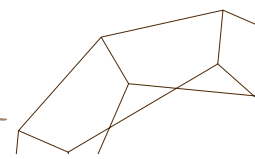




PROFUNDERE SCIENTIAM



nr 19
luty 2025

BIULETYN CENTRUM STUDIÓW ZAAWANSOWANYCH POLITECHNIKI WARSZAWSKIEJ



Współczesne baterie – hit czy kit?

O ogromnych potrzebach, możliwościach, ale także ograniczeniach wynikających z technologii wytwarzania baterii, w oparciu o badania własne, materiałów i technologii, prowadzone przez zespół pod kierunkiem prof. Władysława Wieczorka na Wydziale Chemicznym Politechniki Warszawskiej, we współpracy z wieloma krajowymi i zagranicznymi jednostkami naukowymi oraz z przemysłem – zagadnienia dyskutowane w ramach cyklu seminariów Horyzonty Dyscyplin Nauki – strategia rozwoju nauki w Politechnice Warszawskiej

Żyjemy w czasach przełomowych. Coraz bardziej uświadamiamy sobie, że ludzkość doszła do punktu, w którym niezbędne są istotne zmiany kierunków rozwoju. Procesy towarzyszące dotychczasowemu rozwojowi wywołują skutki uboczne, które mogą w dramatyczny sposób zahamować postęp a może nawet unicestwić cywilizację. Chaotyczny i niekontrolowany rozwój powoduje wyczerpywanie zasobów, które nie są nieskończone, a z drugiej strony produkujemy odpady, których nagromadzenie już wywołuje negatywne skutki o charakterze globalnym. Na pierwszy rzut oka wydaje się, że jak w greckiej tragedii nie ma dobrego wyjścia – albo drastycznie obniżymy poziom życia już teraz, albo utoniemy płynąc luksusowym statkiem, słuchając orkiestry grającej do końca.

Na szczęście nie brak optymistów, którzy nie zgadzają się z taką katastroficzną oceną i przekonują, że problem

da się rozwiązać. Problem jest rozwiązywalny, ponieważ sprowadza się w gruncie rzeczy do rosnącego zapotrzebowania na energię, której źródło, przekazywane przez promieniowanie słoneczne, może być uznane za nieograniczone, jeśli odniesiemy to do ludzkiej perspektywy. Problem jest rozwiązywalny, ponieważ staje się problemem naukowym, w którym redukuje się do magazynowania i przetwarzania energii, a nie jej wytwarzania. Postęp w tej dziedzinie jest bardzo szybki, chociaż wciąż niewystarczający. Znaczący udział w badaniach nad takimi urządzeniami ma Wydział Chemiczny Politechniki Warszawskiej.

Współczesne baterie – problemy badawcze

Najbardziej wydajne urządzenia do magazynowania energii, z punktu

W NUMERZE

między innymi:

- *Współczesne baterie – hit czy kit?* – prof. Władysław Wieczorek, prof. Marek Marcinek, prof. Janusz Płocharski (s. 1, dokończenie s. 6)
- *Iluzja terażniejszości* – prof. Stanisław Janeczko (s. 1, dokończenie s. 23)
- *Zapewnianie jakości oprogramowania z uczeniem maszynowym* – prof. Agnieszka Jastrzębska (s. 12)
- *Upcyding AI* – Karolina Bogacka (s. 16)
- *(f)wiadomość* – dr hab. Jacek Rogala (s. 20)
- *Projekt MONIKA* – prof. Mirosław Rodzewicz (s. 26)
- *Dwie twarze sztucznej inteligencji* – Piotr Sowiński (s. 33)

ILUZJA TERAŻNIEJSZOŚCI

Stanisław Janeczko

Nasz sensoryczny świat to świat terażniejszości. Wiemy o tym, że żyjemy w czasie i że uchwycenie obiektywnych i subiektywnych cech tego zjawiska jest fundamentem realnego poznania rzeczywistości. Słynna definicja „prawdziwego i absolutnego” czasu Newtona, który „na podstawie własnej natury i sam z siebie płynie równomiernie bez żadnej relacji do czegokolwiek na zewnątrz” określa, w pewien sposób, obiektywne zjawisko będące fundamentem, tak dla praw rządzących materią, jak i dla naszego sposobu doświadczenia czasu. Natomiast „Względny, subiektywny i zwyczajny czas jest zewnętrzną, przekazaną nam

{CIĄG DALSZY NA S. 6}

{CIĄG DALSZY NA S. 23}

DZIAŁALNOŚĆ CENTRUM STUDIÓW ZAAWANSOWANYCH PW

czyli najważniejsze wydarzenia i najbliższe plany CSZ PW

PROFESOROWIE WIZYTUJĄCY

W 2024 roku Centrum Studiów Zaawansowanych gościło wielu wybitnych gości ze świata nauki.

W maju 2024 r. prof. Sir Michael Berry z Uniwersytetu w Bristolu (Wielka Brytania) odebrał Wyróżnienie CSZ PW „Kosmos Pitagorasa” i z tej okazji wygłosił odczyt pt. *Four geometrical-optics illusions* oraz wziął udział w Dyspucie Pitagorejkiej pt. *Internal and external limits of knowledge*.

W związku z organizacją międzynarodowych warsztatów Polish-Japanese Singularity Theory Working Days, we wrześniu 2024 r. gośćmi Centrum oraz Wydziału Matematyki i Nauk Informacyjnych PW byli m.in. Osamu Saeki (Kyushu University), Kentaro Saji (Kobe University), Goo Ishikawa (Hokkaido University), Federico Sánchez-Bringas (Universidad Nacional Autónoma de México), Yosef Yomdin (Weizmann Institute of Science), Takashi Nishimura (Yokohama National University), Takahiro Yamamoto (Tokyo Gakugei University), Asahi Tsuchida (Shiga University), Hiroshi Teramoto (Kansai University), Masato Tanabe (Hokkaido University), Masatomo Takahashi (Muroran Institute of Technology), Runa Shimada (Kobe University), Ryosuke Ota (Kyushu University), Nozomi Nakatsuyama (Muroran Institute of Technology), Naomichi Nakajima (Shibaura Institute of Technology), Lucía Ivonne Hernández Martínez (Universidad Autónoma de la Ciudad de México).

KONWERSATORIUM PW

W ramach Konwersatorium Politechniki Warszawskiej, 7 maja 2024 r. odbył się odczyt specjalny Laureata Wyróżnienia CSZ PW „Kosmos Pitagorasa”, prof. Sir Michaela Berry’ego z Uniwersytetu w Bristolu, pt. *Four geometrical-optics illusions*.

KONWERSATORIUM PW – TOPTECHNIKA

Centrum Studiów Zaawansowanych Politechniki Warszawskiej kontynuuje cykl spotkań konwersatoryjnych – TOPTECHNIKA, których celem jest przybliżenie środowisku akademickiemu najnowszych osiągnięć techniki poprzez wystąpienia wybitnych naukowców – praktyków z renomowanych instytucji i firm. W 2024 r. spotkaliśmy się dwa razy:

21 marca 2024 r. – Jerzy Orkiszewski, US-Polish Trade Council (USPTC), USA, *Wyzwania napotymane przy wprowadzeniu sztucznej inteligencji do urządzeń medycznych*.

Prelegent omówił problemy napotymane przy wprowadzaniu rozwiązań AI do terapeutycznych urządzeń medycznych. Jako przykład posłużył nowatorski laser medyczny przeznaczony do leczenia szeregu schorzeń dermatologicznych. Przedstawiony został proces definiowania funkcji systemu, wyboru technologii AI i jej implementacji w urządzeniu. Wykład był oparty o rzeczywisty proces projektowy mający obecnie miejsce w USA.

12 grudnia 2024 r. – prof. Jerzy Rużyłło, Penn State University, USA, *Inżynieria półprzewodnikowa na przestrzeni lat*.

