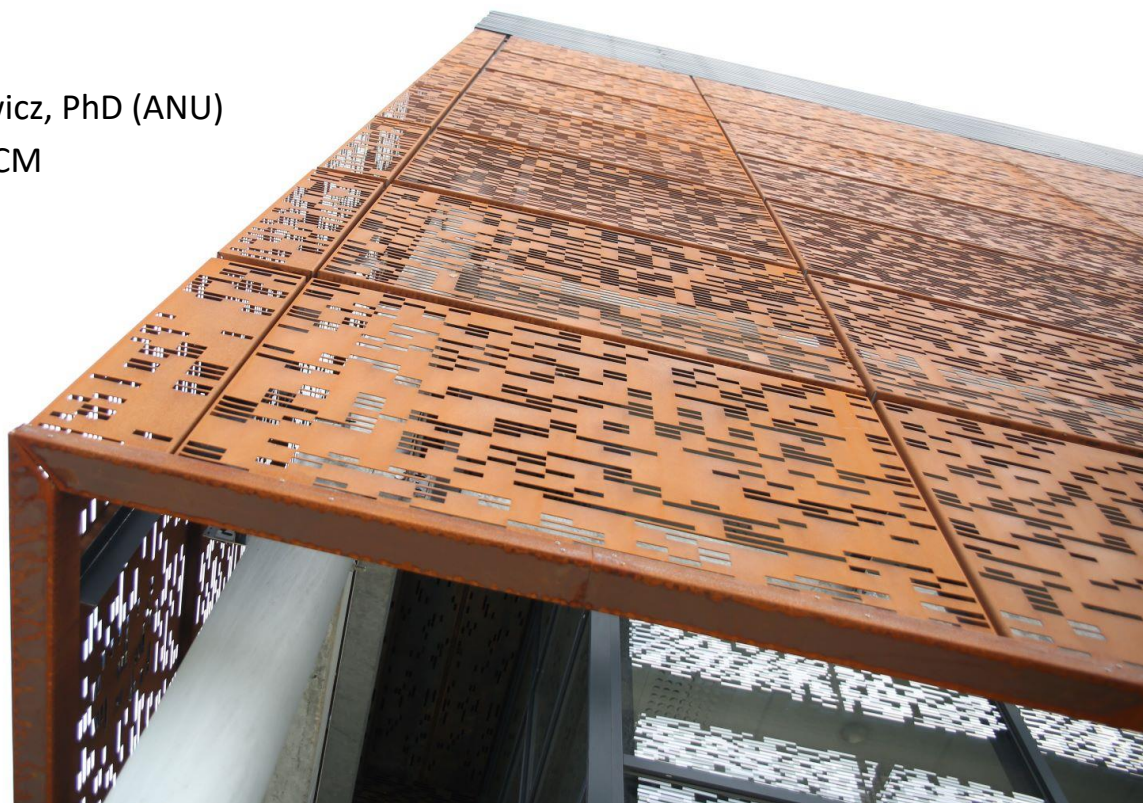




**DOROBK INTERDYSCYPLINARNEGO
CENTRUM MODELOWANIA
MATEMATYCZNEGO I KOMPUTEROWEGO (ICM)
UNIwersYTETU WARSZAWSKIEGO
W OBSZARZE NAUK BIOLOGICZNYCH I MEDYCZNYCH**

Materiał do dyskusji
o rozwijaniu współpracy
pomiędzy WUM a ICM, UW

Marek Michalewicz, PhD (ANU)
Dyrektor (p.o.) ICM



1. KRÓTKI WSTĘP O ICM	4
2. DZIAŁALNOŚĆ I DOROBIEK NAUKOWY ICM W OBSZARZE NAUK BIOLOGICZNYCH I MEDYCZNYCH	4
2.1 Sztuczna Inteligencja (AI), Uczenie Maszynowe (ML) oraz Obrazowanie Medyczne. Działalność Grupy Sztucznej Inteligencji, Wizualizacji i Analizy Obrazowej w Diagnostyce Medycznej	4
2.1.1. AI w obrazowaniu medycznym	5
2.1.2. Wizualizacja medyczna, AR/VR i VisNow	5
2.1.3. Projekt „START: Nowatorskie metody inżynierii tkankowej wspomagające gojenie i regenerację ścięgien i więzadeł”	5
2.1.4. Projekt „TRAFIC: Nowatorska metoda czynnościowej oceny zwężenia w tętnicy wieńcowej za pomocą modelowania przepływów in silico, na podstawie obrazu wielorzędowej tomografii komputerowej”	5
2.1.5. Projekt „INPLA: Integracja wielomodalnych informacji dynamicznych, obrazowych i funkcjonalnych w diagnostyce i planowaniu terapii kończyny dolnej”	5
2.1.6. Projekt „STOPA: Modelowanie matematyczne w planowaniu zabiegów osteotomii kończyny dolnej”	6
2.1.7. AR w zabiegach kardiologii interwencyjnej	6
2.1.8. Uogólnione prawo Murray’a dla bifurkacji tętnic wieńcowych	6
2.1.9. Analiza kształtu komórek w obrazowaniu mikroskopowym	6
2.1.10. Segmentacja struktur mózgowych w wodogłowie	6
2.1.11. Biomechanika	6
2.1.12. Rehabilitacja w udarach mózgu	6
2.1.13. Modelowanie chodu	6
2.1.14. Projekt „KARDIONET: Rozwój zaawansowanych metod obrazowania medycznego, rozproszonej akwizycji, archiwizacji i teletransmisji w zintegrowanym systemie opieki kardiologicznej ostrej zespołów wieńcowych”	7
2.1.15. Projekt „CePT: Centrum Badań Przedklinicznych i Technologii”	7
2.1.16. Rekonstrukcja ruchu lewej komory serca na podstawie 4D USG	7
2.2 Matematyczne modelowanie procesów termodynamicznych i biologicznych w skali makromolekularnej. Działalność dr Zuzanny Szymańskiej	7
2.2.1 Modelowanie matematyczne procesu topnienia DNA”	7
2.2.2. Modelowanie matematyczne roli białka CDC6 w procesie inicjacji mitozy	7
2.2.3. Estymacja bayesowska nielokalnej funkcji proliferacji	8
2.2.4. Modelowanie matematyczne procesu przebudowy włókien kolagenowych	8
2.2.5. Modelowanie matematyczne procesu przebudowy włókien kolagenowych – the role of tendon specific fibroblasts	9
2.3 Modelowanie matematyczne indukcji białka Hsp70 w warunkach szoku termicznego. Działalność dr Grzegorza Dudziuka	9
2.4 Epidemiologia oraz geografia zdrowia. Działalność dr Anety Afelt	10
2.4.1. Zika&Co: Understanding the dynamics of emergence of Aedes-born diseases	10
2.4.2. ARYZA - Assessing the Risk of emergence of Yellow fever and Zika in Asia	11
2.4.3. SOCIAL - Sociology Impact On Aedes-Borne Diseases Local Outbreaks	11
2.4.4. Analiza zmienności czasowej i przestrzennej malarii w Chinach, wieloletnie	12
2.4.5 Malaria dynamic depending on climate fluctuation, China case	12
2.4.6. Inne projekty	13
2.5 Metody informatyczne, aplikacje i API w bioinformatyce i genomice. Działalność Prof. Piotra Bały	13
2.5.1. HEAP - Human Exposome Assessment Platform	13
2.5.2. NIASC - Nordic Information for Action eScience Center	13
2.5.3. Aplikacje bioinformatyczne w PL-Grid	13
2.5.4. Analiza danych medycznych w środowisku gridowym	13

Działalność ICM w Naukach Biologicznych i Medycznych

2.5.5. Genetyczne bazy danych w projekcie OCEAN (POiG)	13
2.5.6. Kardionet (Fundusze Norweskie 2008-2011)	13
2.6 Bioinformatyka, algorytmy bioinformatyczne oraz inna działalność Witolda Rudnickiego	14
2.7 Systemy uczenia maszynowego i selekcji w biologii oraz inna działalność i Mirona Kursy	14
2.6.1 Boruta – system do ekstrakcji istotnych czynników	14
2.6.2 FlowForest dla cytometrii przepływowej	14
2.6.4 Fenomenologiczne analizy mechanizmów neurochemicznych u szczura, w kontekście wpływu substancji, cravingu i interakcji społecznych	15
2.8 Computational modelling of tissue, cell colonies and cancer growth. Działalność dr Macieja Cytowskiego	15
2.9 Działalność innych pracowników ICM	16
3. DZIAŁALNOŚĆ EDUKACYJNA ICM	16
3.1 Inżynieria Obliczeniowa	16
3.2 OMICS Data Science: Bioinformatics and Analysis of large-scale biomedical data	17
3.3 Warsaw Team na światowych konkursach Student Cluster Competition	18
4. DZIAŁALNOŚĆ WSPIERAJĄCA NAUKĘ	19
4.1 Wirtualna Biblioteka Nauki i otwarty dostęp	19
4.2 Konferencje oraz promocja biologii i medycyny obliczeniowej	19
4.2.1 Computational Challenges in Life Science, Melbourne, Australia	19
4.2.2 Joint Fujitsu-A*STAR Computational Bio-Medicine Forum, Singapur	19
4.2.3 Supercomputing Frontiers i Supercomputing Frontiers Europe	20
5. INFRASTRUKTURA	22
5.1 Serwerownie	22
5.2 Sprzęt obliczeniowy	23
5.3 Przechowywanie danych	23
5.4 Sieci Komputerowe	23
5.5 Wizualizacja	24
6. EKSPERTYZA	25
7. PODSUMOWANIE	30
8. PUBLIKACJE	31

I. Krótki wstęp o ICM

Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego (ICM UW), powstało w 1993 roku jako podstawowa, ogólnouczelniana jednostka naukowa Uniwersytetu Warszawskiego. ICM zaangażowane jest w interdyscyplinarne badania naukowe oparte na modelowaniu matematycznym, symulacjach i modelowaniu komputerowym, obliczeniach w wielu i w wielkiej skali oraz dydaktyką w powyższych obszarach. Oprócz zaangażowania w obszarze opisanym w poniższym dokumencie, pracownicy ICM zajmują się Numeryczną Prognozą Pogody wraz z publikowaniem wyników prognoz w znanym serwisie pogodowym *meteo.pl*; problemami związanymi z lotnictwem cywilnym, modelowaniem procesów społecznych i seismologią, zapewnieniem dostępu polskich naukowców do literatury naukowej poprzez utrzymywanie Wirtualnej Biblioteki Nauki, otwartym dostępem do literatury naukowej, oraz wielu innymi problemami.

ICM zarządza dwoma centrami danych w Warszawie. Oddane do użytku w 2016 roku Centrum Technologii ICM (CT-ICM) w Białołęce ma ok. 10.000 m² powierzchni technicznej w której znajdują się dwa superkomputery światowej klasy: Petaflopowy Cray XC-40 (Okeanos) do tradycyjnych intensywnych obliczeń numerycznych, oraz klaster Huawei (Enigma) do analityki wielkich danych i obliczeń chmurowych. W serwerowni CT-ICM znajduje się również sprzęt do przechowywania danych (data storage) na ok. 20 PB danych. W nowozbudowanej auli wykładowej dla 70-ciu słuchaczy znajduje się unikalny sprzęt wizualizacyjny z 16-oma monitorami i oprogramowaniem pozwalającym na transmisję wykładów czy obrazów z całego świata.

W obszarze usług typu High Performance Computing (HPC) i usług chmurowych, ICM wspomaga ok. 1000 użytkowników z całej Polski korzystających z naszych superkomputerów i infrastruktury obliczeniowej, sieciowej oraz przechowywania wielkich danych.

ICM zatrudnia obecnie 110 pracowników.

2. Działalność i dorobek naukowy ICM w obszarze Nauk Biologicznych i Medycznych

2.1 Sztuczna Inteligencja (AI), Uczenie Maszynowe (ML) oraz Obrazowanie Medyczne. Działalność Grupy Sztucznej Inteligencji, Wizualizacji i Analizy Obrazowej w Diagnostyce Medycznej ¹

Członkowie grupy: Bartosz Borucki (lider), dr Krzysztof Nowiński, dr Jakub Zieliński, dr Jędrzej Nowosielski, Norbert Kapiński

2.1.1. AI w obrazowaniu medycznym²

Opracowywanie i wykorzystanie technik głębokiego uczenia się (DL) oraz uczenia maszynowego (ML) w problemach związanych z obrazowaniem medycznym i radiologią. Od poprawy jakości obrazów, poprzez segmentację, detekcję, automatyzację diagnostyki, wspomaganie decyzji, aż po strukturyzację raportowania, optymalizację pracy radiologicznej (np. triage) i automatyzację analizy i generowania opisów badań. Głównie w zakresie MRI, CT i USG, w większości obszarów anatomicznych. Obecne główne obszary zainteresowania: ortopedia / kończyna dolna, klatka piersiowa / oraz płuca.

2.1.2. Wizualizacja medyczna, AR/VR i VisNow³

Wykorzystanie technik wizualnych w zastosowaniach medycznych. Zarówno w zakresie edukacyjnym, jak i klinicznym czy interwencyjnym. Rozwój własnego oprogramowania do celów wizualizacji i prototypowania w medycynie (VisNow). Rozwój rozwiązań aplikacyjnych, w tym AR/VR.

2.1.3. Projekt „START: Nowatorskie metody inżynierii tkankowej wspomagające gojenie i regenerację ścięgien i więzadeł”⁴

Opracowanie narzędzi do obiektywnej i ilościowej oceny stanu ścięgna Achillesa na bazie obrazowania MRI i USG, w celu monitorowania procesów gojenia.

Opracowana metoda, oparta o AI, pozwala generować ustrukturyzowane raporty na podstawie prostej sekwencji obrazowej MRI.

Opracowanie zaawansowanych ważonych rekonstrukcji krzywych wielo-ekspozycyjnych dla mapowania MRI (np. protokołu T2MAP) z minimalizacją szumu.

Opracowanie matematycznych i obliczeniowych modeli symulacyjnych dla procesów gojenia tkanek kolagenowych o potencjalnym zastosowaniu w optymalizacji procesów gojenia.

2.1.4. Projekt „TRAFIC: Nowatorska metoda czynnościowej oceny zwężenia w tętnicy wieńcowej za pomocą modelowania przepływów in silico, na podstawie obrazu wielorzędowej tomografii komputerowej”⁵

Opracowanie metod precyzyjnej rekonstrukcji geometrii drzewa naczyniowego tętnic wieńcowych serca, wraz z informacją subrozdzielczą, z uwzględnieniem geometrii zwężeń i zwapnień. Opracowanie niskowymiarowej metody symulacji przepływu krwi w układzie naczyń wieńcowych serca, wraz z symulacją 3D przepływu w odcinku zwężonym. Opracowanie całościowej metody oceny istotności oceny czynnościowej zwężenia (FFR) na podstawie obrazu CCTA. Zakres prac z obszaru rekonstrukcji geometrii układu tętnic wieńcowych serca w obrazie CT bez- i nisko-kontrastowym.

2.1.5. Projekt „INPLA: Integracja wielomodalnych informacji dynamicznych, obrazowych i funkcjonalnych w diagnostyce i planowaniu terapii kończyny dolnej”⁶

Zakres prac w obszarze gromadzenia i analizy informacji biomechanicznej oraz modelowania chodu i biegu. Koncepcja redukcji zakresu badań oraz detekcji wzorców i nowych informacji diagnostycznych w oparciu o metody uczenia maszynowego i analizę BigData.