W swoim wystąpieniu prof. J. Rużyłło rozważał inżynierię półprzewodnikową jako dziedzinę, która swoim zakresem obejmuje tematy związane z materiałami i przyrządami półprzewodnikowymi, metodami zaawansowanych technologii ich wytwarzania, jak i z metodami charakteryzacji materiałów półprzewodnikowych oraz monitorowania procesów stosowanych w produkcji przyrządów półprzewodnikowych. Celem była również dyskusja zagadnień determinujących rozwój omawianych wyżej elementów



inżynierii półprzewodnikowej na przestrzeni ostatnich lat oraz podzielenie się ze słuchaczami refleksjami na temat przewidywanych kierunków ich dalszego rozwoju.

KOSMOS PITAGORASA

7 maja 2024 roku odbyła się uroczystość wręczenia Wyróżnienia CSZ PW „Kosmos Pitagorasa” dla prof. Sir Michaela Berry'ego, brytyjskiego fizyka teoretycznego, laureata takich prestiżowych nagród jak Medal Lorentza, Royal Medal i Nagroda Wolfa w dziedzinie fizyki. Profesor współpracował z Andriejem Gejmem, późniejszym laureatem Nagrody Nobla. Zajmuje się zagadnieniami z pogranicza fizyki klasycznej, mechaniki kwantowej i optyki. Jest odkrywcą fazy Berry'ego.

Z tej okazji laureat wygłosi odczyt pt. „Four geometrical-optics illusions”. Uroczystość odbyła się w Sali Senatu Politechniki Warszawskiej i zgromadziła wielu gości zarówno z naszej uczelni jak również innych ośrodków naukowych.

WYKŁADY UOD CSZ PW

W Uczelnianej Ofercie Dydaktycznej Centrum Studiów Zaawansowanych Politechniki Warszawskiej w roku akademickim 2023/2024 zorganizowano 3 wykłady podstawowe oraz 14 specjalnych, także w formie laboratoriów. W wykładach uczestniczyło kilkaset osób, głównie doktorantów szkoły doktorskiej PW.

Słuchacze mieli możliwość wyboru spośród różnorodnej oferty tematycznej, zarówno kontynuacji jak i nowości:

1. *Modele matematyczne procesów i przemian*, prof. Stanisław Janeczko (PW);
2. *Terapie XXI wieku*, prof. Ewa Bartnik (UW);
3. *Układy złożone. Przypadkowość i przewidywalność*, prof. Marek Kuś (PAN);
4. *Monitorowanie wysiłku fizycznego*, dr inż. Monika Petelczyc (PW);
5. *Podstawy rozpoznawania obrazów*, dr hab. inż. Agnieszka Jastrzębska, prof. uczelni (PW);
6. *Rozmowy i Rozumowania*, prof. Stanisław Janeczko (PW);
7. *Psychologia wpływu społecznego – jak wpływamy na innych i jak inni wpływają na nas*, dr Leszek Mellibruda (Akademia Handlowa Nauk Stosowanych);
8. *Jak wydobycić potencjał twórczy grupy? Techniki pracy twórczej w grupie*, dr Bartłomiej Skowron (PW);

9. *W poszukiwaniu teraźniejszości*, mgr inż. Barbara Majerska, mgr inż. Aleksandra Przywózka (PW);
10. *Przedsiębiorstwo 4.0. Koncepcja-model-praktyka*, prof. Mieczysław Morawski (PW);
11. *Rysunek odręczny jako narzędzie pracy inżyniera*, dr inż. arch. Joanna Pętkowska-Hankel (PW);
12. *Deformanci są wśród nas, czyli psychoprofilaktyka zniekształceń osobowości*, dr Leszek Mellibruda (Akademia Handlowa Nauk Stosowanych);
13. *Podstawy wnioskowania statystycznego dla inżynierów*, prof. Anna Dembińska (PW);
14. *Wizyjne techniki skanowania i przetwarzania danych 3D*, prof. Robert Sitnik (PW);
15. *Rozmowy i rozumowania*, prof. Stanisław Janeczko (PW).

*) Szczegółowy wykaz archiwalnych i realizowanych przedmiotów znajduje się na stronie www.konwersatorium.pw.edu.pl/oferta, dostępnej również ze strony Centrum – www.csz.pw.edu.pl.

HORYZONTY DYSCYPLIN NAUKI

Zapoczątkowany w kwietniu 2023 r. z inicjatywy Prorektora ds. Nauki, prof. Mariusza Malinowskiego i przy koordynacji prof. Janusza Zawiły-Niedźwieckiego cykl seminariów pn. „Horyzonty Dyscyplin Nauki”, miał swoją kontynuację w 2024 r. Spotkania mają na celu przedstawienie osiągnięć zespołów reprezentujących poszczególne dyscypliny nauki w Politechnice Warszawskiej. Dotychczas odbyło się 10 spotkań w formule online – seminaria, które miały miejsce w 2024 r. to:

18.01.2024 – *Współczesne baterie – hit czy kit?*; dyscyplina: Nauki chemiczne; prelegenci: prof. dr hab. inż. Władysław Wierzchowski, prof. dr hab. inż. Janusz Płocharski, prof. dr hab. inż. Marek Marcinek (WCh PW); koreferenci: prof. dr hab. Paweł Kulesza (Wydział Chemii UW), dr hab. inż. Krzysztof Fic (Wydział Technologii Chemicznej, Politechnika Poznańska); wydarzenie poprowadził prof. dr hab. inż. Tadeusz Hofman (WCh PW). | → zob. s. 1

4.04.2024 – *Rola modelowania i badań in-situ w zakresie zmian klimatu i jakości powietrza atmosferycznego w celu ochrony środowiska i zdrowia*; dyscyplina: Inżynieria środowiska, górnictwo i energetyka; prelegentki: dr inż. Magdalena Reizer, dr inż. Katarzyna Maciejewska (WIBHiŚ PW); koreferenci: dr inż. Krzysztof Skotak (IOŚ – PIB) i Łukasz Adamkiewicz (Europejskie

Centrum Czystego Powietrza); wydarzenie poprowadził dr hab. inż. Mirosław Seredyński, prof. PW (WMEiL).

18.04.2024 – *Mikroskopia bez znaczników (grant ERC)*; dyscyplina: Automatyka, elektronika, elektrotechnika i technologie kosmiczne; prelegent: dr hab. inż. Maciej Trusiak (WMech. PW); koreferent: dr inż. Piotr Zdańkowski (WMech. PW); wydarzenie poprowadził prof. dr hab. inż. Desiré Rasolomampionona (WE PW).

16.05.2024 – *Miasta przyszłości. Utopia czy konieczność?*; dyscyplina: Architektura i urbanistyka; prelegent: prof. dr hab. arch. Elżbieta Ryńska (WA PW); koreferent: prof. dr hab. inż. Jacek Szoltysek z Uniwersytetu Ekonomicznego w Katowicach.

21.11.2024 – *Współpraca lekarz-inżynier*; dyscyplina: Inżynieria biomedyczna; prelegent: prof. Marcel Młynczak (WMech. PW); wydarzenie poprowadził prof. Tadeusz Pałko (WMech. PW).

16.01.2025 – *Pomiędzy łatwym a trudnym*, dr hab. inż. Paweł Rzażewski (WMI NI PW) oraz *Dwa problemy z fizyki rozwiązane ścisłe*, dr hab. Tomasz Cieślak (WMI NI PW); dyscyplina: Matematyka; wydarzenie poprowadził prof. dr hab. Grzegorz Świątek (WMI NI PW).

SYMPOZJUM CENTRUM STUDIÓW ZAAWANSOWANYCH

Warsztaty Ścisłego Myślenia

W dniach 22-24 marca 2024 r. w Europejskim Centrum Edukacji Geologicznej Uniwersytetu Warszawskiego odbyły się Warsztaty Ścisłego Myślenia. Zostały zorganizowane przez Centrum Studiów Zaawansowanych PW we współpracy z Centrum Badania Ryzyka Systemowego UW ze szczególną myślą o studentach indywidualnych studiów politechnicznych – studiów ID. | → zob. s. 10

II Warsztaty Koncepcyjne CBRS – CSZ w Chęcinach

Również w Europejskim Centrum Edukacji Geologicznej UW w Chęcinach, w dniach 10-12 maja 2024 r. odbyły się II Warsztaty Koncepcyjne Centrum Studiów Zaawansowanych Politechniki Warszawskiej oraz Centrum Badania Ryzyka Systemowego UW. Spotkanie zorganizowano pod hasłem „Procesy i przemiany w układach złożonych (cz. 2)”. W warsztatach udział wzięli naukowcy z wielu jednostek naukowych. | → zob. s. 10

Dydaktyka dla poznania, rozumienia i działania

W dniach 1-4 września 2024 r., w Ośrodku Wypoczynkowy PW w Wildze odbyło się sympozjum „Dydaktyka dla poznania, rozumienia i działania” poświęcone odkrywaniu nowych możliwości, ale również identyfikowaniu słabych punktów oraz wyzwań w procesach dydaktycznych uczelni wyższych. Główny trzon uczestników Sympozjum stanowili laureaci konkursów na granty dydaktyczne w PW – członkowie zespołów, którzy dostrzegli potencjał zmian i podjęli wysiłek unowocześnienia sfery dydaktycznej Politechniki Warszawskiej. | → zob. s. 11

Polish-Japanese Singularity Theory Working Days

Międzynarodowe warsztaty “Polish-Japanese Singularity Theory Working Days” zostały zorganizowane przy współpracy Centrum Studiów Zaawansowanych Politechniki Warszawskiej, Institute of Mathematics for Industry, Kyushu University, Department of Mathematics, Hokkaido University oraz Wydziału Matematyki i Nauk Informatycznych Politechniki Warszawskiej. Odbywały się w dniach 9-14 września 2024 r. na Wydziale Matematyki i Nauk Informatycznych PW. Warsztaty składały się z regularnych wykładów oraz mini-wykładów, w tym: → *Faces of Mathematical Modeling*, Łukasz Błaszczyk;

- *Singularity Theory in Super-resolution and Whitney extensions*, Yosef Yomdin;
- *Why do We Need Modern Geometry?*, Michał Zwierzyński .

STUDIA ID

Program elitarnych studiów id dla najzdolniejszych studentów i kandydatów na studia jest realizowany na wcześniejszych zasadach. Pod opieką tutorów młodzi naukowcy kreują swoje indywidualne ścieżki. | → zob. s. 31

INTERDYSCYPLINARNE SEMINARIUM STUDIÓW ID

W ramach aktywizacji studentów realizujących indywidualny tryb studiów id, został uruchomiony cykl „Interdyscyplinarnych seminariów studiów id”.

Do poprowadzenia spotkań zapraszane są wybitne osoby świata nauki o wysokim potencjale popularyzatorskim, aktywne w dziedzinach, które intensywnie się rozwijają i stanowią ciekawą perspektywę dla młodych naukowców.

W minionym roku odbyły się kolejne spotkania z tego cyklu. Wśród zaproszonych gości znaleźli się: prof. Stanisław Janeczko oraz prof. Agnieszka Jastrzębska z Wydziału Matematyki i Nauk Informatycznych PW, prof. Ewa Bartnik z Wydziału Biologii UW, prof. Adam Proń z Wydziału Chemicznego PW, prof. Małgorzata Kujawińska z Wydziału Mechatroniki PW, Sir Michael Berry z Uniwersytetu w Bristolu, dr Anna Pięcińska z Instytutu Historii

Sztuki UW, prof. Marek Kuś z Centrum Fizyki Teoretycznej PAN.

SEMINARIUM „MATEMATYCZNE I INFORMATYCZNE GRANICE ALGORYTMICZNOŚCI”

Centrum Studiów Zaawansowanych oraz Wydział Matematyki i Nauk Informatycznych PW zainaugurowały, wraz z rozpoczęciem roku akademickiego 2024/2025, cykl seminariów poświęconych sztucznej inteligencji. Dziedzina inteligencji obliczeniowej bije rekordy popularności. Jej postęp widzimy zarówno w nowych polach zastosowań jak i coraz to doskonalszych algorytmach. Rozwój jest szybki i fascynujący – na przecięciu matematyki i informatyki powstał obszar wiedzy, o którym można powiedzieć, że wpływa (prawie) na każdego i trudno przejść obok niego obojętnie. Na kolejnych spotkaniach z cyklu poruszane są m.in. zagadnienia reprezentacji danych, złożoności obliczeniowej i wiele innych. Zawsze znajdzie się również miejsce na dyskusję. Do tej pory odbyły się trzy seminaria, na każdym po dwa wykłady. Oto przedstawione i dyskutowane tematy:

- *Wybrane zagadnienia rozwoju oprogramowania z elementami wizji komputerowej*, Agnieszka Jastrzębska, WMiNI PW
- *O pewnych związkach matematyki z teorią złożoności obliczeniowej*, Zbigniew Lonc, WMiNI PW





- *Pojęcia trudności w informatyce*, Paweł Rzażewski, WMiNI PW
- *Klasyfikacja binarna na podstawie danych pozytywnych i nieetykietowanych*, Paweł Teisseyre, WMiNI PW
- *Elementy geometrii czasoprzestrzeni*, Tomasz Miller, CKBI UJ
- *O potrzebie modelowania w statystyce i uczeniu maszynowym*, Jan Mielniczuk, WMiNI PW

Seminarium cieszy się bardzo dużym zainteresowaniem i mamy nadzieję, że jeszcze wiele takich spotkań przed nami. Zapraszamy do śledzenia ogłoszeń na stronie Centrum
→ www.csz.pw.edu.pl

SPOTKANIE ŚWIĄTECZNE

Na zakończenie roku 2024 zostało zorganizowane spotkanie pod przewodnictwem Dyrektora Centrum, prof. Stanisława Janeczko oraz Koordynatora studiów ID, dr Barbary Roszkowskiej-Lech. Było ono przewidziane z myślą o integracji uczestników programu studiów ID, z udziałem kadry naukowej PW, w tym Członków Rady Naukowej Studiów id oraz Rady Programowej Centrum Studiów Zaawansowanych, którzy odebrali swoje nominacje na nową kadencję. Podczas spotkania odbyło się także uroczyste wręczenie wyróżnienia „Genius Cognito” prof. Leonowi Gradoniowi za wieloletnie zaangażowanie i współpracę z Centrum. W programie nie zabrakło oczywiście świątecznych potraw i słodkości.

MOBILITY PW

Kontynuację, w 2024 roku miał również program Mobility PW trwający

od 1 września 2021 r. Jest on realizowany w ramach przedsięwzięcia „Inicjatywa doskonałości – uczelnia badawcza” oraz projektów SEED i PROM PW 2 na zasadach określonych wspólnym regulaminem oraz zgodnych z wytycznymi dla poszczególnych projektów. Program otworzył nowe możliwości ubiegania się o dofinansowanie mobilności naukowych doktorantom i pracownikom Politechniki Warszawskiej. Jego realizacja odbywa się poprzez ogłaszanie konkursów ocenianych przez Komisję Konkursową powołaną na okres realizacji Mobility PW. Celem przedsięwzięcia, a co za tym idzie każdego z konkursów jest doskonalenie kompetencji doktorantów i pracowników Politechniki Warszawskiej poprzez udział w aktywnościach zagranicznych, a także stymulowanie wymiany doświadczeń poprzez wizyty cudzoziemców. Finansowanie przyznawane jest najwyższej ocenionym wnioskowi, przy uwzględnieniu środków dostępnych dla poszczególnych typów aktywności. Program jest koordynowany przez Centrum Studiów Zaawansowanych. Do tej pory ogłoszono i rozstrzygnięto 13 konkursów. Finansowanie mobilności przyznano 198 osobom – 100 doktorantom PW, 2 doktorantom zagranicznym i 96 pracownikom PW.

INNOWACJE DYDAKTYCZNE

Poza działaniami monitorującymi projekty wybrane do realizacji w roku akademickim 2023/2024, została przeprowadzona kolejna, już czwarta, edycja konkursu na granty dydaktyczne PW. Zbieranie wniosków, wzorem

poprzednich konkursów, odbywało się poprzez formularz dostępny na stronie www.badawcza.pw.edu.pl i zakończyło się 29 września 2024 roku. Wnioski zostały ocenione przez Komisję Konkursową. W wyniku rozstrzygnięcia czwartego konkursu na granty dydaktyczne w PW, przyznano dofinansowanie 11 projektom. Ich realizacja przypada na rok akademicki 2024/2025.