2.1.6. Projekt „STOPA: Modelowanie matematyczne w planowaniu zabiegów osteotomii kończyny dolnej”⁷

Opracowanie automatycznej metody pomiarów geometrii układu kostnego stopy w oparciu o dane obrazowe MRI lub CT. Opracowanie metody automatycznej oceny deformacji układu kostnego stopy i ich klasyfikacji, a także opracowanie metody doboru typu zabiegu i parametrów zabiegu osteotomii w leczeniu palucha koślawego, opartych o uczenie maszynowe. Uzgadnianie wolumetryczne wielomodalnych zbiorów danych.

2.1.7. AR w zabiegach kardiologii interwencyjnej⁸

Walidacja stosowalności rozwiązań typu rozszerzonej rzeczywistości (AR) w zabiegach kardiologii interwencyjnej, w szczególności CTO. Opracowano aplikację (wraz z modelem sterowania głosowego i gestowego) do prezentacji danych zrekonstruowanych na podstawie przedzabiegowego CT.

2.1.8. Uogólnione prawo Murray’a dla bifurkacji tętnic wieńcowych⁹

Badania w zakresie relacji między średnicami naczyń w bifurkacjach naczyń wieńcowych i związku między średnicą i długością naczyń wieńcowych. Uogólnienie prawa Murray'a i sprowadza się do prawa Finet'a przy założeniu odwrotnej proporcjonalności między średnicą naczynia a jego długością. Obserwacje na podstawie danych obrazowych CCTA dla drzew lewej tętnicy wieńcowej.

2.1.9. Analiza kształtu komórek w obrazowaniu mikroskopowym

Opracowano metody wieloskalowej oceny kształtu komórek widocznych w obrazowaniu mikroskopowym. Otrzymane wskaźniki kształtu umożliwiają klasyfikację i analizę z wykorzystaniem uczenia maszynowego.

2.1.10. Segmentacja struktur mózgowych w wodogłowie¹⁰

Opracowanie narzędzi i metod segmentacji i pomiarów geometrii struktur mózgowych w problemach wodogłowa, na bazie obrazowania medycznego.

2.1.11. Biomechanika¹¹

Opracowanie uśrednionego modelu chodu zależnego od prędkości poruszania się.

2.1.12. Rehabilitacja w udarach mózgu

Wspomaganie rehabilitacji pacjentów po przebytych udarach z wykorzystaniem modelowania matematycznego, telerehabilitacji i funkcjonalnej elektrostymulacji. Postawiono zagadnienie opracowania metod obliczeniowych związanych z tworzeniem spersonalizowanych modeli układu mięśniowo-szkieletowego pacjentów oraz przestrzeni deskryptorów opisujących dysfunkcję chodu. Działania te miały posłużyć optymalizacji wykorzystania synergii rehabilitacji w placówce medycznej oraz w domu przy wykorzystaniu nowoczesnych technologii telemedycznych.

2.1.13. Modelowanie chodu¹²

Celem projektu miało być wyznaczanie wielkości parametrów mechanicznych w stawie pod skokowym podczas chodu oraz opracowanie adekwatnego do opisanych problemów anatomicznych modelu matematycznego stopy.

2.1.14. Projekt „KARDIONET: Rozwój zaawansowanych metod obrazowania medycznego, rozproszonej akwizycji, archiwizacji i teletransmisji w zintegrowanym systemie opieki kardiologicznej ostrych zespołów wieńcowych”

2.1.15. Projekt „CePT: Centrum Badań Przedklinicznych i Technologii”
Rolą ICM w projekcie CePT było wdrożenie laboratoriów obliczeń numerycznych, składowania danych oraz wizualizacji, na potrzeby badań przedklinicznych.

2.1.16. Rekonstrukcja ruchu lewej komory serca na podstawie 4D USG¹³

2.2 Matematyczne modelowanie procesów termodynamicznych i biologicznych w skali makromolekularnej. Działalność dr Zuzanny Szymańskiej

2.2.1 Modelowanie matematyczne procesu topnienia DNA¹⁴

Autorzy: Mateusz Dębowski (MIM UW), Mirosław Lachowicz (MIM UW), Zuzanna Szymańska (ICM UW; IM PAN)

Opis badań:

W pracy zajmujemy się mikroskopowym opisem cieplnej denaturacji DNA. Zaproponowany model składa się z liniowego równania całkowo-różniczkowego opisującego stopień otwartości poszczególnych wiązań między parami zasad DNA. Opisujący proces polega na zrywaniu wiązań wodorowych pomiędzy komplementarnymi parami zasad, co powoduje rozdzielenie dwóch nici DNA. W przeciwieństwie do poprzednich prób modelowania tego zjawiska uwzględniamy stany wszystkich par zasad, co w rzeczywistości narzuca mikroskopowy charakter tego podejścia. Model jest liniowym nieautonomicznym równaniem różniczkowo-całkowym opisującym dynamikę gęstości prawdopodobieństwa, które charakteryzuje odległości między zasadami w obrębie poszczególnych nukleotydów. Nieautonomiczna struktura równania wynika z zależności od temperatury, która w ogólnym przypadku może być (daną) funkcją czasu. Według naszej wiedzy jest to pierwszy model uwzględniający nie tylko siłę podwójnych i potrójnych wiązań wodorowych między komplementarnymi zasadami, ale także oddziaływania stosowe między sąsiednimi parami zasad. W niniejszym artykule omówiono podstawowe właściwości modelu, w tym wyniki dotyczące istnienia, jednoznaczności i stabilności. Przeprowadzone symulacje numeryczne odtwarzają sigmoidalny kształt krzywych topnienia DNA i pokazują pojawienie się eksperymentalnie obserwowanych tzw. bąbli denaturacyjnych.

2.2.2. Modelowanie matematyczne roli białka CDC6 w procesie inicjacji mitozy¹⁵

Autorzy: Mateusz Dębowski (MIM UW), Zuzanna Szymańska (ICM UW; IM PAN), Jacek K. Kubiak (CNRS; WIHE), Mirosław Lachowicz (MIM UW)

Opis badań:

W pracy przedstawiamy hipotezę dotyczącą roli białka CDC6 w przechodzeniu komórek eukariotycznych do fazy mitozy w cyklu komórkowym. Nowe dane eksperymentalne sugerują, że obecność CDC6 wpływa hamująco na pojawianie się aktywnych kompleksów CDK1 i cykliny B pod koniec fazy G2 cyklu komórkowego [elDika et al. 2014]. W pracy proponujemy model matematyczny dynamiki białek: kinazy CDK1, cykliny B, fosfatazy CDC25 i inhibitora CDC6, o których wiadomo, że biorą udział w regulacji aktywnych kompleksów CDK1 i cykliny B. Formułujemy nową hipotezę, że CDC6 spowalnia aktywację kompleksów CDK1 i cykliny B po wejściu mitotycznym. Nasz model matematyczny oparty na kinetyce akcji masowej sugeruje wyjaśnienie danych eksperymentalnych. Twierdzimy, że dynamika aktywnych kompleksów CDK1 i cykliny B ma charakter podobny do dynamiki diauksycznej wprowadzonej przez Monoda [Monod 1949]. W kategoriach matematycznych stwierdzamy, że istnieje więcej niż jeden punkt przegięcia krzywej określającej dynamikę kompleksów.

2.2.3. Estymacja bayesowska nielokalnej funkcji proliferacji

Autorzy: Zuzanna Szymańska (ICM UW; IM PAN), Błażej Miasojedow (MIM UW), Piotr Gwiazda (IM PAN)

Opis badań:

Pomimo oczywistych korzyści płynących ze stosowania opisu matematycznego procesów biologicznych, większość modeli obarczonych jest istotnym ograniczeniem stosowalności polegającym na braku prawdziwego związku z danymi eksperymentalnymi. Dotyczy to w szczególności modeli związanych z opisem procesów nowotworzenia. Sposobem na pokonanie tych ograniczeń jest rozwój odpowiednich metod matematycznych umożliwiających skuteczne estymowanie parametrów modelu. W pracy proponujemy przede wszystkim nielokalny model proliferacji komórek, a następnie proponujemy Bayesowski model dla nieznanymi parametrów zaproponowanej funkcji wraz z dowodem stabilności metod numerycznych zastosowanych do rozwiązywania modelu. Wykazujemy dokładność estymatorów dla rzeczywistych danych eksperymentalnych.

2.2.4. Modelowanie matematyczne procesu przebudowy włókien kolagenowych

Autorzy: Jose Antonio Carrillo (Imperial College), Martin Parisot (Inria), Zuzanna Szymańska (ICM UW; IM PAN)

Opis badań: Urazy ścięgien stanowią bardzo poważne wyzwanie kliniczne dla współczesnej medycyny. Z reguły goją się one bardzo powoli, a chora tkanka rzadko odzyskuje strukturę i wytrzymałość mechaniczną sprzed urazu. Dodatkową trudność stanowi fakt, że ścięgno stanowi niedostatecznie zbadaną

tkankę. Jest tak dlatego, że biopsja zdrowych tkanek ścięgien jest zasadniczo niemożliwa, a prawie wszystkie istniejące dane eksperymentalne pochodzą z badań na zdrowych zwierzętach, którym mechanicznie uszkodzono ścięgna. Dopiero lepsze zrozumienie procesu gojenia pozwoli na opracowanie lepszych strategii leczenia, na przykład w ramach medycyny regeneracyjnej. Proponujemy nowy model składający się z równania różniczkowo-całkowego opisującego proces uliniawiania włókien kolagenowych. Rozważamy funkcję $f(t, x, \phi)$ opisującą gęstość włókien kolagenu, tj. prawdopodobieństwo znalezienia włókna kolagenu w chwili $t > 0$ w punkcie x o orientacji ϕ . Przez orientację ϕ rozumiemy kąt ostry między styczną do włókna w punkcie x a osią zdrowego ścięgna.

2.2.5. Modelowanie matematyczne procesu przebudowy włókien kolagenowych – the role of tendon specific fibroblasts

Autorzy: Mirosław Lachowicz (MIM UW), Henryk Leszczyński (WM UG), Zuzanna Szymańska (ICM UW; IM PAN)

Opis badań: Przedmiotem modelowania jest proces uliniawiania włókien kolagenu, który zachodzi w trakcie gojenia ścięgna. Proces ten polega głównie na zastąpieniu nieuporządkowanego kolagenu typu III uporządkowanym kolagenem typu I, w którym pośredniczą fibroblasty właściwe dla ścięgien, tzw. tenocyty. Rozważamy dwie zmienne: opisujące odpowiednio rozkład włókien kolagenu i gęstość komórek ścięgien. Dokładniej, funkcja $f(t, x, v)$ opisuje gęstość prawdopodobieństwa znalezienia włókna kolagenowego w chwili $t > 0$ w punkcie x o orientacji v . Zmienna v , którą nazywamy orientacją, opisuje kierunek włókien kolagenu w danym miejscu. Z punktu widzenia monitorowania procesu gojenia ścięgna v jest kluczową zmienną, ponieważ proces gojenia polega przede wszystkim na przywróceniu równoległej struktury włókien kolagenowych. Funkcja $c(t, x, \alpha)$ opisuje gęstość komórek ścięgien, tj. gęstość prawdopodobieństwa znalezienia komórki ścięgna w chwili $t > 0$ w punkcie x na etapie dojrzwania α . Zmienna α opisuje stopień specjalizacji/dojrzałości komórek ścięgien w danym miejscu.

2.3 Modelowanie matematyczne indukcji białka Hsp70 w warunkach szoku termicznego¹⁶. Działalność dr Grzegorza Dudziuka

Autorzy: Dudziuk Grzegorz (ICM UW), Wronowska Weronika (Wydział Biologii UW; CeNT UW), Gambin Anna (MIM UW), Szymańska Zuzanna (ICM UW; IM PAN), Rybiński Mikołaj (D-BSSE ETH Zurich)

Opis badań: Zachowanie homeostazy na poziomie komórkowym jest złożonym procesem, w który zaangażowanych jest wiele czynników. Jednym z nich jest białko Hsp70, które wspomaga „naprawę” lub degradację uszkodzonych białek, których obecność jest niekorzystna dla środowiska komórkowego. Ilość uszkodzonych białek w komórkach zwiększa się między innymi wskutek ekspozycji komórki na szok termiczny. W odpowiedzi na obecność zwiększonej

puli uszkodzonych białek komórka zwiększa syntezę Hsp70, co umożliwia obsługę negatywnych efektów szoku. Ten mechanizm regulacji syntezy Hsp70 w zależności od proteotoksycznych warunków występujących między innymi w czasie szoku termicznego jest częścią bardziej ogólnego mechanizmu reakcji komórkowej na stres (ang. *heat shock response*).

W ramach badań zaproponowaliśmy model indukcji syntezy białka Hsp70 w warunkach szoku termicznego. Model jako dane wejściowe przyjmuje temperaturę szoku i tej podstawie umożliwia obliczenie zmian w syntezie Hsp70 w czasie szoku, jak również zmian w ilości niektórych innych białek biorących udział w ścieżce indukcji Hsp70. Możliwe jest także wykorzystanie jako danych wejściowych dowolnej funkcji temperatury od czasu, dzięki czemu model umożliwia symulowanie bardzo różnorodnych scenariuszy szoku termicznego. Parametry modelu zostały skalibrowane przez dopasowanie dynamiki modelu do dostępnych w literaturze danych eksperymentalnych uzyskanych *in vitro* na komórkach HeLa. Model wymaga dalszych udoskonaleń i walidacji, jednak co ciekawe niektóre przewidywania modelu w obecnej jego postaci są zgodne z danymi eksperymentalnymi.

Wspomniany model skonstruowaliśmy rozszerzając istniejące wcześniej modele. W porównaniu do tych wcześniejszych modeli użyliśmy nieco zmodyfikowanej struktury modelu oraz zadbaliliśmy, aby parametry modelu czyniły zadość biologicznie uzasadnionym więzom.

2.4 Epidemiologia oraz geografia zdrowia. Działalność dr Anety Afelt

2.4.1. Zika&Co: Understanding the dynamics of emergence of Aedes-born diseases

French-Indonesian Project (2016-2018)

Autorzy: Roger Frutos, Tri Baskoro Satoto Tunggul, Sylvie Manguin, Aneta Afelt, Kadek Mulyawan, Sofia Mubarika

Opis badań¹⁷¹⁸: Celem projektu było powiązanie coraz lepiej poznawanych obszarów występowania chorób wirusowych przenoszonych przez komary grupy *Aedes* z warunkami środowiska przyrodniczego i antropogenicznego w warunkach klimatu tropikalnego, na przykładzie Indonezji. Zostały wytypowane obszary badawcze rozlokowane w zróżnicowanych warunkach przyrodniczych i geograficznych, przeprowadzono kampanie poboru terenowego komarów oraz larw i jaj. W tych samych lokacjach przeprowadzono kartowanie przyrodnicze warunków ekologicznych dla komarów oraz warunków socjo-ekonomicznych i struktury zagospodarowania przestrzennego.

Spodziewanym wynikiem projektu jest powiązanie warunków ekosystemowych *Aedes* z zachorowalnością ludzi na choroby przez nie przenoszone. Przy czym warunki ekosystemowe zostały zdefiniowane nie tylko jako warunki przyrodnicze, ale przede wszystkim warunki tworzone przez przekształcenie środowiska. Pierwsze wyniki projektu wskazują na relację między szczególnymi rozwiązaniami zagospodarowania terenu charakterystycznymi dla grup etnicznych

i kulturowych a zachorowaniami. Obszarem mojej odpowiedzialności merytorycznej w projekcie było prowadzenie bazy danych terenowych, laboratoryjnych, statystycznych, danych przestrzennych oraz analiza wieloczynnikowa całości materiału (**GIS and socio-economic data collection and analysis**). Rezultaty projektu jeszcze nie są w pełni opublikowane. Jednym z niepublikowanych wyników projektu jest autorskie opracowanie narzędzia do zdalnej kontroli prowadzonych badań terenowych wraz z zasilaniem bazy danych. Narzędzie zostało rozwinięte i udoskonalone w projekcie WHO ARYZA.

2.4.2. ARYZA - Assessing the Risk of emergence of Yellow fever and Zika in Asia

World Health Organization South-East Asia Region project (2017-2019)

Autorzy: Autorzy: Roger Frutos, Tri Baskoro Satoto Tunggul, Sylvie Manguin, Theeraphap Chareonviriyaphap, Aneta Afelt, Mike Banks

Opis badań: Projekt zakładał wytypowanie na podstawie istniejących rejestrów zachorowań na choroby wirusowe przenoszone przez komary poligonów badawczych o wysokim prawdopodobieństwie wystąpienia Ziki i Yellow fever. Wirusy powodujące wystąpienie zachorowań przenoszone są przez specyficzne grupy komarów. PO analizie danych epidemiologicznych, do badań wytypowano 5 poligonów (2 w Tajlandii, 3 w Indonezji). W wyniki prowadzonych badań stwierdzono występowanie Ziki w Indonezji, w żadnym z poligonów nie wykryto Yellow fever. Rezultatem projektu jest opracowanie narzędzia do zdalnej kontroli badań terenowych poboru komarów z zachowaniem pełnej georeferencyjności materiału badawczego oraz zredukowaniu możliwości błędu przypadkowego, losowego (m.in. częsty problem identyfikacji próbek materiału biologicznego, niewłaściwa lokalizacja miejsca poboru).

2.4.3. SOCIAL - Sociology Impact On Aedes-Borne Diseases Local Outbreaks

French-Indonesian Project (2018-2020)

Autorzy: Philippe Cadene, Tri Baskoro Satoto Tunggul, Sylvie Manguin, Aneta Afelt, Triwibowo Ambar Garjito, Roger Frutos,

Opis badań: Po uzyskaniu pierwszych wyników badań w projekcie Zika&Co wskazujących na silną zależność między środowiskiem antropogenicznym tworzonym przez człowieka w miastach (w szczególności w strefie suburbiów), projekt SOCIAL zakłada szczegółowe badania zależności między zachorowalnością ludzi a warunkami ich codziennego życia (miejsce zamieszkania i praca). Wytypowane zostały w Indonezji miasta-objekty badawcze, rozlokowane na 5 wyspach, wyróżniające się największą dynamiką rozwoju populacji i zmianami w zagospodarowaniu przestrzennym. Jednym z założeń projektu jest wytypowanie gospodarstw domowych do stałego monitoringu zdrowia ludności oraz populacji komarów. Zaplanowane przejście do badań miejskich i wywiadów bezpośrednich jest nową metodą realizowaną w tym

regionie. Prace są w trakcie realizacji. Zakres mojej odpowiedzialności dotyczy gromadzenia danych (w tym przestrzennych) i ich przetwarzania, w tym analiz statystycznych.

2.4.4. Analiza zmienności czasowej i przestrzennej malarii w Chinach, wielolecie

(wielolecie 2004-2020) (w trakcie realizacji)

Współpraca z Mironem Kursą, Grzegorzem Dudziukiem i Martą Hallay, ICM; po stronie Chin Shaozen Zang, Dept. of Malaria, National Institute of Parasitic Diseases. Chinese Center for Disease Control and Prevention

Opis badań^{19,20}: Chiny w 2018 r. oficjalnie zgłosiły wyeliminowanie rodzimej malarii. Jest to rezultatem narodowego programu zwalczania malarii w ramach wieloletniego programu eradykacji malarii prowadzonego przez WHO. Chiny przystąpiły do projektu w 2002 r, a od 2004 r. prowadzona jest skrupulatna baza danych dotycząca zachorowań. Okres 2004-2018 jest jednym z najlepiej na świecie udokumentowanych i możliwie kompletnych danych o zachorowaniach na malarię w skali dużej części kontynentu. Ponadto, w systemie gromadzone są na bieżąco również informacje o zachorowaniach pacjentów na malarię poza granicami ChRL, wraz ze wskazaniem regionu. Tak kompletna i długoczasowa baza danych pozwala na wykonanie szeregu analiz czasowych, przestrzennych, środowiskowych związanych z dynamiką zachorowań oraz profilem pacjentów.

2.4.5 Malaria dynamic depending on climate fluctuation, China case

(praca planowana na marzec-maj 2020: Stypendium rządu Francji, Campus France).

Planowana współpraca: Miron Kursą, Marta Hallay, Grzegorz Dudzik (ICM), Therese Libourel, Carmen Gervet (Espace-DEV)

Opis badań: Warunki życia komarów przenoszących malarię uzależnione są generalnie od dwóch czynników – (1) warunków rozrodu (regulowaną m.in. przez temperaturę atmosferyczną) oraz (2) warunków życia (regulowaną możliwością znalezienia żywicieli; ludzie, zwierzęta). Zakładając iż dostęp do żywicieli jest stały, postawione zostało pytanie: czy krótkookresowe fluktuacje klimatyczne wpływają na „okienko zachorowalności”, czyli okres aktywności komarów w ciągu roku. Przenalizowane zostaną dane o temperaturze powietrza oraz wilgotności w zestawieniu z pierwszym i ostatnim notowanym w danym roku przypadkiem zachorowalności na malarię. Dane o zachorowaniach pochodzą z bazy CDC, Chiny, zaś dane klimatyczne z chińskiego serwisu meteorologicznego.

2.4.6. Inne projekty

PREDICT, USDA – współpraca dr A. Afelt z zespołem francusko-kambodżańskim zaangażowanym w projekt detekcji wirusów przenoszonych przez nietoperze(2012-2016). Wynikiem jest 5 publikacji: 3 bezpośrednio powiązane z projektem^{21 22 23} oraz 2 przygotowane jako publikacje przeglądowe, podsumowanie wyników i wiedzy^{24 25}.

**2.5 Metody informatyczne, aplikacje i API w bioinformatyce i genomice.
Działalność Prof. Piotra Bały**

2.5.1. HEAP - Human Exposome Assessment Platform

(H2020, 2020-2024)

Projekt zajmuje się badaniem nad wpływem środowiska na zdrowie człowieka i wchodzi w skład klastra kilkunastu projektów finansowanych przez UE w tym zakresie. W ramach projektu ICM będzie rozwijał oprogramowanie do analiz genetycznych i zajmował się jego zrównolegleniem i dostosowaniem do uruchamiania na komputerach dużej mocy.

2.5.2. NIASC - Nordic Information for Action eScience Center ²⁶

(project NordForsk, 2014-2018)

Zrównoleglenie i optymalizacja aplikacji bioinformatycznych oraz ich zastosowanie do wyszukiwania w DNA człowieka sekwencji pochodzących od wirusów.

2.5.3. Aplikacje bioinformatyczne w PL-Grid ²⁷

Opracowanie systemu do składowania i przetwarzania sekwencji DNA uzyskiwanych z sekwencera Roche. System, zbudowany w oparciu o technologię UNICORE pozwalał na automatyczne przekazywanie danych z sekwencera na serwery ICM a następnie na uruchamianie na zasobach obliczeniowych ICM analiz z wykorzystaniem oprogramowania Roche. Stworzony system pozwolił na znaczące skrócenie czasu analiz DNA. Projekt był realizowany wspólnie z Zakładem Medycyny Sądowej Collegium Medicum UMK w Bydgoszczy.

2.5.4. Analiza danych medycznych w środowisku gridowym

Opracowanie systemu do analizy danych CT pozwalających na określenie kształtu i rozmiaru zatok. System zrealizowany w postaci serwisu webowego opartego o technologię UNICORE i UNICORE Portal ²⁸. Analiza korelacji pomiędzy występowaniem incydentów kardiologicznych (zawały, udary) a pogodą²⁹.

2.5.5. Genetyczne bazy danych w projekcie OCEAN (POiG)

Opracowanie, w oparciu o technologię UNICORE bezpiecznych metod składowania i analizy danych genetycznych. W ramach projektu OCEAN składowano ok 1TB danych medycznych z WUM (prof. Płoski) oraz ok 0.6TB z WCO (Wielkopolskie Centrum Onkologii). Niestety współpraca w obu przypadkach się nie rozwinęła.

2.5.6. Kardionet (Fundusze Norweskie 2008-2011)

Opracowanie nowoczesnego systemu informatycznego do składowania elektronicznego rekordu pacjenta kardiologicznego. Opracowanie metod zdalnego (za pomocą tabletów) dostępu do historii choroby dla zespołów wyjazdowych. Projekt realizowany wspólnie z WUM ³⁰.

2.6 Bioinformatyka, algorytmy bioinformatyczne oraz inna działalność Witolda Rudnickiego³¹

Udział w challenge'ach DREAM - analiza różnych rodzajów danych biomedycznych z wykorzystaniem metod uczenia maszynowego, publikacje w ramach konsorcjum DREAM³²

Analiza regulacji genetycznej przy użyciu metod uczenia maszynowego³³

Analiza własności aminokwasów ważnych z ewolucyjnego punktu widzenia na podstawie macierzy podobieństwa BLOSUM³⁴.

Równoległe implementacje podstawowego algorytmu bioinformatycznego, służącego do przyrównania sekwencji białkowych³⁵

2.7 Systemy uczenia maszynowego i selekcji w biologii oraz inna działalność i Mirona Kursy

2.6.1 Boruta¹ – system do ekstrakcji istotnych czynników

Boruta^{36 37 38}, oparty o systemy uczące się jest systemem do ekstrakcji istotnych czynników, w tym uwikłanych w złożone zależności. Boruta używany jest obecnie przez setki naukowców w wielu dziedzinach wiedzy. Może być stosowany zarówno na małych danych, na przykład klinicznych, jak i na zbiorach "omicznych" idących w miliony cech, również kombinowanych danych różnego typu.

Boruta ma mocne fundamenty teoretyczne, dobrze sprawdza się w empirycznych testach, a co ważne, jest w stanie generować negatywne odpowiedzi jeśli nie ma żadnych istotnych statystycznie zależności, podczas gdy większość konkurencyjnych algorytmów zawsze podsuwa jakieś, w tym wypadku błędne i mylące, rozwiązanie.

2.6.2 FlowForest dla cytometrii przepływowej

FlowForest^{39 40}, system uczący się dedykowany do bezpośredniej analizy danych z eksperymentów cytometrii przepływowej. Samoistnie tworzy gate'y na podstawie wejściowych danych. Oznacza populacje zdarzeń uczestniczące w przewidywaniu. *Obecnie w fazie rozwojowej, zabiegam o zbiory danych potrzebne do dalszego rozwoju (w szczególności ze starych, bezkonluzywnych badań).*

Co ciekawe, istnieją skuteczne uogólnienia na dane innego typu. Dla przykładu, mam obiecujące wyniki rozpoznawania subtelnych różnic w potencjałach wywołanych w danych elektroencefalograficznych. Co ważne, niezależne od różnic osobniczych, oraz dające się utrzymać nawet przy praktycznym braku typowego pre-processingu, co dowodzi na wysoką autonomię metody i odporność na typowe zaburzenia.

¹ Projekt prowadzony wspólnie z Witoldem Rudnickim

Dr Kursa pracuje również nad problemami fMRI i innymi elektrofizjologicznymi szeregami czasowymi (EKG, holtery). Praca w toku.

2.6.4 Fenomenologiczne analizy mechanizmów neurochemicznych u szczura, w kontekście wpływu substancji, cravingu i interakcji społecznych

Projekt^{41 42} w ścisłej współpracy z Instytutem Biologii Eksperymentalnej im. Nenckiego PAN. Metodologicznie interesują mnie zastosowania algorytmów rekonstrukcji sieci regulacyjnych w tym kontekście. W szczególności, zastosowanie teorii informacji do filtrowania, poprawnej interpretacji wymuszeń oraz modularyzacji, jak również biologiczna interpretacja miar centryczności.

2.8 Computational modelling of tissue, cell colonies and cancer growth. Działalność dr Macieja Cytowskiego²

Autorzy: Maciej Cytowski, Zuzanna Szymanska

Projekt Timothy⁴³

Opis badań: Metody eksperymentalne bardzo często nie wystarczają do zbudowania spójnej, usystematyzowanej teorii opisującej zjawiska biologiczne. Odkrywanie złożoności procesów biologicznych może wymagać wykonania dużej liczby kosztownych eksperymentów. Ponadto pełne zrozumienie wysoce nieliniowych mechanizmów rządzących tymi procesami wydaje się być niemal niemożliwe przy użyciu wyłącznie metod eksperymentalnych. Jednym z szeroko stosowanych narzędzi uzupełniających tradycyjne podejście eksperymentalne, jest modelowanie matematyczne, które, z zastosowaniem szeregu metod analitycznych i obliczeniowych, ma na celu przedstawienie matematycznego opisu procesów biologicznych.

Głównym celem projektu było stworzenie metodologii i zaawansowanych algorytmów obliczeniowych służących do opisywania wybranych zjawisk biologicznych, w nieosiągalnej dotychczas skali. W rezultacie rozwinięte zostało nowatorskie narzędzie Timothy umożliwiające przeprowadzanie symulacji układów komórkowych w skali przestrzennej sięgającej 1cm^3 (ponad 10^9 komórek). Zaproponowany model bazuje na indywidualnym opisie komórki. Oznacza to, że stan każdej komórki w rozpatrywanym układzie biologicznym opisany jest za pomocą jej indywidualnych procesów i własności. Po raz pierwszy w ramach pojedynczego modelu możliwe jest badanie i łączenie dynamiki wewnątrzkomórkowej z procesami zachodzącymi w skali makroskopowej. Zaproponowane narzędzie jest generycznym modelem umożliwiającym przeprowadzanie wielkoskalowych symulacji różnych klas procesów biologicznych, w szczególności może być przystosowane do opisu specyficznych zjawisk takich jak wzrost guza łitego w skali o znaczeniu klinicznym, rozwój biofilmu czy proces waskulogenezy^{44 45 46 47 48 49 50 51}.

² obecnie w Pawsey Centre, Perth, Australia

2.9 Działalność innych pracowników ICM

Student *Inżynierii Obliczeniowej* w ICM, Karol Niedzielewski w swojej pracy magisterskiej z 2019 roku zaimplementował bibliotekę ParalleX do zrównoleglania kodów na komputerach wielkiej mocy obliczeniowej w kontekście wykrywania cech wielowymiarowych w zbiorach genomicznych. Jego praca "Multidimensional Feature Selection and High Performance ParalleX"⁵² która ukazała się w Springer Nature Computer Science 24-go października 2019 była już prawie 900 razy ściągana z portalu wydawcy Springer Nature.

3. Działalność edukacyjna ICM

3.1 Inżynieria Obliczeniowa

Inżynieria obliczeniowa to kierunek studiów stacjonarnych drugiego stopnia o profilu praktycznym uruchomiony przez ICM UW w semestrze zimowym 2016/2017. Oferta studiów skierowana jest przede wszystkim do absolwentów studiów inżynierskich (studia pierwszego stopnia, 7-semestralne, zakończone uzyskaniem tytułu inżyniera) lub magisterskich. Studia trwają trzy semestry, kończą się uzyskaniem tytułu zawodowego magistra. Kierunek inżynieria obliczeniowa przyporządkowany jest do dyscypliny informatyka w dziedzinie nauk ścisłych i przyrodniczych.