Ważnym wydarzeniem podsumowującym do tej pory zrealizowane granty było sympozjum „Dydaktyka dla poznania, rozumienia i działania”, które odbyło się 1-4 września 2024 w Ośrodku Wypoczynkowy PW w Wildze.

| → zob. s. 11

WYDAWNICTWA

Publikacje CSZ, w tym liczne książki naukowe, można nabyć w księgarniach Oficyny Wydawniczej PW w Gmachu Głównym i przy ul. Noakowskiego 18/20 w Warszawie.

*) Przegląd wszystkich dotychczasowych pozycji można odnaleźć pod adresem: www.csz.pw.edu.pl/ Wydawnictwa.

NAJBLIŻSZE PLANY

SYMPOZJA

W 2025 roku planowane są 3 wyjazdowe sympozja:

- marzec 2025 – „Warsztaty ścisłego myślenia”, z udziałem naukowców Politechniki Warszawskiej, Centrum Badania Ryzyka Systemowego UW oraz uczestników studiów id;
- maj 2025 – „Warsztaty koncepcyjne CBRS-CSZ”;
- jesień 2025 – międzynarodowe warsztaty matematyczne.

SEMINARIA

Z niecierpliwością czekamy na nowe spotkania w naszych cyklach, o których będziemy informować z odpowiednim wyprzedzeniem.

STUDIA ID

Swoją kontynuację będzie miał również program studiów dla najzdolniejszych na Politechnice Warszawskiej. Planowane są kolejne procesy rekrutacyjne i wiele działań towarzyszących urozmaiceniu oferty dla całej grupy, w tym wyjazdy i spotkania dyskusyjno-integrujące.

widzenia masowego zastosowania, to baterie, ogniwa paliwowe i superkondensatory. Obecnie, spośród nich, jednymi z najbardziej obiecujących są akumulatory litowo-jonowe. Pomimo trzydziestu lat komercyjnego istnienia technologii litowo-jonowej i jej uznania poprzez przyznanie Nagrody Nobla w 2019 roku (J.B. Goodenough, M.S. Whittingham i A. Yoshino), technologia ta wciąż ma wady utrudniające jej pełne rozpowszechnienie. Istnieje zapotrzebowanie na nowe elementy składowe akumulatorów, tj. elektrody i elektrolity, które zagwarantowałyby większą pojemność, wzrost bezpieczeństwa i szerszy zakres stabilności elektrochemicznej przy jednoczesnym wydłużeniu żywotności urządzenia. Wobec ograniczonych ilości surowców niezbędnych do budowy akumulatorów pilną sprawą jest zaprojektowanie nowych elektrod i elektrolitów w oparciu o łatwo dostępne materiały, niski koszt i łatwość wytwarzania. Zdano sobie sprawę, że zjawiska występujące na granicy faz, głównie pomiędzy anodą a elektrolitem, mają ogromne znaczenie dla efektywnego działania akumulatorów oraz zrozumienia, kontrolowania i projektowania ich funkcji. Te powierzchnie zetknięcia się faz (interfejsy) są miejscami, które ułatwiają magazynowanie energii w akumulatorach, ale także w nich rozpoczyna się wiele procesów degradacyjnych. W przypadku akumulatorów litowo-jonowych stwierdzono, że jeden z głównych mechanizmów starzenia, czyli wzrostu tak zwanej interfejsy stałego elektrolitu (SEI) na anodach grafitowych, w istotny sposób wpływa na stabilność podczas cykli pracy.

„Wyniki badań prowadzonych na Wydziale Chemicznym opisano w ponad 160 artykułach w wiodących czasopiśmie międzynarodowych, które do dziś cytowane były ponad 5000 razy, ponadto w 34 patentach, w tym w 12 patentach międzynarodowych”

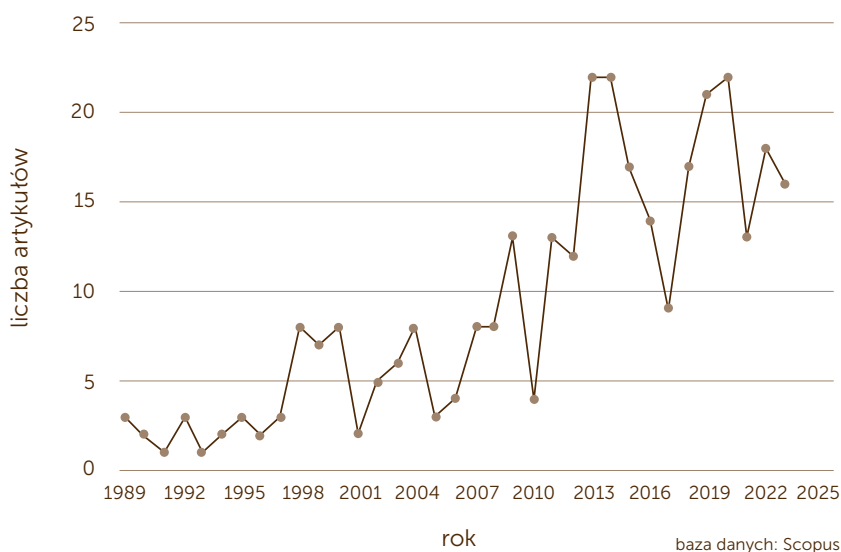
Reaktywność międzyfazowa może zostać zredukowana dzięki tworzeniu się warstwy tzw. pasywującej międzyfazy, która umożliwia dyfuzję jonów, ale jest przeszkodą dla elektronów. Taka konstrukcja pomoże przedłużyć stabilność termodynamiczną elektrolitów organicznych stosowanych w akumulatorach.

Do kontrolowania wydajności interfejsów w akumulatorach litowo-jonowych stosuje się dwie główne metody. Do elektrolitu dodaje się różne składniki budujące SEI. Związki te ulegają degradacji przy potencjałach wyższych niż potencjał degradacji głównych składników elektrolitu (anionów soli i rozpuszczalnika), na których tworzy

się warstwa ochronna SEI. Inną strategią jest sztuczne utworzenie warstwy SEI pokrywającej powierzchnię anody [1].

Prace Wydziału Chemicznego PW nad elektrolitami

Naukowcy z Wydziału Chemicznego Politechniki Warszawskiej są światowymi liderami w dziedzinie chemii materiałów magazynujących energię i inżynierii akumulatorów litowo-jonowych. Odnosi się to szczególnie do projektowania i rozwoju nowych elektrolitów i ich składników stosowanych w akumulatorach litowo-jonowych. Analizowane są również inne układy magazynowania energii. Wyniki badań prowadzonych na Wydziale Chemicznym opisano w ponad 170 artykułach w wiodących czasopiśmie międzynarodowych, które do dziś cytowane były ponad 6500 razy, ponadto w 34 patentach, w tym w 12 patentach międzynarodowych. Ostatnio zaprojektowano i opracowano syntezę nowych soli należących do grupy tzw. soli Hückela, stosowanych jako dodatki do elektrolitów w akumulatorach litowo-jonowych [2]. Produkcja soli została skomercjalizowana w 2018 roku przez francuską firmę chemiczną Arkema. Sól sprzedawana jest pod nazwą handlową Foranext. Opisane badania stały się podstawą do stworzenia europejskiego programu nauczania na poziomie magisterskim, znanego jako MESC (*Materials for Energy Conversion and Storage*). Jako rozszerzenie istniejącego programu magisterskiego, w 2020 roku utworzono nowy program doktorancki o nazwie *Destiny*, wspierany przez Komisję



↑ Rys. 1 Liczba artykułów zawierających cytowania publikacji: J. Płocharski, W. Wieczorek, „PEO based composite solid electrolyte containing nasicon”, *Solid State Ionics*, Vol. 28–30, Part 2 (1988), pp 979–982. Do grudnia 2023 roku ta praca miała 324 cytowania

Europejską w ramach środków Funduszu Marii Skłodowskiej-Curie.