Program studiów dotyczy zastosowania obliczeń naukowych, w tym obliczeń wykorzystujących systemy wielkoskalowe, do rozwiązywania zaawansowanych problemów naukowych i technicznych.

W trakcie studiów na kierunku inżynieria obliczeniowa studenci uzyskują podstawową wiedzę z zakresu systemów wielkoskalowych, ich architektury, a także zarządzania nimi i wykorzystania w wybranych dziedzinach. Zapoznają się z metodami przetwarzania, analizy i wizualizacji wielkich danych w tym metodami uczenia maszynowego i sztucznej inteligencji. Poznają metody i paradygmaty programowania systemów wielkoskalowych ze szczególnym uwzględnieniem programowania równoległego i rozproszonego, w tym przetwarzania w chmurze. Zdobywają wiedzę na temat wybranej dziedziny zastosowań obliczeń wielkoskalowych, wykorzystywanych algorytmów i metod obliczeniowych, a także biorą udział w projektach obliczeniowych z wykorzystaniem superkomputerów.

Jednym z istotnych obszarów zastosowań jest biologia i medycyna. Studenci mają możliwość wyboru wykładów dotyczącej wizualizacji danych wielomodalnych, w tym danych medycznych, czy wykładów dotyczących chemii obliczeniowej, modelowania molekularnego czy bioinformatyki.

Nauka na kierunku inżynieria obliczeniowa, ze względu na unikatowy charakter i niewielką liczbę studentów, jest oparta o realizację projektów z wykorzystaniem istniejącej infrastruktury obliczeniowej. Studenci mają do dyspozycji laboratorium komputerowe, korzystają z superkomputerów Cray XC40 (Okeanos), klastra

obliczeniowego (Hydra) oraz systemu analityki dużych danych (Enigma). Oprócz uczestnictwa w tradycyjnych wykładach studenci w ramach ćwiczeń biorą udział w zadaniach związanych z bieżącym utrzymaniem systemów obliczeniowych ICM. Studenci są włączani w realizację projektów badawczych i rozwojowych prowadzonych w ICM, mogą także realizować własne projekty. Duży nacisk kładziony jest na rozwój kompetencji miękkich wymaganych od absolwentów kierunków informatycznych takich jak umiejętność prezentacji, pracy grupowej, funkcjonowania w interdyscyplinarnym zespole.

Studenci inżynierii obliczeniowej biorą udział w konkursach i hackathonach zajmując czołowe lokaty.

3.2 OMICS Data Science: Bioinformatics and Analysis of large-scale biomedical data

Program Kursu Omics Data Science - Bioinformatyka i Analiza wielkoskalowych danych biomedycznych został opracowany przez radę programową i wprowadzony w 2019 jako część projektu "*Genetically conditioned diseases - education and diagnostics (EDUGEN)*" współfinansowanego ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego, Program Operacyjny Wiedza Edukacja Rozwój.

Oś priorytetowa: IV. Innowacja społeczna i współpraca ponadnarodowa Działania: 4.3 we współpracy pomiędzy ICM UW a Instytutem Matki i Dziecka w Warszawie.

Kierownikiem programu jest pracownica ICM dr Katarzyna Suski-Grabowski. Dr Katarzyna Suski-Grabowski jest absolwentką Uniwersytetu Paris-Sud, Wydział genetyki molekularnej, gdzie obroniła rozprawę doktorską na temat replikacji DNA. W latach 2003-2010 prowadziła badania w Memorial Sloan Kettering Cancer Center w Nowym Jorku a następnie w Weatherall Institute of Molecular Medicine Cancer Research Oxford University. Badania te dotyczyły strukturalnych blokad w DNA przy użyciu metod genetycznych, biochemicznych oraz analiz bioinformatycznych. W latach 2011-2017 dr Katarzyna Suski-Grabowski kierowała grupą badawczą w Instytucie Biochemii i Biofizyki Polskiej Akademii Nauk. Wraz z zespołem składającym się z studentów studiów magisterskich i studiów doktoranckich badała wpływ uszkodzenia niewiernej DNA-polymerazy na replikację DNA używając metody sekwencjonowania oraz analizując wyniki dzięki składaniu i porównywaniu sekwencji *in silico*⁵³. Obecnie dr Katarzyna Suski-Grabowski jest kierownikiem kursu bioinformatycznego Omics Data Science prowadzonego przez Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego oraz koordynatorem Flagship 1 na europejskim Uniwersytecie 4EU+.

Ponadto projekt OMICS ma na celu zbudowanie nowego środowiska naukowców, lekarzy i techników medycznych którzy będą biegli w nowych technikach obliczeniowych oraz analizy danych w nowoczesnych dziedzinach biologii systemowej, medycyny i genetyki. Program przygotowuje uczestników do zmierzenia się ze specyficznymi zagadnieniami OMICS w kontekście genetyki medycznej.

Zajęcia wprowadzające do analizy danych wysokoprzepustowych:

Działalność ICM w Naukach Biologicznych i Medycznych

- Badania wysokoprzepustowe w medycynie – wprowadzenie
- Infrastruktura w badaniach wysokoprzepustowych
- Podstawowe narzędzia informatyczne (Python)
- Podstawowe bazy danych i narzędzia do analizy danych wysokoprzepustowych
- Podstawy analizy z wykorzystaniem pakietu R
- Zastosowanie narzędzi Big Data w analizach omicznych
- Modelowanie Deep Learning w badaniach biomedycznych
- Etyczne aspekty biomedycznych badań wysokoprzepustowych

Ten interdyscyplinarny program składa się ze wszystkich nauk ilościowych OMICS (OMICS data science) i jest przygotowany dla studentów ze wszystkich dziedzin nauki. Studenci zdobywają wiedzę z analizy wysokoprzepustowej:

Praktyczne wykorzystanie analiz wysokoprzepustowych – zajęcia praktyczne poprzedzone wstępem teoretycznym

- Genomika
- Transkryptomika
- Metagenomika/ Mikrobiom
- Epigenomika
- Proteomika
- Metabolomika

3.3 *Warsaw Team* na światowych konkursach Student Cluster Competition

ICM stworzył w grudniu 2016 roku *Warsaw Team*, zespół studentów polskich którzy biorą udział w światowych konkursach Student Cluster Competition (SCC). Trzy niezależne konkursy odbywają się corocznie w Chinach (kwiecień-maj), Niemczech (czerwiec) i USA (listopad). *Warsaw Team* brał już udział w siedmiu światowych finałach i obecnie w rankingu światowym znajduje się na 23-cim miejscu, oraz 6-tym miejscu w obszarze Europy, Bliskiego Wschodu i Afryki (EMEA).

Warte podkreślenia w tym kontekście jest to, że studenci mają często za zadanie uruchomić i przyspieszyć najbardziej zaawansowane oprogramowania naukowe, często w dziedzinie obliczeń biologicznych. Na przykład w konkursie Chińskim ASC 2017 jednym z programów był *FALCON*: eksperymentalny PacBio diploid assembler, na Supercomputing 2017 (SC17) studenci przyspieszali działanie programu *MrBayes*, programu do inferencji Bayesowskiej i wyboru modelu dla szerokiego wyboru modeli filogenetycznych i ewolucyjnych. Na konkursie w Chinach ASC w 2018 roku jednym z zadań konkursowych był program *RELION* – “złoty standard” stworzony do analizy obrazów w Cryo-EM, a rok później na ASC 2019 studenci zajmowali się programem *Redbean*: A fuzzy Bruijn graph approach to long noisy reads assembly.

4. Działalność wspierająca naukę

4.1 Wirtualna Biblioteka Nauki i *otwarty dostęp*

Od kilkunastu lat ICM, wspierane finansowaniem Ministerstwa Nauki i Szkolnictwa Wyższego, zaangażowane jest w utrzymanie Wirtualnej Biblioteki Nauki. ICM odpowiedzialne jest za negocjowanie licencji krajowych dla prawie 600 instytucji akademickich i naukowych w kraju, oraz dużej liczby licencji konsorcyjnych, z udziałem od kilkunastu do kilkudziesięciu partnerów. Liczba czasopism objętych aktualnymi licencjami sięga ponad 8.000 tytułów. Liczba e-książek dostępnych w ramach WBN wynosi ok. 120.000.

W 2019 roku, po bardzo trudnych negocjacjach prowadzonych przez ICM z wydawnictwem Elsevier do krajowej licencji czasopism dołączone zostały czasopisma z grupy Cell i Lancet.

Negocjacje doprowadziły również do tego, że ok. 5.000 artykułów złożonych corocznie w czasopismach wydawnictwa Springer przez polskich autorów jest przedpłaconych z możliwością publikowania w otwartym dostępie (Open Access), a w czasopismach wydawnictwa Elsevier liczba ta wynosi 1.000 publikacji w 2020 roku i 1.500 publikacji w 2021.

4.2 Konferencje oraz promocja biologii i medycyny obliczeniowej

4.2.1 Computational Challenges in Life Science, Melbourne, Australia

Dyrektor ICM, Dr Marek Michalewicz, od roku 1994 jest aktywnie zaangażowany w promowanie obliczeń superkomputerowych w biologii i medycynie. W roku 1994 był inicjatorem i głównym organizatorem pierwszej w Australii konferencji gromadzącej naukowców zajmujących się biologią lub medycyną obliczeniową *Computational Challenges in Life Science*³ (Melbourne, 12-13 maj, 1994). Druga edycja *Computational Challenges in Life Science* odbyła się 5-7-go lutego 1996 roku w Melbourne. Prace przedstawione w drugiej edycji zostały wydane w dwóch tomach pod redakcją dr Michalewicza: *Plants to Ecosystems*⁴ oraz *Humans to proteins*⁵. Działalność ta była małym zalążkiem prowadzącym do stworzenia w 2009 roku w Melbourne Wiktoriańskiej Inicjatywy Obliczeniowych Nauk o Życiu (Victorian Life Sciences Computation Initiative) która to w 2017 roku przeobraziła się w Melbourne Bioinformatics.

4.2.2 Joint Fujitsu-A*STAR Computational Bio-Medicine Forum, Singapur

³ Programy tych konferencji znajdują się w załączniku.

⁴ [“Plants to ecosystems”, Advances in Computational Life Sciences Vol. I.](#)

Marek T. Michalewicz, Editor, CSIRO Publishing, Melbourne, Feb 1997, Pages: 144, ISBN: 0643059423;

⁵ [“Humans to proteins”, Advances in Computational Life Sciences Vol. II.,](#)

Marek T. Michalewicz, Editor, CSIRO Publishing, Melbourne, March 1998, Pages: 228, ISBN: 0643060332 (v.2)

25-go sierpnia 2009 roku dr Michalewicz zorganizował w Singapurze *Joint Fujitsu-A*STAR Computational Bio-Medicine Forum*⁶

4.2.3 Supercomputing Frontiers i Supercomputing Frontiers Europe

Dr Michalewicz zainicjował w Singapurze w 2015 roku serię konferencji pod nazwą Supercomputing Frontiers. Pierwsze trzy edycje (2015-2017) odbyły się w Singapurze, po czym od roku 2018 konferencje odbywają się w Warszawie pod nazwą Supercomputing Frontiers Europe. Wśród tematów każdej z edycji ważną pozycję zawsze zajmowała biologia i medycyna obliczeniowa.

Oto tytuły prac z biologii i medycyny obliczeniowej przedstawione na Supercomputing Frontiers w latach 2015-2019 oraz planowanych na szóstą edycję w 23-25 marca 2020:

Supercomputing Frontiers, Singapur (2015-2017)

2015

Keynote: *Exascale Challenges in Computational Genomics*, Rick Stevens, University of Chicago & Argonne National Laboratory, USA

2016

Keynote: *Could A Brain Ever Have A Mind?*, Baroness Susan Greenfield, Oxford University, United Kingdom

Invited: *The Potential for Exascale to Transform Healthcare*, Patricia Kovatch, Icahn School of Medicine, Mount Sinai Hospital, USA

2017

Enabling Large Scale Modelling of Personalised Stem Cells Therapy Through High Performance Computing, Joanna Jurek, Dominika Bakalarz & Maciej Cytowski, Ulster University, UK, Oxford University, UK, and ICM, University of Warsaw, Poland

A Clinical Pathogen Identification Pipeline Based on HPC Platform, Haoran Ma & Kenneth Ban, National University of Singapore

Using High Performance Computing to Create and, Freely Deliver, the Asian Genomic Database Necessary for Precision Medicine in This Population, Saumya Jamuar & Jonathan Picker, Global Gene Corporation, Singapore

Data Science in High Resolution Imaging of Sample Heterogeneity, Duane Loh, Colin Teo, Abhik Datta, Zhou Shen and Deepan Balakrishnan, National University of Singapore

Multi-GPU Deep Learning Techniques for Tendon Healing in Regenerative Stem Cells Based Medicine, Norbert Kapinski, Jakub Zielinski, Bartosz Borucki & Krzysztof Nowinski, ICM, University of Warsaw, Poland

⁶ Program symposium znajduje się w załączniku.

Mathematical Modelling and Large-Scale Computational Optimisation Elucidates the Link Between C. Elegans Neural Circuit Morphology and Worms Behaviour, Franciszek Rakowski, Piotr Bala, Łukasz Górski & Jan Karbowski, ICM, University of Warsaw, Poland

Supercomputing Frontiers Europe, Warszawa (2018-2019)

2018

Keynote: *Precision Medicine as an Accelerator for Next Generation Supercomputing*, Dimitri Kusnezov, Department of Energy, USA

Invited: *Advanced theoretical techniques to overcome drug-resistant bacteria*, Joanna Sułkowska, University of Warsaw, Poland

Accelerating scientific discoveries in protein bioinformatics with Microsoft Azure cloud platform, Dariusz Mrozek, Microsoft

High performance computing with coarse-grained model of biological macromolecules, Emilia A. Lubecka et al., University of Gdańsk, Poland

2019

Invited: *Studying brain structure and function through detailed computational modelling and topological analysis*, Michael Reiman, École Polytechnique Fédérale de Lausanne, Switzerland

Invited: *Computational Challenges in understanding the structure and function of the microbiome*, Tomasz Kosciółek, University of California San Diego, USA

Invited: *Sequence Similarity Search for Large-scale Metagenomic Data using Liquid Immersion Cooling Supercomputer*, Yoshinori Kimura, Infinite Curation, Japan

The effect of crowded environment on the dynamics of the hepatitis C virus protease NS3/4a, Natalia Ostrowska, Michael Feig and Joanna Trylska

Transport of vitamin B12 through the outer membrane protein BtuB in E.coli, Tomasz Pieńko and Joanna Trylska, University of Warsaw, Poland

SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction, Michał Boniecki et al., International Institute of Molecular and Cell Biology in Warsaw, Poland

Supercomputing Frontiers, Warszawa, 23-25 marzec 2020

Keynote: *Human Microbiome Project*, Rob Knight, University of California San Diego,

Invited: Machine learning-guided multiscale modeling of platelet aggregations, Yuefan Deng, Stony Brook University, NY,

Invited: *GRAPE Supercomputer and Biosystems computational research*, Makoto Taiji, RIKEN Center for Biosystems Dynamics Research,

Invited: *Functional dynamics of biomolecules with supercomputers*, Joanna Trylska, Centre of New Technologies University of Warsaw,

5. Infrastruktura

5.1 Serwerownie

ICM zarządza dwoma centrami danych w Warszawie. Oddane do użytku w 2016 roku Centrum Technologii ICM (CT-ICM) w Białoleęce ma ok. 10.000 m² powierzchni technicznej i jest jednym a najnowocześniejszych i najlepiej wyposażonych naukowych centów obliczeniowych w Polsce. Centrum zabezpieczone jest w układy ciągłego zasilania energii elektrycznej na poziomie 5.5 MW, układ chłodzenia komputerów wodą lodową i glikolem, trzy niezależne układy bezpieczeństwa przeciwpożarowego oraz układy zabezpieczenia dostępu.

Druga, starsza serwerownia znajduje się w budynku IBIB PAN przy ul. Pawińskiego. Mieści ona starszy sprzęt obliczeniowy, zasoby chmówowe oraz storage.



Figure 1. Nowe Centrum Technologii ICM na Białoleęce.

5.2 Sprzęt obliczeniowy

W CT ICM na Białogłęce znajdują się dwa superkomputery Petaflopowy Cray XC-40 (Okeanos) do tradycyjnych intensywnych obliczeń numerycznych, oraz klaster Huawei (Enigma) do analityki wielkich danych i obliczeń chmurowych. W serwerowni na Pawińskiego znajduje się klaster Huawei (Topola) przeznaczony do obliczeń mniejszej skali.



Figure 2. Jeden z najpotężniejszych komputerów w Polsce: Cray-XC40 - Okeanos w CT ICM w Białogłęce.

W ostatnim roku ICM zbudował klaster składający się z sześciu serwerów wyposażonych w najnowszą generację akceleratorów GPU (Nvidia Tesla V100) oraz unikalny komputer NEC Aurora Tsubasa z ośmioma kartami wektorowymi. W tym roku klaster zostanie rozbudowany o węzeł z dużą ilością pamięci RAM (1.5TB) i kolejnymi akceleratorami. Sumarycznie ICM posiada zasoby obliczeniowe ok. 2 PFLOPS

5.3 Przechowywanie danych

Oprócz zasobów obliczeniowych ICM posiada zasoby storage: 10 PB dostępne na komputerze Okeanos, sumarycznie 4 PB na lokalnych dyskach w węzłach Enigmy, oraz 6 PB storage obiektowego przeznaczonego do długotrwałego przechowywania cennych danych.

5.4 Sieci Komputerowe

ICM UW jest członkiem konsorcjum Pionier zrzeszającym 21 krajowych sieci

Działalność ICM w Naukach Biologicznych i Medycznych

miejskich i czterech innych akademickich centrów superkomputerowych. Ogólnopolskie partnerstwo w zakresie sieci światłowodowych umożliwiło dostęp do nowoczesnej sieci o znacznych przepustowościach i licznych połączeniach międzynarodowych.

Dwie serwerownie ICM są połączone pomiędzy sobą oraz z pozostałymi centrami obliczeniowymi i siecią Pionier łączami o przepustowości 100Gb/s. Fizyczne połączenia do użytkowników sieci oraz pomiędzy węzłami szkieletowymi wykonane są w ramach własnej sieci światłowodowej, tak aby zapewnić redundantny układ połączeń, przy zachowaniu geograficznego odseparowania tras światłowodowych.

Dzięki dwóm serwerowniom, ICM jest w stanie znacznie ograniczyć ryzyko wystąpienia awarii lub zminimalizować skutki w przypadku ich wystąpienia przez odpowiednie rozlokowanie komponentów infrastruktury obliczeniowej i sieciowej.

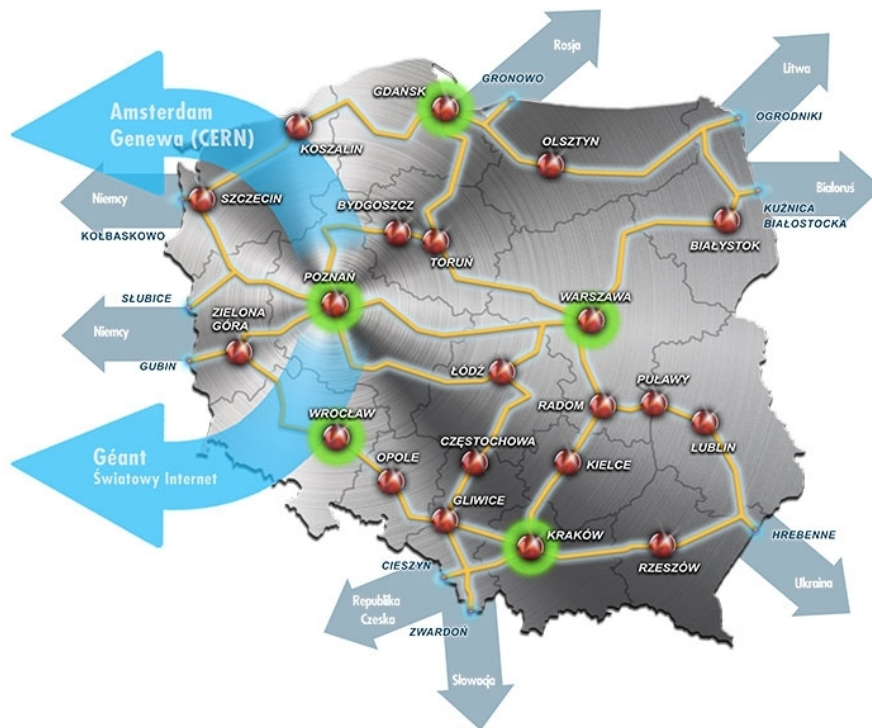


Figure 3. Ogólny zarys topologii akademickiej sieci światłowodowej PIONIER. Zielone punkty oznaczają umiejscowienie pięciu akademickich centrów superkomputerowych.

5.5 Wizualizacja

W nowozbudowanej auli wykładowej w Centrum Technologii ICM dla 70-ciu słuchaczy znajduje się unikalny sprzęt wizualizacyjny. Ściana wideo o rozdzielczości 8K i przekątnej 220" (16 monitorów) zasilana jest treścią obrazową bezpośrednio z dedykowanego serwera wyposażonego w 4 karty GPU, 48 rdzeni CPU i 1 TB pamięci operacyjnej. Oprogramowanie zapewnia możliwość interaktywnej pracy wizualnej z dużymi danymi ([m.in.](#) VisNow), a także zdalną pracę wspólną ([m.in.](#) SAGE2) czy transmisję wykładów czy obrazów z całego świata.

6. Ekspertyza

ICM jest zaangażowane w badania, eksplorację i rozwój nowych technik obliczeniowych i najnowszych rozwiązań technologicznych.

W latach 2009-2016 obecny dyrektor ICM, dr Marek Michalewicz zajmował stanowisko CEO (dyrektora naczelnego) w A*STAR Computational Resource Centre w Singapurze. Jest to największe w rejonie Azji Południowo-Wschodniej naukowe centrum obliczeniowe należące do organizacji naukowej A*STAR (Agency for Science, Technology and Research) w Singapurze. A*STAR zatrudnia ponad 5,5 tysiąca pracowników w 14 instytutach naukowych. Siedem z tych instytutów należy do tzw. BMRC: Bio-Medical Research Council. Dr Michalewicz był odpowiedzialny za stworzenie Narodowego Centrum Superkomputerowego w Singapurze (NSCC). NSCC było inwestycją rządu 100 mln dolarów singapurskich i rozpoczęło swoją działalność na początku 2015 roku. Architektura rozwiązań zaproponowanych przez dr Michalewicza była w szczególności skierowana dla użytkowników Singapurskiego Instytutu Genomiki (Genome Institute of Singapore - GIS) i innych Instytutów BMRC. Wśród unikalnych rozwiązań stworzonych i wprowadzonych pod kierownictwem dr Marka Michalewicza w Singapurze należy wymienić:

1. Sieci komunikacyjne przesyłania danych:

Zaprojektowanie unikalnej w skali światowej architektury połączeń dwóch kampusów A*STAR oddległych o ok. 1,2 km: Biopolis, gdzie mieści się Genome Institute of Singapore (GIS) i Fusionopolis 1 (siedziba NSCC). Całkowita przepustowość sieci jest bez precedensu światowego i wynosiła w 2016 roku 1.18 Tbps.

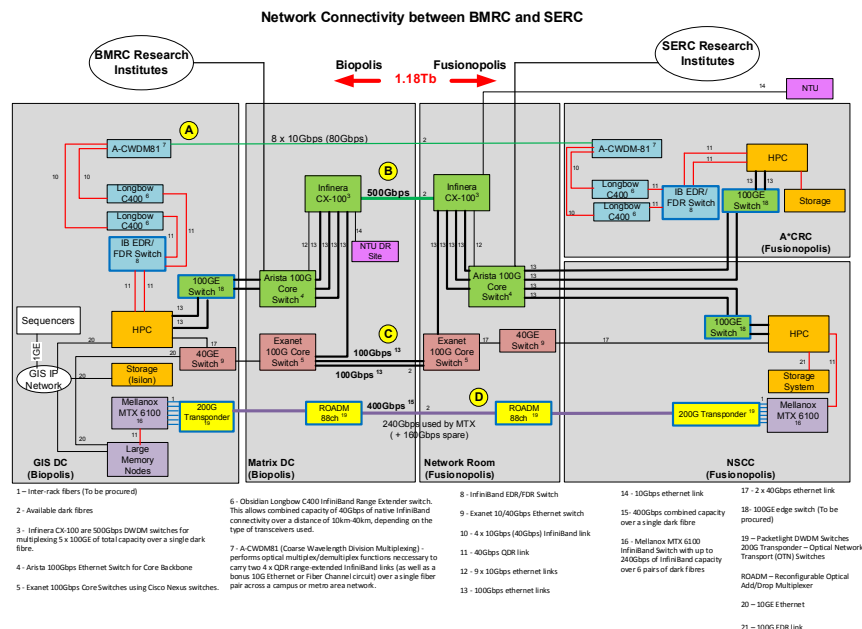


Figure 4. Połączenie sieciowe pomiędzy Biopolis i Fusionopolis w Singapurze. 1,18 Tbps zrealizowane poprzez pięć połączeń, w tym 500 Gbps CloudXpress-1 (Infinera, tcp/ip) oraz bezpośrednie połączenie InfiniBand (MetroX, Mellanox).

Działalność ICM w Naukach Biologicznych i Medycznych

Jednym z ciekawszych zastosowań połączenia pomiędzy GIS a NSCC w Singapurze było bezpośrednie połączenie sekwenatorów Illumina ze storagem NSCC odległego o 1,2 km używając protokołu InfiniBand.

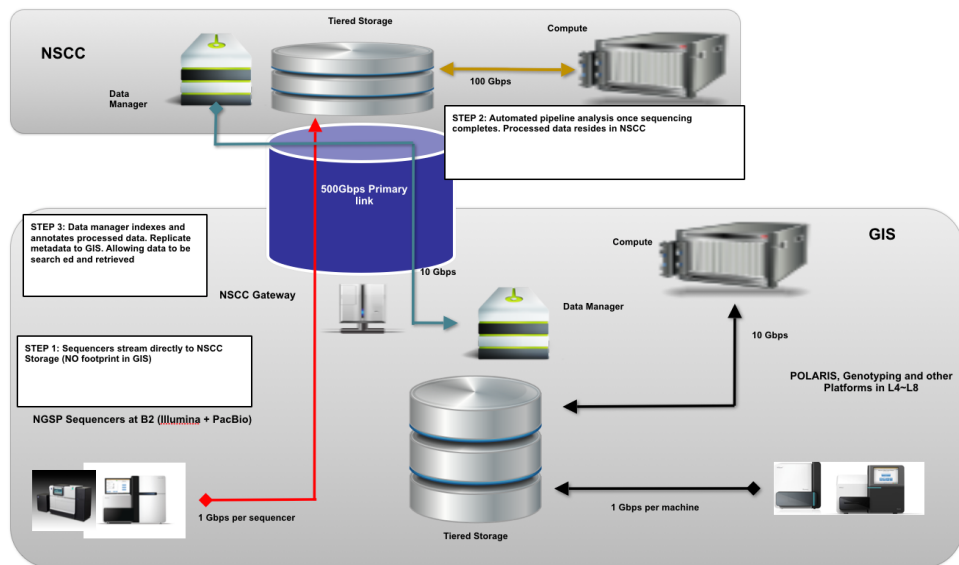


Figure 5. Sekwenator Illumina w GIS przekazuje bezpośrednio dane doświadczelne do storageu NSCC i dalej do analizy na superkomputerze - bez żadnych kroków pośrednich

2. InfiniCortex

W latach 2014-2016 dr Michalewicz był inicjatorem i liderem projektu InfiniCortex – stworzenia współbieżnego komputera o globalnym zasięgu, mieszczącego się na czterech kontynentach i w siedmiu krajach (Singapur, Australia, Japonia, Kanada, USA, Francja, Polska) połączonego globalną siecią InfiniBand z przepustowością 100gbps.

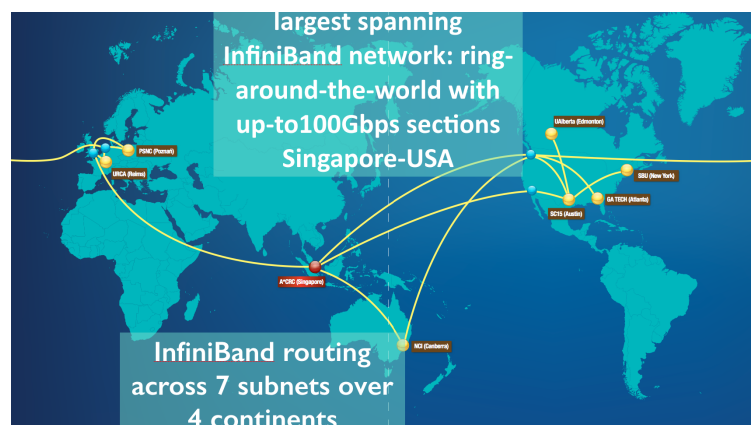


Figure 6. InfiniCortex w 2015, prezentowany na konferencji Supercomputing 15 w Nowym Orleanie. Globalne połączenie InfiniBand z przepustowością 100gbps na czterech kontynentach. Po raz pierwszy pokazano również tworzenie pod-sieci InfiniBandu o globalnym zasięgu.

Działalność ICM w Naukach Biologicznych i Medycznych

Osiągnięcie to jest bezprecedensowe i nie powtórzone do dzisiaj. InfiniCortex to nie tylko wyjątkowa infrastruktura służąca do obliczeń rozproszonych i do przesyłania danych o zasięgu światowym, lecz również platforma na której zademonstrowano wiele zastosowań biologicznych i medycznych:

- i. przesyłanie danych genomicznych pomiędzy Australią i Singapurem, via Seattle. Zbiór genomiczny o wadze 1,14 Tb został przesłany używając połączenia InfiniCortex w czasie 24 minut. Normalny transfer przy użyciu standardowego połączenia tcp/ip wymaga 12,5 godziny.