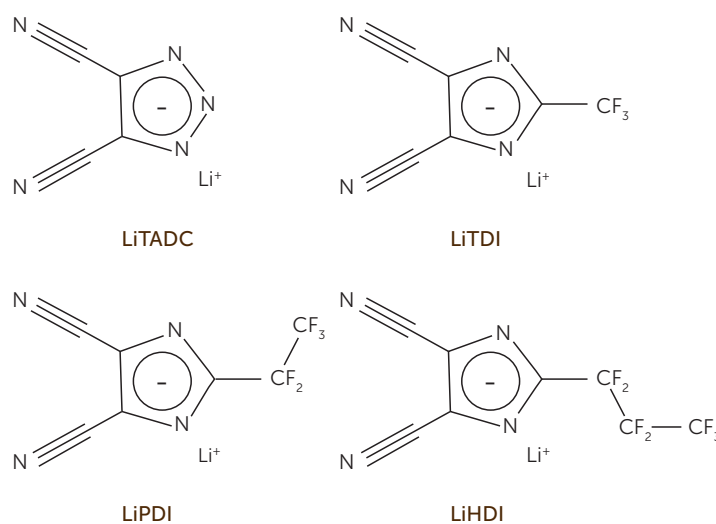
Badania prowadzone na Wydziale Chemicznym Politechniki Warszawskiej rozpoczęły się w połowie lat 80-tych XX wieku. Jedną z pierwszych publikacji poświęconych elektrolitom stałym ukazała się we wrześniu 1988 r. Rysunek 1 prezentujący przyrost cytowań tej pracy na przestrzeni ponad trzech dekad ilustruje również wzrost zainteresowania elektrolitami polimerowymi. Z ciekawostek można w tym miejscu dodać, że obecnie w bazie Scopus jest ok 9 tys. publikacji im poświęconych, a jedynie w 2023 r. ukazało się 881 prac pod hasłem *Composite Polymer Electrolytes*. Wracając na Wydział Chemiczny PW, badania początkowo skupiały się na kompozytowych elektrolitach polimerowych. Szczególny nacisk położono na rozwinięcie teorii transportu jonów w materiałach kompozytowych polimerowo-ceramicznych. Obecnie jest ona powszechnie wykorzystywana do projektowania nowych układów kompozytowych. Kolejnym tematem były tzw. receptory anionowe - ich celem było zwiększenie przewodności, jak i liczby przenoszenia kationów litu w kompozytowych elektrolitach polimerowych. Zastosowanie elektrolitów kompozytowych w akumulatorach litowych zapewnia znaczne zmniejszenie oporności na styku elektroda-elektrolit, a tym samym poprawia bezpieczeństwo i żywotność akumulatora. Następnie podjęto badania nad nową klasą stałych przewodników protonowych, zbudowanych na bazie niewodnych żelowych elektrolitów polimerowych. Przewodniki te mogą być zastosowane w urządzeniach elektrochromowych. Koncepcja niewodnych żelowych elektrolitów polimerowych przewodzących protony jest obecnie stosowana w superkondensatorach pracujących w warunkach poniżej temperatury otoczenia. Najnowsze badania zaowocowały syntezą wspomnianej wcześniej rodziny soli typu Hückela, które znalazły zastosowanie w elektrolitach akumulatorów litowo- i sodowo-jonowych (Rys. 2).

Sole Hückela: sole, których organiczne aniony spełniają tzw. regułę Hückela. Mają charakter aromatyczny, względnie duże promienie oraz zdelokalizowany ładunek elektryczny. Tego typu sole nie były komercyjnie dostępne. Należało je wymyślić i zrobić.

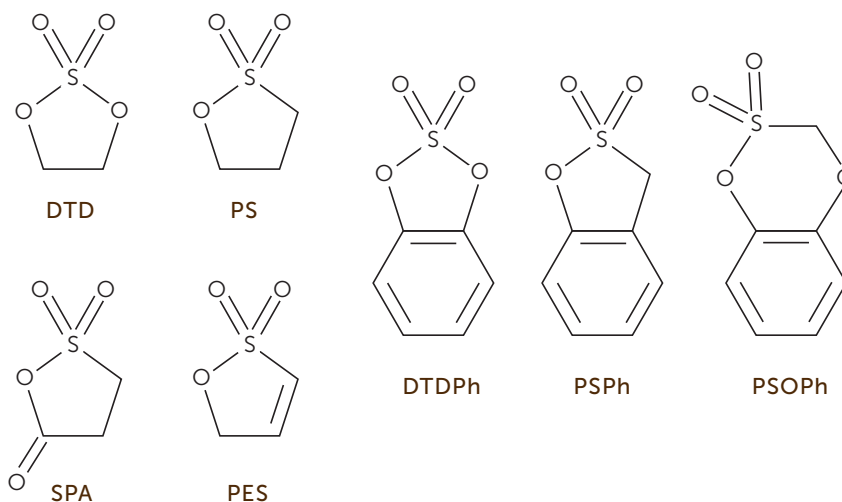
Sole Hückela są solami przewodzącymi opartymi na anionach, które osłabiają oddziaływanie koordynacyjne pomiędzy anionem i kationem, co powoduje zwiększenie przewodności jonowej i liczby przenoszenia kationów. Sole te są dodawane do elektrolitu jako pojedynczy dodatek lub w mieszaninie z innymi przyjaznymi dla środowiska składnikami zawierającymi organiczne związki siarki [3]. Te ostatnie również zostały opracowane na Wydziale Chemicznym PW (Rys. 3).

Dodanie soli Hückela do elektrolitu wywołuje również wiele innych pozytywnych skutków. Zwiększa zakres stabilności temperaturowej w porównaniu do komponentów już dostępnych na rynku. Zmniejsza się tendencja do tworzenia asocjatów i wzrasta tolerancja na wodę. Sole te działają również jako modyfikatory SEI i wewnętrzne środki

suszące. Mogą być bardzo skutecznym pochłaniaczem nie tylko śladów wody, ale także fluorowodoru (HF), czego wynikiem jest znaczne wydłużenie żywotności akumulatora i poprawa bezpieczeństwa jego pracy [4,5]. Ogranicza to ilości komponentów potrzebnych do produkcji akumulatorów, co jest wymogiem zrównoważonego rozwoju branży akumulatorowej. Możliwe jest również stosowanie nowatorskich materiałów elektrodowych, a to w perspektywie daje szansę poprawy bezpieczeństwa oraz zwiększenia pojemności i gęstości energii. Powoduje to również co najmniej pięciokrotne zmniejszenie ilości fluoru stosowanego w elektrolitach w porównaniu do istniejących układów i to bez pogarszania parametrów użytkowych akumulatorów. Prowadzi to nie tylko do obniżenia kosztów produkcji, ale także ogranicza użycie szkodliwego dla

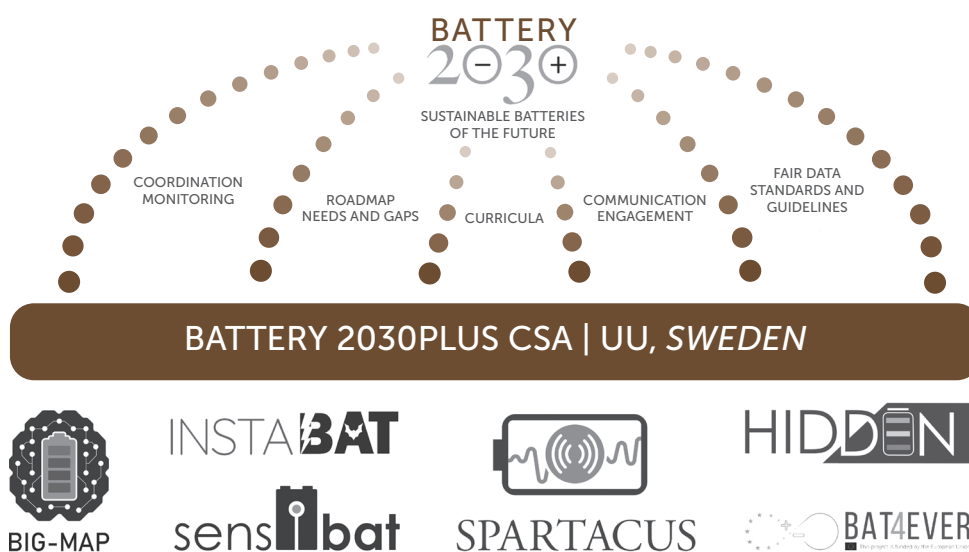


↑ Rys. 2 M. Armand et al., „Development of Hückel Type Anions: From Molecular Modeling to Industrial Commercialization. A Success Story”, *J. Electrochem. Soc.* (2020)167 070562