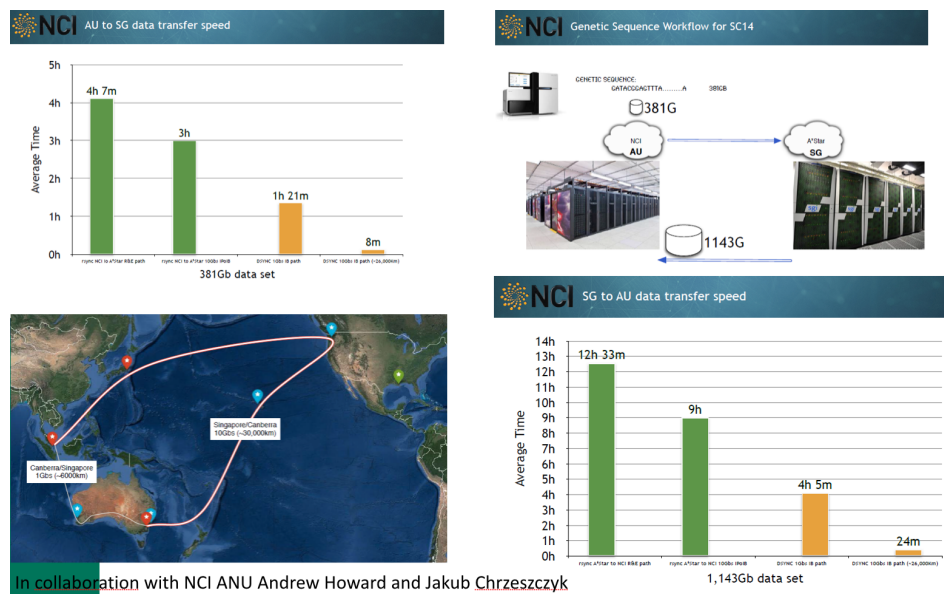


Figure 7. Transfer danych genomicznych pomiędzy National Computing Infrastructure na Australijskim Uniwersytecie Narodowym (NCI ANU) w Kanberze a Singapurem, przy użyciu InfiniCortex'a (protokół InfiniBand na odległości globalne).

- ii. InfiniCloud.

Na podłożu infrastruktury InfiniCortex'a został stworzony globalny wirtualny komputer *InfiniCloud*, który był konfigurowany na przyciśnięcie klawisza i znajdował się w czterech miejscach na świecie: Australii, Singapurze, Francji i USA. InfiniCloud pposłużył do zademonstrowania zastosowania z genomiki raka: "cancer mutation calling pipeline".

Działalność ICM w Naukach Biologicznych i Medycznych

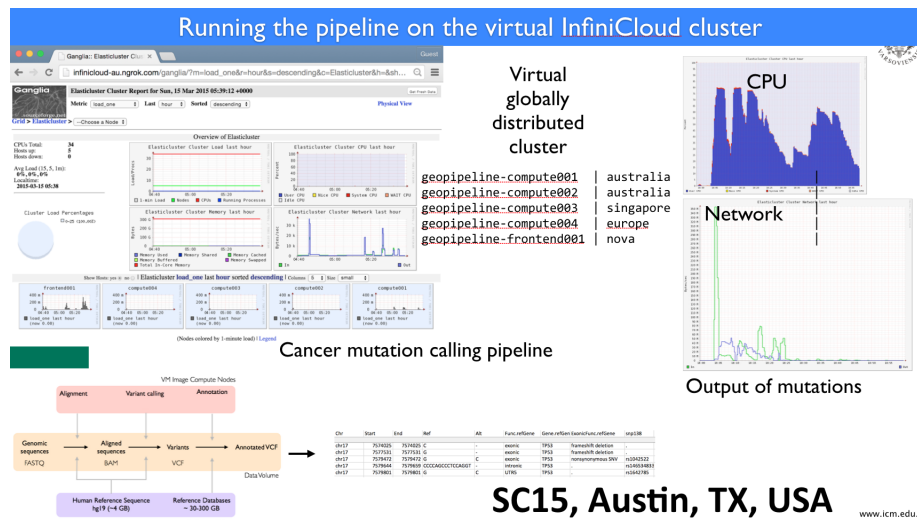


Figure 8. InfiniCloud - globalny wirtualny komputer zastosowany do obliczeń w genomice mutacji raka.

iii. Segmentacja “w locie” dużych obrazów patologicznych tkanek (Streaming Segmentation of Large Pathology Tissue Images). Projekt ten został zrealizowany z partnerami ze Stony Brook University, Oak Ridge National Laboratory i Georgia Institute of Technology (GT). Olbrzymie obrazy patologiczne tkanek w skali do 120.000 x 120.000 pikseli przechowywane w Singapurze były w czasie rzeczywistym przesyłane używając połączenia InfiniCortex do GT do analizy. Wyniki z GT przekazywane były do tabletu (czytaj: lekarze przy pacjencie) w Nowym Orleanie.

Streaming Segmentation of Large Pathology Tissue Images

Abstract:

Demo showed nuclear segmentation on a high resolution whole slide tissue color image (53,000x36,000 pixels, RGB) stored on a cluster at ACRC in Singapore.

The image is partitioned into tiles. The tiles that contain enough tissue data are streamed from multiple cluster nodes via the ORNL ADIOS system over Longbow connections to a cluster at Georgia Tech (GT) in the US. The tiles are processed on the GT cluster using the SBU RT middleware as they are received to segment nuclei. Segmented nuclei and the tiles are assembled into an image.

Team:

Tahsin Kurc, SBU and ORNL Scott Klasky, ORNL Jong Choi, ORNL Joel Saltz, SBU

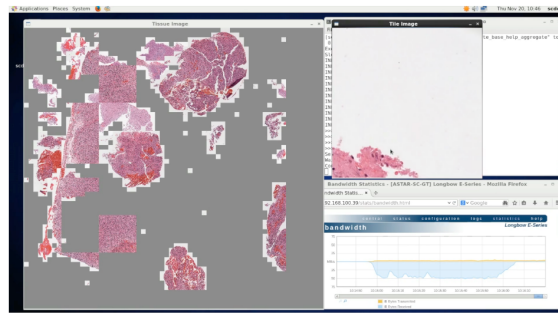


Figure 9. Segmentacja obrazów patologii tkanek "w locie" na odległościach transkontynentalnych.

3. Architektura superkomputera Aspire 1 oraz zasobów do przechowywania danych naukowych (Storage) w NSCC w Singapurze zaprojektowana przez dr Michalewicza i współpracowników została stworzona w dużej mierze aby zaspokoić potrzeby Genomic Institute of Singapore i innych instytutów bio-medycznych A*STAR. Unikalne cechy tego rozwiązania to:

- bezpośrednie połączenie InfiniBand between GIS a NSCC

Działalność ICM w Naukach Biologicznych i Medycznych

- ii. bezprecedensowo szybki I/O (Input/Output) pomiędzy superkomputerem a storage'em: 500 Gbps używając 10 x Infinite Memory Engine jednostek tzw. burst buffer.
- iii. dziesięć serwerów z bardzo dużą pamięcią współdzieloną przeznaczonych specjalnie do zadań genomiki i biologii w czasie rzeczywistym (poza systemem kolejkowym). Jeden serwer z 6TB pamięci RAM, 4x z 2TB i 5x 1TB.
- iv. Hierarchiczne zarządzanie storage'em (HSM) i specjalne rozwiązania do zarządzania rozproszonymi meta-danymi (Arcitecta).
- v. Specjalne partycje w storage'u dla danych genomicznych.
- vi. Wydzielona część storage'u o wielkości 500TB z zabezpieczeniem anonimowości (secure storage) specjalnie dla wrażliwych danych genomicznych i medycznych do długoterminowego przechowywania.
- vii. Bezpośrednie nody loginowe połączone InfiniBandem z NSCC w trzech uniwersytetach (NUS, NTU, SUTD) a w przyszłości w kampusach wielu szpitali Singapuru – jeszcze jedno rozwiązanie unikalne w skali światowej.

ICM zaangażowany jest od kilku lat w wyjątkowe projekty technologiczne stawiające nas w lidze liderów zastosowań rozwiązań sieciowych i obliczeniowych w skali światowej. Oto kilka przykładów:

1. Połączenie dwóch centrów danych ICM (Ochota – Białółka) na odległości ok. 20 km z przepustowością 1.2 Tbps używając najnowocześniejszego sprzętu demonstracyjnego CloudXpress-2 firmy Infinera (ICM jako pierwsze w Europie, tuż po Amazon, Facebook i Google).
2. Data Transfer Nodes (DTN) – w połączeniu z transferem danych i obliczeniami w skali globalnej. Demonstracja na Supercomputing 2018 w USA połączenia z Pawsey Centre w Perth, Australia i uruchomienia skonteneryzowanych programów alternatywnie na zasobach w Australii lub w Warszawie.
3. Ustanowienie połączeń InfiniBand pomiędzy ICM a TASK Gdańsk, oraz ICM a NCBJ, Świerk i zbudowanie daleko rozproszonych współbieżnych komputerów rozproszonych
4. Transfer danych pomiędzy Warszawą a Singapurem na nowym połączeniu 100Gbps CEA-1 (Collaboration Asia Europe 1) z szybkością 100Gbps we współpracy z firmą zettar.

W 2020 roku ICM, w ramach programu Zintegrowanego Programu Rozwoju, będzie gościł profesora wizytującego Yuefan'a Deng'a ze Stony Brook University, USA. Prof Yuefan Deng od lat prowadzi wieloskalowe badania obliczeniowe przepływu krwi z uwzględnieniem dynamiki płytek krwi i krwinek. Prof Deng jest również założycielem i jednym z dyrektorów nowopowstałego [Institute for Engineering Driven Medicine](#) w Stony Brook University (razem z Prof. Joel Saltz).

7. Podsumowanie

Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego od 27-miu lat zaangażowane jest w różnorodne badania związane ściśle z biologią i medycyną. W ciągu ostatnich lat nasiliła się również działalność wprowadzania nowych technologii obliczeniowych, nowych algorytmów, w szczególności sztucznej inteligencji i uczenia maszynowego (AI & ML), nowoczesnych metod przesyłania olbrzymich danych naukowych, przechowywania danych i ich analizy przy użyciu różnorodnych metod oraz nowych architektur takich jak karty graficzne GPU czy procesory wektorowe.

ICM rozwija współpracę naukową i technologiczną z wieloma partnerami naukowymi: uniwersytetami, centrami badań i obliczeń wielkiej skali, ze szpitalami badawczymi i organizacjami rządowymi zajmującymi się zdrowiem lub medycyną na całym świecie.

Obecnie ICM pracuje nad zacieśnieniem wewnętrznej współpracy z Centrum Nowych Technologii UW, oraz innymi jednostkami UW; z Politechniką Warszawską; Instytutem Matki i Dziecka oraz szpitalami Warszawskimi.

ICM zaprasza Warszawski Uniwersytet Medyczny do rozmów na temat możliwej współpracy naukowej, wymiany doświadczeń w zakresie planowania, projektowania architektur i rozwiązań IT, realizacji przetargów i zakupów sprzętu oraz utrzymania infrastruktury informatycznej i obliczeniowej.

W tym celu proponujemy organizację wspólnych, jednodniowych warsztatów, które służyłyby przedstawieniu przez obie strony naszych osiągnięć i aspiracji, oraz wyszukaniu możliwych obszarów współpracy lub wymiany doświadczeń.

Proponujemy aby warsztaty takie odbyły się w naszym nowopowstałym audytorium w Centrum Technologii ICM przy ul. Kupieckiej 32 w Białolece w pierwszej połowie 2020. ICM podjąby się organizacji tego wydarzenia oraz zaprosiłyby uczestników do zwiedzenia naszej serwerowni i całego zaplecza technologicznego.

8. Publikacje

¹ B.A.Borucki, Ł.Bolikowski, "Quo vadis, CARS? First steps towards text-mining-based analysis of topics in the International Journal of Computer Assisted Radiology and Surgery", Int. J. of Computer Assisted Radiology and Surgery, Vol.10, Supp.1, June 2015, pp 222-223

P.Regulski, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Assessment of anisotropic denoiser enhanced cone beam CT for patient dose reduction", Int. J. of Computer Assisted Radiology and Surgery, Vol.10, Supp.1, June 2015, pp 295-296

M.M.Rahman, P.K.M.Kumar, B.Borucki, K.S.Nowiński, M.S.Uddin, "Speckle noise reduction of ultrasound images using Extra-Energy Reduction function", Proc. of Int. Conf. on Informatics Electronics Vision ICIEV 2013, IEEE, 2013, pp 1-6

Gąsior JS1, Sacha J, Pawłowski M, Zieliński J, Jeleń P, Tomik A, Książczyk TM, Werner B, Dąbrowski. Normative Values for Heart Rate Variability Parameters in School-Aged Children: Simple Approach Considering Differences in Average Heart Rate. Front Physiol. 2018 Oct 24;9:1495.

Przemysław Jerzy Kwasiborski, Paweł Kowalczyk, Jakub Zieliński, Jacek Przybylski, Andrzej Cwetsch, Znaczenie powinowactwa hemoglobiny do tlenu w adaptacji do hipoksemii. Polski Merkuriusz Lekarski 2010, 28 (166): 260-264

Anna M. Badowska-Kozakiewicz, Elżbieta Malicka, Anna Rodo, Jakub Zieliński, Rola komórek nacieku zapalnego w nowotworach gruczołu sutkowego suk. Życie Weterenaryjne, 84 (12): 986 - 990

Jakub Zieliński, Czy wykonywanie zawodu lekarza znacząco skraca życie? Przegląd Lekarski 2010, 67 (12): 1340-1341

Jakub Zieliński, Jacek Przybylski, Ile wody tracimy z oddechem?, Pneumonologia i Alergologia Polska 2012; 80 (3): 339–342

T. Lewandowski, E. Głodkowska-Mrówka, R. Biernacka, K. Roszkowska-Purska, M. Machnicki, J. Zieliński, P. Włodarski, T. Stokłosa, "EGFR expression but not mutation of EGFR, KRAS, BRAF, PIK3CA and MSI is prognostic factor in gastric cancer after surgical resection", European Journal of Cancer 2013; 49 (Suppl 2): S639-S639

Jakub S. Gąsior, Jerzy Sacha, Piotr J. Jeleń, Jakub Zieliński, Jacek Przybylski, Heart Rate and Respiratory Rate Influence on Heart Rate Variability Repeatability: Effects of the Correction for the Prevailing Heart Rate, Frontiers in Physiology 2016; 7: 356

² N.Kapiński, J.M.Nowosielski, M.E.Marchwiany, J.Zieliński, B.Ciszkowska-Łysoń, B.A.Borucki, T.Trzeciński, K.S.Nowiński "Late fusion of deep learning and hand-crafted features for Achilles tendon healing monitoring", Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, 2019

P.Woźnicki, P.Przybyszewski, N.Kapiński, J.Zieliński, B.Ciszkowska-Łysoń, B.A.Borucki, T.Trzeciński, K.S.Nowiński "Monitoring Achilles tendon healing progress in ultrasound

imaging with convolutional neural networks", Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, 2019

N.Kapiński, J.Zieliński, B.A.Borucki, T.Trzeciński, B.Ciszkowska-Łysoń, U.Zdanowicz, R.Śmigielski, K.S.Nowiński "Monitoring of the Achilles tendon healing process: can artificial intelligence be helpful?", Acta Bioeng Biomech., Vol.21, No.1, January 2019, pp 103-111

N.Kapiński, J.Zieliński, B.A.Borucki, T.Trzeciński, B.Ciszkowska-Łysoń, K.S.Nowiński "Estimating Achilles tendon healing progress with convolutional neural networks", in A.F.Fangi et.al. eds., Medical Image Computing and Computer Assisted Intervention - MICCAI 2018, 2018, pp 949-957

N.Kapiński, J.Zieliński, B.A.Borucki, "AchillesDL: Convolutional Neural Nets Help to Evaluate Tendons Regeneration Process in Sport Medicine", Nvidia GPU Technology Conference, Silicon Valley, March 2018

N.Kapiński, J.Zieliński, B.A.Borucki, K.S.Nowiński, "MRI-based deep learning for in-situ monitoring of Achilles tendon regeneration process", Int. J. of Computer Assisted Radiology and Surgery, Vol.12, Supp.1, June 2017, pp 57-58

N.Kapiński, J.Zieliński, B.A.Borucki, "Assessing the Applicability of Deep Learning Techniques in Computer Assisted Radiology for Diagnosing Tendons", Nvidia GPU Technology Conference, Silicon Valley, May 2017

N.Kapiński, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Multi-GPU Deep Learning Techniques for Tendon Healing in Regenerative Stem Cells based Medicine", Supercomputing Frontiers Conference, Singapore, March 2017

³ M.P.Opolski, I.M.Michałowska, B.A.Borucki, B.Nicińska, Ł.Szumowski, M.Sterliński "Augmented-reality computed tomography-guided transcatheter pacemaker implantation in dextrocardia and congenitally corrected transposition of great arteries", Cardiology Journal, Vol.25 (3), 2018, pp 412-413

M.P.Opolski, A.Dębski, B.A.Borucki, A.D.Staruch, C.Kęпка, J.K.Rokicki, B.Sieradzki, A.Witkowski "Feasibility and safety of augmented-reality glass for computed tomography-assisted percutaneous revascularization of coronary chronic total occlusion: A single center prospective pilot study", Journal of Cardiovascular Computed Tomography, Vol.11(6), 2017, pp 489-496

M.P.Opolski, A.Dębski, B.A.Borucki, M.Szpak, A.D.Staruch, C.Kęпка, A.Witkowski, "First-in-Man Computed Tomography-Guided Percutaneous Revascularization of Coronary Chronic Total Occlusion Using a Wearable Computer: Proof of Concept", Canadian Journal of Cardiology, Vol.32(6), June 2016, pp 829.e11-829.e13

N.Kapiński, B.A.Borucki, K.S.Nowiński, "Automatic Parameterization of OpenSim Models and Biomechanical Data Integration in Open Source Visualization Platform VisNow", 21st Congress of the European Society of Biomechanics, July 2015

⁴ N.Kapiński, J.M.Nowosielski, M.E.Marchwiany, J.Zieliński, B.Ciszkowska-Łysoń, B.A.Borucki, T.Trzeciński, K.S.Nowiński "Late fusion of deep learning and hand-crafted features for Achilles tendon healing monitoring", Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, 2019

- P.Woźnicki, P.Przybyszewski, N.Kapiński, J.Zieliński, B.Ciszkowska-Łysoń, B.A.Borucki, T.Trzciński, K.S.Nowiński "Monitoring Achilles tendon healing progress in ultrasound imaging with convolutional neural networks", Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, 2019
- N.Kapiński, J.Zieliński, B.A.Borucki, T.Trzciński, B.Ciszkowska-Łysoń, U.Zdanowicz, R.Śmigielski, K.S.Nowiński "Monitoring of the Achilles tendon healing process: can artificial intelligence be helpful?", Acta Bioeng Biomech., Vol.21, No.1, January 2019, pp 103-111
- N.Kapiński, J.Zieliński, B.A.Borucki, T.Trzciński, B.Ciszkowska-Łysoń, K.S.Nowiński "Estimating Achilles tendon healing progress with convolutional neural networks", in A.F.Fangi et.al. eds., Medical Image Computing and Computer Assisted Intervention - MICCAI 2018, 2018, pp 949-957
- N.Kapiński, J.Zieliński, B.A.Borucki, "AchillesDL: Convolutional Neural Nets Help to Evaluate Tendons Regeneration Process in Sport Medicine", Nvidia GPU Technology Conference, Silicon Valley, March 2018
- J.Nowosielski, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Multidimensional Haralick's feature space analysis for assessment of the Achilles tendon in MR imaging", Int. J. of Computer Assisted Radiology and Surgery, Vol.12, Supp.1, June 2017, pp 218-220
- P.Regulski, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Comparison of noise reducing T2map reconstruction methods in MRI imaging of Achilles tendon", Int. J. of Computer Assisted Radiology and Surgery, Vol.12, Supp.1, June 2017, pp 15-16
- N.Kapiński, J.Zieliński, B.A.Borucki, K.S.Nowiński, "MRI-based deep learning for in-situ monitoring of Achilles tendon regeneration process", Int. J. of Computer Assisted Radiology and Surgery, Vol.12, Supp.1, June 2017, pp 57-58
- N.Kapiński, J.Zieliński, B.A.Borucki, "Assessing the Applicability of Deep Learning Techniques in Computer Assisted Radiology for Diagnosing Tendons", Nvidia GPU Technology Conference, Silicon Valley, May 2017
- N.Kapiński, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Multi-GPU Deep Learning Techniques for Tendon Healing in Regenerative Stem Cells based Medicine", Supercomputing Frontiers Conference, Singapore, March 2017
- J.Zieliński, P.Regulski, N.Kapiński, B.Ciszkowska-Łysoń, B.A.Borucki, K.S.Nowiński, "Assessment of the Achilles tendon by Haralick's texture features of MRI images preliminary study", Int. J. of Computer Assisted Radiology and Surgery, Vol.11, Supp.1, June 2016, pp 194-195
- ⁵ P.Regulski, J.Zieliński, B.A.Borucki, K.S.Nowiński, "Assessment of anisotropic denoiser enhanced cone beam CT for patient dose reduction", Int. J. of Computer Assisted Radiology and Surgery, Vol.10, Supp.1, June 2015, pp 295-296
- Zieliński, J. & Nowosielski, J.M. Finet's Law as a Special Case of the Generalised Murray's Law SN COMPUT. SCI. (2020) 1: 42.
- Nowiński K.S., Zieliński J., (2014). Multi-step anisotropic denoiser scheme applied for cardiac non-contrast CT images. International Journal of Computer Assisted Radiology and Surgery. 9. S263-264

⁶ N.Kapiński, B.A.Borucki, K.S.Nowiński, "Automatic Parameterization of OpenSim Models and Biomechanical Data Integration in Open Source Visualization Platform VisNow", 21st Congress of the European Society of Biomechanics, July 2015

N.Kapiński, B.Borucki, K.S.Nowiński, "Error assessment and minimization in 4D motion tracking for functional orthopaedics diagnostics", *Int. J. of Computer Assisted Radiology and Surgery*, Vol.8, Supp.1, June 2013, pp 157-159

⁷ B.Borucki, K.S.Nowiński, P.Adamczyk, J.M.Laskowski, "Automatic classification of hallux valgus deformations with the use of automatic evaluation of geometric descriptors", *Int. J. of Computer Assisted Radiology and Surgery*, Proceedings of CARS 2012, Vol.7, Supp.1, June 2012, pp 196-198

M.Wychowański, K.S. Nowiński, B.Borucki, A.Wit, P.Adamczyk, J.M.Laskowski, "Wybrane problemy metrologiczne oceny morfologicznej i funkcjonalnej stopy w leczeniu palucha kołowego", *PAK*, 2012, Vol.1, pp 48-51

B.Borucki, K.S.Nowiński, M.Chlebiej, A.Rutkowski, P.Adamczyk, J.M.Laskowski, "Automated geometric features evaluation method for normal foot skeleton model", *Annales UMCS Informatica*, Vol.4, 2011, pp 61-75

⁸ M.P.Opolski, I.M.Michałowska, B.A.Borucki, B.Nicińska, Ł.Szumowski, M.Sterliński "Augmented-reality computed tomography-guided transcatheter pacemaker implantation in dextrocardia and congenitally corrected transposition of great arteries", *Cardiology Journal*, Vol.25 (3), 2018, pp 412-413

M.P.Opolski, A.Dębski, B.A.Borucki, A.D.Staruch, C.Kępka, J.K.Rokicki, B.Sieradzki, A.Witkowski "Feasibility and safety of augmented-reality glass for computed tomography-assisted percutaneous revascularization of coronary chronic total occlusion: A single center prospective pilot study", *Journal of Cardiovascular Computed Tomography*, Vol.11(6), 2017, pp 489-496

M.P.Opolski, A.Dębski, B.A.Borucki, M.Szpak, A.D.Staruch, C.Kępka, A.Witkowski, "First-in-Man Computed Tomography-Guided Percutaneous Revascularization of Coronary Chronic Total Occlusion Using a Wearable Computer: Proof of Concept", *Canadian Journal of Cardiology*, Vol.32(6), June 2016, pp 829.e11-829.e13

⁹ Zieliński, J. & Nowosielski, J.M. Finet's Law as a Special Case of the Generalised Murray's Law *SN COMPUT. SCI.* (2020) 1: 42

¹⁰ Czerwosz, L., Szczepek E., Nowiński K.S., Sokołowska B., Jurkiewicz J., Czernicki Z., Koszewski W., (2017). Discriminant Analysis of Intracranial Volumetric Variables in Patients with Normal Pressure Hydrocephalus and Brain Atrophy. 10.1007/5584_2017_75.

Szczepek E., Czerwosz L., Nowiński K.S., Dmowska-Pycka A., Czernicki Z., Jurkiewicz J., (2015). The usefulness of the evaluation of volumetric and posturographic parameters in the differential diagnosis of hydrocephalus. *Wiadomości lekarskie (Warsaw, Poland : 1960)*. 68. 145-52.

Szczepek E., Czerwosz L., Nowiński K.S., Czernicki Z., Jurkiewicz J., (2015). Original article Analysis of intracranial volume ratios by means of cerebrospinal fluid deployment indicators. *Folia neuropathologica / Association of Polish Neuropathologists and Medical*

Research Centre, Polish Academy of Sciences. 53. 121-7. 10.5114/fn.2015.52408.

Szczepek E., Czerwosz L., Nowiński K.S., Jurkiewicz J., Czernicki Z., (2014). Evaluation of Volumetric Changes in Differential Diagnosis of Brain Atrophy and Active Hydrocephalus. *Advances in experimental medicine and biology*. 840. 10.1007/5584_2014_85

¹¹ N.Kapiński, B.A.Borucki, K.S.Nowiński, "Automatic Parameterization of OpenSim Models and Biomechanical Data Integration in Open Source Visualization Platform VisNow", 21st Congress of the European Society of Biomechanics, July 2015

N.Kapiński, B.Borucki, K.S.Nowiński, "Error assessment and minimization in 4D motion tracking for functional orthopaedics diagnostics", *Int. J. of Computer Assisted Radiology and Surgery*, Vol.8, Supp.1, June 2013, pp 157-159

¹² N.Kapiński, B.A.Borucki, K.S.Nowiński, "Automatic Parameterization of OpenSim Models and Biomechanical Data Integration in Open Source Visualization Platform VisNow", 21st Congress of the European Society of Biomechanics, July 2015

N.Kapiński, B.Borucki, K.S.Nowiński, "Error assessment and minimization in 4D motion tracking for functional orthopaedics diagnostics", *Int. J. of Computer Assisted Radiology and Surgery*, Vol.8, Supp.1, June 2013, pp 157-159

¹³ Chlebiej M., Nowiński K.S., Scislo P., Bała P., (2008). Development of heart motion reconstruction framework based on the 4D echocardiographic data. *Annales UMCS, Informatica*. 8. 43-50. 10.2478/v10065-008-0023-3

¹⁴ Mateusz Dębowski, Mirosław Lachowicz, Zuzanna Szymańska, *Microscopic description of DNA thermal denaturation*, *Appl. Math. Comput.* 361 (2019) 47-60 [doi: 10.1016/j.amc.2019.05.001](https://doi.org/10.1016/j.amc.2019.05.001)

¹⁵ Mateusz Dębowski, Zuzanna Szymańska, Jacek Z. Kubiak i Mirosław Lachowicz, *Mathematical model explaining the role of CDC6 in the diauxic growth of CDK1 activity during the M-phase of the cell cycle*, *Cells* 8 (2019) 1537 [doi: 10.3390/cells8121537](https://doi.org/10.3390/cells8121537)

¹⁶ Dudziuk G., Wronowska W., Gambin A., Szymańska Z., Rybiński M., „Biologically sound formal model of Hsp70 heat induction”, *J. Theor. Biol.* 478 (7), 74-101, 2019.

¹⁷ Triwibowo Ambar Garjito, Widiarti, Yusnita Mirna Anggraeni, Sitti Alfiah, Tri Baskoro Tunggul Satoto, Achmad Farchanny, Gina Samaan, **Aneta Afelt**, Sylvie Manguin, Roger Frutos, Tjandra Yoga Aditama, 2018, Japanese encephalitis in Indonesia: An update on epidemiology and transmission ecology, *Acta Tropica* 187, pp.240-247

¹⁸ Pham Thi Kim Lien, Laurence Briant, Laurent Gavotte, Pierrick Labbe, Marco Perriat-Sanguinet, Emmanuel Cornillot, Vu Trong Duo, Nguyen Thi Ye, Tran Vu Phong, Nguyen Van Soai, Christian Devaux, **Aneta Afelt**, Tran Chi Cuong, Phan Thi Nga, Tran Nhu Duong, Roger Frutos. 2017. Incidence of dengue and chikungunya virus in mosquitoes and human patients in border provinces of Vietnam, *Parasites & Vectors* 10(1)

¹⁹ Shaosen Zhang, Shaohua Guo, Xinyu Feng, **Aneta Afelt**, Roger Frutos, Shuisen Zhou, Sylvie Manguin. 2017. Anopheles vectors in Mainland China while approaching malaria elimination, *Trends in Parasitology*, 33(11).