↑ Rys. 3 Dodatki zawierające siarkę

LARGE-SCALE RESEARCH INITIATIVE



↑ Rys. 4 „Wymiar społeczny” baterii

środowiska fluoru. Komponenty typu soli Hückela są aktualnie stosowane w wielu procesach wytwarzania akumulatorów „postlitowo-jonowych”, takich jak sodowo-jonowe, litowo-siarkowe czy litowo-powietrzne.

Należy zauważyć iż przez kolejne lata pracowaliśmy nad podwyższeniem przewodności polimerowych elektrolitów, niestety bez istotnego, a właściwie, wystarczającego przełomu, który zapewniłby rozwiązanie problemów bezpieczeństwa energetycznego przy zrównoważeniu środowiskowym. Nadal, ani nam, ani też nikomu innemu nie udało się opracować stałego

elektrolitu polimerowego, który miałby wystarczająco wysoką przewodność i stabilność umożliwiające praktyczne zastosowanie w elektrochemicznych źródłach prądu. Niemniej, powstał duży zespół naukowy pod kierownictwem prof. Władysława Wieczorka, publikacje, doktoraty, habilitacje, profesury i olbrzymi potencjał współpracy z najlepszymi zespołami na świecie. Obecnie Wydział Chemiczny Politechniki Warszawskiej jest partnerem w projekcie 4 Horyzont 2020, pełniąc kierowniczą rolę w opracowywaniu nowatorskich receptur elektrolitów. Podjęta została współpraca z polskim

partnerem przemysłowym PKN Orlen. Naszym celem jest wykorzystanie produktów ubocznych lub zapasowych pochodzących z innych procesów chemicznych realizowanych w zakładach przemysłowych. Są to – siarka i sadza z przemysłu rafineryjnego, oligoglikole i węglany cykliczne z PKN Orlen czy dwutlenek tytanu z przemysłu nawozów rolniczych.

Surowcowe bariery rozwoju

Rozwój ogniw, baterii oraz magazynowanie energii w skali makro (elektromobilność, duże magazyny) napotyka poważne zagrożenie niedoboru surowców. Przewiduje się, że globalny popyt na surowce do produkcji akumulatorów, takie jak nikiel, grafit i lit, w 2040 roku wzrośnie kilkunastokrotnie w porównaniu do roku 2020. Chiny pozostaną głównym dostawcą surowców do produkcji baterii w 2030 roku, chociaż globalne dostawy tych materiałów będą coraz bardziej zróżnicowane.

Jednak niedobory podaży pozostaną. W perspektywie średnioterminowej spodziewane są deficyty litu już w obecnym czasie, podczas gdy zrównoważenie podaży i popytu w przypadku niklu, grafitu i manganu będzie trwało do roku 2030. Oczekuje się, że do 2035 roku produkcja ogniw akumulatorowych w UE pokryje zapotrzebowanie na pojazdy elektryczne i magazynowanie energii. Jest jednak prawdopodobne, że UE będzie w różnym stopniu uzależniona od importu surowców niezbędnych do produkcji poszczególnych komponentów

„Nadal, ani nam, ani też nikomu innemu nie udało się opracować stałego elektrolitu polimerowego, który miałby wystarczająco wysoką przewodność i stabilność umożliwiające praktyczne zastosowanie w elektrochemicznych źródłach prądu”

„Baterie są kluczowym czynnikiem umożliwiającym wzrost bezpieczeństwa energetycznego, zmniejszenie negatywnego wpływu technologii na środowisko w różnych obszarach...”

elektrolitu – soli i dodatków stabilizujących granice faz elektroda-elektrolit.

Australia i Kanada to kraje o największych możliwościach zapewnienia UE dodatkowych dostaw o względnie umiarkowanym ryzyku, dla prawie wszystkich surowców niezbędnych do produkcji baterii. Zwiększenie obiegu zamkniętego w użytkowaniu komponentów daje szansę na zmniejszenie zależności UE od dostaw. Szacuje się, że do 2040 roku, recykling może zaspokoić odpowiednio 51% i 42% zapotrzebowania UE na kobalt i nikiel.

Strategicznym celem Komisji Europejskiej jest transformacja do społeczeństwa neutralnego pod względem emisji dwutlenku węgla, poprzez redukcję emisji gazów cieplarnianych netto o co najmniej 55 procent do 2030 roku.

Wymaga to używania akumulatorów o bardzo wysokiej wydajności, wykraczającej poza obecne możliwości. Gęstość energii i mocy musi zbliżyć się do teoretycznych limitów, a także należy osiągnąć wyjątkową żywotność i niezawodność, w powiązaniu ze zwiększonym bezpieczeństwem i zrównoważeniem środowiskowym.

Baterie są kluczowym czynnikiem umożliwiającym wzrost bezpieczeństwa energetycznego, zmniejszenie negatywnego wpływu technologii na środowisko w różnych obszarach i przekształcenie społeczeństwa w neutralne dla klimatu, przy jednoczesnym tworzeniu nowych rynków i miejsc pracy. Przyjęta w ramach Unii Europejskiej strategia działania to tworzenie konsorcjów zrzeszających środowiska naukowe i producentów baterii. Jednym z flagowych programów unijnych, w którym Politechnika Warszawska aktywnie uczestniczy, jest program Battery 2030+ (Rys. 4). Kluczowym założeniem tego programu jest zdefiniowanie strategii produkcji nowych generacji wysokowydajnych, bezpiecznych i zrównoważonych baterii w Europie. Te nowatorskie technologie akumulatorowe będą miały wpływ na społeczeństwo i środowisko na wielu poziomach.

Podejście oparte na współpracy w ramach projektu Battery 2030+ daje silny impuls do wspólnego działania na rzecz Europy. Dzięki zintegrowanemu podejściu na dużą skalę nasze zasoby zostaną jak najlepiej wykorzystane i przyspieszą innowacje. Taka inicjatywa jest również potrzebna, aby przyciągnąć talenty i kompetencje niezbędne do osiągnięcia celów technicznych oraz wesprzeć europejski przemysł wykwalifikowaną siłą roboczą. Programy edukacyjne i informacyjne są zatem istotną częścią aspektu społecznego tej inicjatywy.

BIBLIOGRAFIA

- [1] P. Jankowski, W. Wiczeorek, P. Johansson, Functional ionic liquids: Cationic SEI-formers for lithium batteries, *Energy Storage Materials*, 2019, 20, 108-117
- [2] M. Armand, P. Johansson, W. Wiczeorek, Review – Development of Huckel type anions. From molecular modeling to industrial commercialization. A success story, *J. Electrochem. Soc.* 2020, 167, 70562
- [3] C. Wölke, A. Benayad, T.-L. Lai, F. Hanke, G. Baraldi, M. Echeverría, E. Esen, E. Ayerbe, A.R. Neale, J. Everitt, L.J. Hardwick, P. Yan, M. Poterata, W. Wiczeorek, M. Winter, I. Cekic-Laskovic, Single Versus Blended Electrolyte Additives: Impact of a Sulfur-Based Electrolyte Additive on Electrode Cross-Talk and Electrochemical Performance of LiNiO₂||Graphite Cells, *Advanced Energy Materials*, 2024, 202402152
- [4] F. Lindgren, Ch. Xu, J. Maibach, A.M. Andersson, M. Marcinek, L. Niedzicki, T. Gustafsson, F. Bjorefors, K. Edstrom, A hard X-ray photoelectron spectroscopy study on the solid electrolyteinterphase of a lithium 4,5-dicyano-2-(trifluoromethyl)imidazolidbased electrolyte for Si-electrodes, *J. Power Sources* 2016, 301, 105-112
- [5] F. Lindgren, Ch. Xu, L. Niedzicki, M. Marcinek, T. Gustafsson, F. Bjorefors, K. Edstrom, R. Younesi, SEI Formation and Interfacial Stability of a Si Electrode in a LiTfDl-salt Based Electrolyte with FEC

and VC Additives for Li-ion Batteries, *ACS Applied Materials & Interfaces* 2016, 8, 15758-15766

Artykuł autorstwa prof. dr hab. inż. Władysława Wiczeorka, prof. dr hab. inż. Marka Marcinka oraz prof. dr hab. inż. Janusza Plocharskiego (Wydział Chemiczny PW) – powstał na podstawie seminarium, pod tym samym tytułem, które odbyło się w cyklu Horyzonty Dyscyplin Nauki, w 2024 roku.