- ²⁰ Shaosen Zhang; Zhengbin Zhou; Tianmu Chen; Xuezhong Wang; Wenqi Shi; Weikang Jiang; Julin Li; Xiaonong Zhou; Roger Frutos; Sylvie Manguin; Shuisen Zhou; **Aneta Afelt**, 2018, Monitoring of malaria vectors at the China-Myanmar border while approaching malaria elimination, *Parasites & Vectors*, 11:511
- ²¹ Lacroix, A., Duong, V., Hul, V., Sorn, S., Holl, D. Omaliss, K., Chea, S., Hassanin, A., Theppangna, W., Silithammavong, S., Afelt, A., Greateorex, Z., Fine, A., Goldstein, T., Keatts, L., Frutos, R., Buchy, P., 2016 Genetic diversity of coronaviruses in bats in Lao PDR and Cambodia. *Infection Genetics and Evolution*, 48 (2017) 10–18.
- ²² Lacroix, A., Duong, V., Hul, V., Sorn, S., Holl, D. Omaliss, K., Chea, S., Theppangna, W., Silithammavong, S., Hassanin, A., Afelt, A., Fine, A., Goldstein, T., Joly, D.O., Greateorex, Z.F., Keatts, L., Frutos, R., Buchy, P. 2016. Diversity of bat astroviruses in Lao PDR and Cambodia. *Infection Genetics and Evolution*. 47 (2017) 41–50
- ²³ Afelt A., Lacroix A., Zawadzka-Pawlewska U., Pokojski W., Buchy P., Frutos R., 2018, Distribution of bat-borne viruses and environment patterns, *Infection Genetics and Evolution* 58(2018), 181-191
- ²⁴ Afelt Aneta, Roger Frutos, Christian Devaux, 2018, Bats, Coronaviruses, and Deforestation: Toward the Emergence of Novel Infectious Diseases? *Frontiers in Microbiology*, 9:702. doi: 10.3389/fmicb.2018.00702
- ²⁵ Aneta Afelt, Christian Devaux, Jordi Serra-Cobo and Roger Frutos, 2018, Bats, Bat-Borne Viruses, and Environmental Changes, in: *Bats*, IntechOpen, DOI: 10.5772/intechopen.74377
- ²⁶ M. Nowicki, D. Bzhalava, P. Bała: Massively Parallel Implementation of Sequence Alignment with BLAST Using PCJ Library *J. Comp. Biology* vol. 25 no. 8 pp. 871-881 (2108) DOI: 10.1089/cmb.2018.007
Z. Bzhalava, A. Tampuu, P. Bała, R. Vicente, J. Dillner: Machine Learning for detection of viral sequences in human metagenomic datasets. *BMC Bioinformatics* vol 19 p. 336 (2018) DOI: 10.1186/s12859-018-2340-x
M. Nowicki, D. Bzhalava, P. Bała: Massively Parallel Sequence Alignment with BLAST through Work Distribution Implemented using PCJ Library In: S. Ibrahim, Kim-Kwang R. Choo, Z. Yan, W. Pedrycz (Eds.) *Algorithms and Architectures for Parallel Processing. ICA3PP 2017. Lecture Notes in Computer Science*, vol 10393. Springer, Cham, 2017, pp. 503-512 DOI: 10.1007/978-3-319-65482-9_36
- ²⁷ M. Borcz, R. Kluszczyński, G. Marczak, K. Benedyczak, P. Bała Grid Services for Genomic Research. Polish National Grid perspective In: *Proceedings of the 2014 International Conference on High Performance Computing & Simulation (HPCS 2014)* W. W. Smari, V. Zeljkovic (Eds.) IEEE 2014 pp. 985-990 ISBN: 978-1-4799-5311-0
K. Skonieczna, B. Malyarchuk, A. Jawien, A. Marszałek, Z. Banaszkiewicz, P. Jarmocik, M. Borcz, P. Bała, T. Grzybowski Heteroplasmic substitutions in the entire mitochondrial genomes of human colon cells detected by ultra-deep 454 sequencing *Forensic Science International: Genetics*, 2014 DOI: 10.1016/j.fsigen.2014.10.021
M. Borcz, R. Kluszczyński, G. Marczak, R. Chrupała, K. Benedyczak, K. Skonieczna, T. Grzybowski, P. Bała: PL-Grid services for bioinformatics community. In: M. Bubak, M. Turafa, K. Wiatr (Eds) *CGW'13 Proceedings, ACK CYFRONET AGH, Kraków*, 2013 pp. 67-68 ISBN 978-83-61433-08-8
M. Borcz, R. Kluszczyński, K. Skonieczna, T. Grzybowski, P. Bała Using PL-Grid infrastructure to store and process DNA sequence data from GS FLX Instrument. In: M.

Bubak, M. Turaßla, K. Wiatr (Eds) CGW'12 Proceedings, ACK CYFRONET AGH, Kraków, 2012 pp. 67-68, ISBN 978-83-61433-06-4

M. Borcz, R. Kluszczynski, P. Bała: Statistical Analysis of Biomolecular Data using UNICORE workflows In: A. Fred, J. Filipe, H. Gamboa (eds.) Proc. of BIOINFORMATICS 2010 ISBN 978-989-674-019-1

²⁸ A. Rutkowski, M. Chlebiej, M. Borcz, R. Kluszczynski, K. Benedyczak, P. Bała Seamless Medical Image Processing on the Grid on the Example of Segmentation and Partition of the Airspaces In: M. Hettinga, A. Smedberg, L. Van Gemert-Pijnen, K. Dyb, A. Granstrom Ekeland (Eds.) *eTELEMED 2015 The Seventh International Conference on eHealth, Telemedicine, and Social Medicine* Lisbon, 2015 ISBN: 978-1-61208-384-1

²⁹ B. Zyglarski, P. Bała, J. Mizerski Searching of Correlation of Weather and Cardiologic Events. Computer Methods for Relation Discovery and Events Prediction. In: *HEALTHINF 2014: proceedings of the International Conference on Health Informatics: ESEO, Angers, Loire Valley, France, 3-6 March, 2014* Scitepress - Science and Technology Publications, 2014 p. 529-534 DOI: 10.5220/0004922805290534

³⁰ J. Sierdzinski, P. Bala, R. Rudowski, M. Grabowski, G. Karpinski, B. Kaczynski KARDIONET: Telecardiology Based on GRID Technology In: *Studies in Health Technology and Informatics Volume 150 (Medical Informatics in a United and Healthy Europe - Proceedings of MIE 2009 - The XXIIInd International Congress of the European Federation for Medical Informatics)*, Eds. Klaus-Peter Adlassnig, Bernd Blobel, John Mantas, Izet Masic pp. 463467 ISBN 978-1-60750-044-5 DOI: 10.3233/978-1-60750-044-5-463

³¹ WR Rudnicki, M Kierczak, J Koronacki, J Komorowski, A statistical method for determining importance of variables in an information system, *International Conference on Rough Sets and Current Trends in Computing*, 557-566, 2006

M Kierczak, K Ginalski, M Damiński, J Koronacki, W Rudnicki, A Rough Set-Based Model of HIV-1 Reverse Transcriptase Resistome, *Bioinformatics and biology insights* 3, 109, 2009

M Kierczak, W Rudnicki, J Komorowski, Construction of Rough Set-Based Classifiers for Predicting HIV Resistance to Nucleoside Reverse Transcriptase Inhibitors, *Granular Computing: At the Junction of Rough Sets and Fuzzy Sets*, 249-258, 2008

MB Kursa, A Jankowski, WR Rudnicki, Boruta a system for feature selection, *Fundamenta Informaticae* 101 (4), 271-285, 2010

MB Kursa, WR Rudnicki, Feature Selection with the Boruta Package, *Journal of Statistical Software* 36 (i11), 2010

K Mnich, WR Rudnicki, All-relevant feature selection using multidimensional filters with exhaustive search, *arXiv preprint arXiv:1705.05756*, 2017

R Piliszek, K Mnich, S Migacz, P Tabaszewski, A Sułcecki, A Polewko-Klim, MDfS: MultiDimensional Feature Selection in R, *R Journal* 11 (1), 198-210, 2019

K Niedzilewski, ME Marchwiany, R Piliszek, M Michalewicz, W Rudnicki, Multidimensional Feature Selection and High Performance ParalleX, *SN Computer Science* 1 (1), 40, 2020

³² MT Weirauch, A Cote, R Norel, M Annala, Y Zhao, TR Riley, ..., Evaluation of methods for modeling transcription factor sequence specificity, *Nature Biotechnology*, 2013

P Meyer, G Siwo, D Zeevi, E Sharon, R Norel, E Segal, G Stolovitzky, ..., Inferring gene expression from ribosomal promoter sequences, a crowdsourcing approach, *Genome research* 23 (11), 1928-1937, 2013

DP Noren, BL Long, R Norel, K Rhissorakrai, K Hess, CW Hu, ..., A Crowdsourcing approach to developing and assessing prediction algorithms for AML prognosis, *PLoS Computational Biology*, 2016

A Keller, RC Gerkin, Y Guan, A Dhurandhar, G Turu, B Szalai, ..., Predicting human olfactory perception from chemical features of odor molecules, *Science* 355 (6327), 820-826 75 2017

M Gönen, B Weir, G Cowley, F Vazquez, Y Guan, A Jaiswal, ..., A Community Challenge for Inferring Genetic Predictors of Gene Essentialities through Analysis of a Functional Screen of Cancer Cell Lines, *Cell Systems* 5 (5) 2017

J Guinney, T Wang, TD Laajala, KK Winner, JC Bare, EC Neto, SA Khan, ..., Prediction of overall survival for patients with metastatic castration-resistant prostate cancer: development of a prognostic model through a crowdsourced challenge with open, *The Lancet Oncology* 18 (1), 132-142, 2017

R Kueffner, N Zach, M Bronfeld, R Norel, N Atassi, V Balagurusamy, ..., Stratification of amyotrophic lateral sclerosis patients: a crowdsourcing approach, *Scientific Reports* 9 (690) 2019

³³ A Podsiadło, M Wrzesień, W Paja, W Rudnicki, B Wilczyński, Active enhancer positions can be accurately predicted from chromatin marks and collective sequence motif data, *BMC Systems Biology* 7 (Suppl 6), S16, 2013

³⁴ WR Rudnicki, T Mroczek, P Cudek, Amino Acid Properties Conserved in Molecular Evolution, *PloS one* 9 (6), e98983, 2014

A Polewko-Klim, W Lesiński, K Mnich, R Piliszek, WR Rudnicki, Integration of multiple types of genetic markers for neuroblastoma may contribute to improved prediction of the overall survival, *Biology direct* 13 (1), 17, 2018

³⁵ Ł Ligowski, W Rudnicki, An efficient implementation of Smith Waterman algorithm on GPU using CUDA, for massively parallel scanning of sequence databases, *Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International ...2009*

WR Rudnicki, A Jankowski, A Modzelewski, A Piotrowski, A Zadrożny, The new SIMD implementation of the Smith-Waterman algorithm on Cell microprocessor, *Fundamenta Informaticae* 96 (1), 181-194, 2009

L Ligowski, WR Rudnicki, Y Liu, B Schmidt, Accurate Scanning of Sequence Data bases with the Smith-Waterman Algorithm, *GPU Computing Gems Emerald Edition*, 155-171, 2011

³⁶ Kursa, M.B., Rudnicki, W.R. Feature selection with the Boruta package (2010). *Journal of Statistical Software*, 36 (11), pp. 1-13. [463 cytowania]

- ³⁷ Kursa, M.B. Robustness of Random Forest-based gene selection methods (2014) BMC Bioinformatics, 15 (8)
- ³⁸ Kursa, M.B., Rudnicki, W.R. A deceiving charm of feature selection: The microarray case study (2011) Advances in Intelligent and Soft Computing, 103, pp. 145-152
- ³⁹ Piotrowski, B.P., Kursa, M.B. Fully automatic classification of flow cytometry data (2018) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11177 LNAI, pp. 3-12
- ⁴⁰ Aghaeepour, N., Finak, G., Hoos, H., Mosmann, et al. Critical assessment of automated flow cytometry data analysis techniques (2013) Nature Methods, 10 (3), pp. 228-238
- ⁴¹ Hamed, A., Kursa, M.B. Inter-individual differences in serotonin and glutamate co-transmission reflect differentiation in context-induced conditioned 50-kHz USVs response after morphine withdrawal (2018) Brain Structure and Function, 223 (7), pp. 3149-3167
- ⁴² Hamed, A., Daszczuk, P., Kursa, M.B., Turzyńska, D., Sobolewska, A., Lehner, M., Boguszewski, P.M., Szyndler, J. Non-parametric analysis of neurochemical effects and Arc expression in amphetamine-induced 50-kHz ultrasonic vocalization (2016) Behavioural Brain Research, 312, pp.174-185
- ⁴³ <https://tissuemodelling.github.io/timothy/>
- ⁴⁴ M.Cytowski, Z.Szymanska, "Large Scale Parallel Simulations of 3-D Cell Colony Dynamics", IEEE Computing in Science and Engineering, vol.16, no.5, pp.86,95, Sept.-Oct. 2014
- ⁴⁵ M.Cytowski, Z.Szymanska, "Large-Scale Parallel Simulations of 3D Cell Colony Dynamics: The Cellular Environment", IEEE Computing in Science and Engineering , vol.17, no.5, pp.44-48, Sept.-Oct. 2015
- ⁴⁶ Szymańska, Z., Cytowski, M., Mitchell, E., Macnamara, C.K., Chaplain, M.A.J., "Computational modelling of cancer growth and invasion: Modelling at multiple scales and multiscale modelling"; Bull. Math. Biol., 805:1366-1403, 2018
- ⁴⁷ Cytowski, M., Szymańska, Z., Umiński, P., Andrejczuk, G., Raszkowski, K., "Implementation of an Agent-Based Parallel Tissue Modelling Framework for the Intel MIC Architecture", Sci. Program., 2017, 1-11, 2017
- ⁴⁸ Cytowski, M., Szymańska, Z., "Large Scale Parallel Simulations of 3-D Cell Colony Dynamics. II. Coupling with continuous description of cellular environment", IEEE Computing in Science & Engineering, 17(5):44-48, 2015
- ⁴⁹ Cytowski, M., Szymańska, Z., "Enabling Large Scale Individual-Based Modelling through High Performance Computing"; EDP Sciences - ITM Web of Conferences, Vol. 5, 2015
- ⁵⁰ Szymańska, Z., Cytowski, M., Mitchell, E., Macnamara, C.K., Chaplain, M.A.J., "Computational modelling of cancer growth and invasion: Modelling at multiple scales and multiscale modelling", Bulletin of Mathematical Biology

- ⁵¹ Cytowski, M., Szymańska, Z., Umiński, P., Andrejczuk, G., Raszkowski, K., "Implementation of an agent-based parallel tissue modelling framework for the Intel MIC architecture", Scientific Programming
- ⁵² Niedzielewski, K., Marchwiany, M.E., Piliszek, R. *et al.* Multidimensional Feature Selection and High Performance ParalleX. *SN COMPUT. SCI.* **1**, 40 (2020) doi:10.1007/s42979-019-0037-5
- ⁵³ Szwajczak, IJ. Fijalkowska and **C. Suski**. (2018) The importance of an interaction network for proper DNA polymerase ζ heterotetramer activity. *Curr Genet.* **64**(3), 575-580
- Szwajczak, IJ. Fijalkowska and **C. Suski**. (2017) The CysB motif of Rev3p involved in the formation of the four-subunit DNA polymerase ζ is required for defective-replisome-induced mutagenesis. *Mol Microbiol.* **106**:659–672
- Grabowska E, U. Wronska, M. Denkiewicz, M. Jaszczur, A. Respondek, M. Alabrudzinska, **C. Suski**, K. Makiela-Dzbenka, P. Jonczyk, IJ. Fijalkowska. (2014) [Proper functioning of the GINS complex is important for the fidelity of DNA replication in yeast.](#) *Mol Microbiol.* **92**:659-80.
- Larsen NB, E. Sass, **C. Suski**, HW. Mankouri, ID. Hickson. (2014) [The Escherichia coli Tus-Ter replication fork barrier causes site-specific DNA replication perturbation in yeast.](#) *Nat Commun.* **5**:3574.
- Suski** and K. J. Mariani (2008) Resolution of converging replication forks by RecQ and topoisomerase III. *Mol Cell.* **30**:779-89.
- Janniere, D. Canceill, **C. Suski**, S. Kanga, B. Dalmais, R. Lestini, AF. Monnier, J. Chapuis, A. Bolotin, M. Titok, EL Chatelier, SD. Ehrlich. (2007) Genetic evidence for a link between glycolysis and DNA replication. *PLoS ONE* **2**:e447.
- Titok, **C. Suski**, B. Dalmais, S.D. Ehrlich and L. Janniere. (2006) The replicative polymerases PolC and DnaE are required for theta replication of the *Bacillus subtilis* plasmid pBS72. *Microbiology.* **152**: 1471-8.
- Panozzo, M. Nawara, **C. Suski**, R. Kucharczyka, M. Skoneczny, A.M. Bécam, J. Rytko and C.J. Herbert (2002) Aerobic and anaerobic NAD⁺ metabolism in *Saccharomyces cerevisiae*. *FEBS letters* **517**: 97-102.
- Dervyn, **C. Suski**, R. Daniel, C. Bruand, J. Chapuis, J. Errington, L. Jannière and S.D. Ehrlich. (2001) Two essential DNA polymerases at the bacterial replication fork. *Science* **294**: 1716-1719.