Konferentami seminarium byli prof. dr hab. Paweł Kulesza (Wydział Chemii UW) oraz dr hab. inż. Krzysztof Fic (Wydział Technologii Chemicznej, Politechnika Poznańska). W wydarzenie poprowadził prof. dr hab. inż. Tadeusz Hofman (Wydział Chemiczny PW).

{ Profesor Władysław Wiczeorek

– szef zespołu na Wydziale Chemicznym PW, od wielu lat zajmującego się badaniami nad wytwarzaniem nowoczesnych chemicznych źródeł prądu, w szczególności baterii. Jego działalność naukowa skupia się na syntezie i badaniu właściwości szerokiej gamy polimerowych elektrolitów pod kątem ich zastosowania w urządzeniach elektrochemicznych, takich jak: baterie litowe i litowo-jonowe, ogniwa paliwowe, układy elektrochromowe. Pełnił wiele prestiżowych funkcji i piastował ważne stanowiska na uczelni. Był prodziekanem, a później dziekanem Wydziału Chemicznego PW, w latach 2008-2016 prorektorem ds. studenckich PW. Odbył kilka naukowych staży zagranicznych m.in. w University of Guelph, Kanada (1993-1996), Chalmers University, Goeteborg (1992), CNR TAE Messina (1989). Od listopada 2018 roku jest członkiem Rady Naukowej (Battery 2030+) przy European Battery Alliance, w której współprzewodniczy zespołowi odpowiedzialnemu za syntezę nowych materiałów i komponentów baterijnych w programie Batflag realizowanym w ramach platformy Battery 2030+ }

Więcej informacji →



WARSZTATY ŚCISŁEGO MYŚLENIA

Poznawanie świata determinowane jest przez działanie narządu myślenia. Dobra kondycja tego narządu, szczególnie u ludzi nauki, wymaga ciągłego rozwijania i ćwiczenia myślenia ścisłego opartego na matematycznej pra-logice naszego języka

W dniach 22-24 marca 2024 r. w Europejskim Centrum Edukacji Geologicznej Uniwersytetu Warszawskiego odbyły się Warsztaty Ścisłego Myślenia. Zostały zorganizowane przez Centrum Studiów Zaawansowanych PW we współpracy z Centrum Badania Ryzyka Systemowego UW ze szczególną myślą o studentach indywidualnych studiów PW.

Ideą przewodnią wyjazdu było przedstawienie tematyczne odczytów przeznaczonych dla szerokiego merytorycznie kręgu uczestników. Otwartych na prezentowanie własnych przemyśleń i badań z uwzględnieniem różnych specjalizacji wśród słuchaczy. Szczególny nacisk położony został na zagadnienia złożoności systemów oraz z pogranicza dziedzin humanistycznych i matematyki, z odniesieniem do modelowania zjawisk w naukach przyrodniczych i społecznych.

Wygłoszone odczyty:

Stanisław Janeczko, *O ścisłym myśleniu w nauce*; Julia Le Bihan, *Dlaczego talie kart należy tasować 7 razy?*; Adam Zuba, *Diody typu OLED oparte na emiterach TADF*;

Michał Dębski, *Matematyka w rezerwacji biletów kolejowych*; Dariusz Baran, *Metoda syntezy hybryd porfiryńowo-cukrowych za pomocą reakcji siła-Sonogashiry*; Marek Kuś, *Topologia w analizie sztuki abstrakcyjnej*; Natalia Grzybicka, *Zbadanie zależności fluktuacji krotności cząstek zidentyfikowanych od prędkości w zderzeniach p+p przy prędkości wiązki 158 GeV/c zebranych przez eksperyment NA61/SHINE*; Agnieszka Jastrzębska, *Zapewnianie jakości w procesie wytwarzania oprogramowania wykorzystującego uczenie maszynowe*; Zofia Kubrak, *Wykorzystanie metaheurystyk inspirowanych naturą w optymalizacji procesów*; Marta Piecyk, *Rekonstrukcja grafów*; Bartosz Godlewski, *Nowe podejście do funkcjonalizacji labilnych w środowisku zasadowym pochodnych cukrowych i porfiryńowych na drodze reakcji Suzukiego-Miyaura promowanej Ag₂O*; Beata Bajno, *Wędrownica po miejscach niedookreślenia*; Bartłomiej Łuszczuk, *„Młody i dynamiczny zespół” badawczy – jak zostać naukowym szefem i się nie przekreślić*; Jacek Rogala, *Temperatura mózgu, płeć i pamięć i co z tego wynika*; Mikołaj Spytek, *Ukryte reprezentacje danych w uczeniu maszynowym i co można z nimi robić*; Piotr Bentkowski, *Modelowanie ekosystemu – wpływ farmaceutyków na system*

troficzny jezior; Józef Koszewski, *Dlaczego i jak wykorzystywać Sztuczną Inteligencję w przemyśle wydobywczym?*; Grzegorz Łach, *Pandemic on a graph*; Mateusz Krzyżiński, *Jak wyjaśniać decyzje sztucznej inteligencji? Metody SHAP oraz SurvSHAP(t)*; Hubert Grochowski, *O L(2,1)-etykietowaniu grafów słów kilka*; Zuzanna Górecka, *Cyberbezpieczeństwo przemysłowych systemów sterowania – nowe wyzwania diagnostyki procesów przemysłowych*; Paweł Naroski, *Kolorowe ścieżki*; Zofia Nowicka, *Technologia w modzie*; Jakub Modrzewski, *Wykorzystanie technologii Unreal Engine do budowy Cyfrowych Bliźniaków*; Jakub Włodarczyk, *Zbadanie zależności fluktuacji krotności cząstek zidentyfikowanych od energii zderzenia w zderzeniach p+p zebranych przez eksperyment NA61/SHINE*; Bartłomiej Jan Sobieski, *Modele dyfuzyjne w modelowaniu generatywnym*; Marcin Ziółkowski, *Przewidywanie zakłóceń wpływających na realizację usług sieci 5G działających w wirtualnym środowisku*; Katarzyna Gorczyca, *Synteza porfiryń o zwiększonej hydrofobowości – rozwój nowej generacji fotouczulaczy przeciwnowotworowych*.

Zespół CSZ

II WARSZTATY KONCEPCYJNE W CHĘCINACH

Warsztaty CSZ PW

Ponownie w Europejskim Centrum Edukacji Geologicznej UW w Chęcinach i we współpracy z Centrum Badania Ryzyka Systemowego UW, w dniach 10-12 maja 2024 r. odbyły się II Warsztaty Konceptyjne Centrum Studiów Zaawansowanych Politechniki Warszawskiej. Ich celem było omówienie pewnych kluczowych i bardzo aktualnych problemów, takich jak wyzwania i zagrożenia związane z rozwojem AI, a także dyskusja nad tymi zagadnieniami z perspektywy różnych dziedzin nauki – zarówno ścisłych jak i humanistycznych. Reprezentujący różne dziedziny nauki członkowie Centrum Badania Ryzyka Systemowego UW mieli okazję porozmawiać i podyskutować z wybitnymi matematykami z Centrum Studiów Zaawansowanych PW oraz humanistami, m.in.

z Instytutu Historii im. Tadeusza Manteuffla PAN.

Tematy wystąpień:

- *Dobrostan – w świetle rozważań o ergonomii i samoświadomości* – Barbara Majerska;
- *Dwa spojrzenia na układy złożone* – Stanisław Janeczko;
- *Klasyczne metody statystyczne czy uczenie maszynowe – czy naprawdę trzeba wybierać?* – Jacek Rogala;
- *Sztuczna inteligencja i big data w ekologii – jak automatyczna analiza obrazu zmienia badania zachowania zwierząt* – Piotr Bentkowski;
- *Czy uczony jest artystą? Rzecz o pięknie w naturze, sztuce i nauce* – Stanisław Bajtlik;
- *Ludzka mądrość i sztuczna inteligencja* – Wiesław Bartkowski;

- *Zastosowanie AI w prospekcji stanowisk archeologicznych: czy jest to możliwe?* – Fabian Welc;
- *Co generatywna AI robi nauce* – Sebastian Szymański;
- *Boty i reklama w Internecie* – Agnieszka Jastrzębska;
- *O projektowaniu sekwencji RNA* – Grzegorz Łach;
- *„Niewidzialny towar”? Metody przyrodnicze w badaniach nad średniowiecznym niewolnictwem* – Marek Jankowiak.

Zespół CSZ

DYDAKTYKA DLA POZNANIA, ROZUMIENIA I DZIAŁANIA

Symposium podsumowujące 3 edycje programu Grantów Dydaktycznych PW



W dniach 1-4 września 2024 w Ośrodku Wypoczynkowy PW w Wildze odbyło się symposium „Dydaktyka dla poznania, rozumienia i działania” poświęcone odkrywaniu nowych możliwości, ale również identyfikowaniu słabych punktów oraz wyzwań w procesach dydaktycznych uczelni wyższych. Pod rozważania poddano zagadnienia dotyczące tego, jak efektywnie kształcić – uczyć i aktywizować naukowo współczesnego studenta – jak przeprojektowywać procesy i pokonywać trudności, które stawia przed dydaktyką akademicką złożona i niejednorodna rzeczywistość społeczna. Główny trzon uczestników Symposium stanowili laureaci konkursów na granty dydaktyczne w PW – członkowie zespołów, którzy dostrzegli potencjał zmian i podjęli wysiłek unowocześnienia sfery dydaktycznej Politechniki Warszawskiej.

Trwający od trzech lat program Grantów Dydaktycznych unaoczniał i wypuklił wiele słabości w dydaktyce z jakimi boryka się uczelnia. Pokazał również drzemiący w wielu zespołach badawczych, potencjał kreatywnego kształcenia i podejmowania nowych, niełatwych zagadnień współczesnej nauki. Wśród trzech aspektów kształcenia: informacyjnego, projektowego i problemowego, dominującą formą jest często jednostronne informowanie uczestników procesu dydaktycznego. Z kolei poziom zaangażowania projektowego nie wydobywa dostatecznie

zaawansowanych twórczych możliwości. Balans pomiędzy „nowatorskimi metodami”, które same w sobie mogą zdominować cały proces, a żmudnym „tradycyjnym nauczaniem”, jest trudny do odnalezienia. Należy również pamiętać, jak ważna i złożona jest dydaktyka problemowa. Wymaga wysokiego stopnia indywidualizacji w kształceniu i wyjątkowo autorskich metod.

Symposium było okazją do wymiany pomysłów oraz dyskusji. Swoje osobiste doświadczenia i przemyślenia zaprezentowali m.in. laureaci Grantów Dydaktycznych. Uczestnicy wysłuchali następujących wystąpień:

- Krzysztof Ejsmont, *Cyfrowy bliźniak (digital twin) w dydaktyce – możliwości i zastosowanie*;
 - Przemysław Korpas, *Bezpieczeństwo Internetu Rzeczy nauczane metodą PBL – wnioski po pierwszej realizacji, wyzwania i proponowane rozwiązania*;
 - Andrzej Kraśniewski, *Innowacyjne studia drugiego stopnia „Inżynieria Internetu Rzeczy” – pierwsze doświadczenia*;
 - Daniel Paczesny, *Living Lab w Wildze – połączenie dydaktyki i nauki*;
 - Andrzej Kraśniewski, *Eksperyment dydaktyczny: przedmiot o nieokreślonej zawartości treściowej*;
 - Daniel Paczesny, *Narzędzia rozwijane przez OKNO PW w obszarze kształcenia na odległość*;
 - Małgorzata Jakubowska, *Utworzenie nowego przedmiotu pt. Elektronika drukowana jako przedmiotu obieralnego dla studentów II stopnia*;
 - Ewa Ura-Bińczyk, *Wykorzystanie narzędzi coachingu grupowego w dydaktyce*;
 - Kamila Sałasińska, Joanna Ryszkowska, *Ocena korzyści kształtowania wiedzy studentów Wydziału Inżynierii Materiałowej w obszarze środowiska za pośrednictwem nowego przedmiotu: Projektowanie wyrobów według cyklu życia*;
 - Jerzy Pokojski, *Engineering Knowledge Management – nauczanie akademickie w kontekście inicjatywy Industry 4.0*;
 - R. Robert Gajewski, *Dylematy i wyzwania współczesnej edukacji*;
 - Anna Kalbarczyk, *Utworzenie stanowisk projektowych z obszaru magazynowania i przetwarzania energii elektrycznej w zgodzie z priorytetami proekologicznej infrastruktury technicznej*;
 - Jakub Możaryn, *Innowacyjne Metody Dydaktyczne i Technologie w Edukacji Inżynierskiej*;
 - Michał Zwierzyński, *Ewolucja sposobu nauczania matematyki akademickiej*.
- Mamy nadzieję, że wysłuchanie prezentacji pozwoliło lepiej zrozumieć zjawiska, mechanizmy i problemy obecne w sferze dydaktyki, a także sformułować wnioski praktyczne służące lepszej i bardziej efektywnej realizacji misji dydaktycznej Uczelni.

Zapewnianie jakości oprogramowania z uczeniem maszynowym

czyli kilka aspektów tworzenia i wdrażania oprogramowania z elementami uczenia maszynowego

„...Systemy, o których mowa w swojej istocie są niczym innym jak oprogramowaniem. Pojawiają się więc dwa pytania: jak określić jakość oprogramowania w systemach z uczeniem maszynowym oraz jak zapewnić tę jakość?”

Uczenie maszynowe bije rekordy popularności. Postęp tej dziedziny widzimy zarówno w coraz to doskonalszych algorytmach, jak i w pojawianiu się nowych obszarów zastosowań. Obserwowany rozwój jest szybki i fascynujący – na przecięciu matematyki i informatyki funkcjonuje obszar wiedzy, o którym można powiedzieć, że wpływa (prawie) na każdego.

Pojawia się też wiele głosów skrajnych wyrażających wysoki entuzjazm lub duże obawy o dalszy rozwój tej dziedziny i jej wpływ na funkcjonowanie człowieka.

W artykule przedstawię wybrane zagadnienia, które w mojej opinii istotnie wpłyną na dalsze losy rozwoju tej dziedziny. Są to: jakość danych, weryfikacja jakości modelu oraz tworzone akty prawne.

Powszechność uczenia maszynowego

Jeśli przyjmiemy, że jednym z kryteriów oceny pomysłu i rozwiązania jest ich użyteczność, przyznać należy, że algorytmy uczenia maszynowego wybierają się w aktualnie rozwijanej

informatyce. Społeczeństwo chętnie akceptuje użycie algorytmów przetwarzania dźwięku, obrazu i innych sygnałów, doceniając oszczędności zasobów ponoszonych w miejscach, w których wdrożono takie rozwiązania. Algorytmy rozpoznawania wzorców pozwalają odkryć nową wiedzę oraz usprawnić lub całkowicie zautomatyzować wiele procesów. Podając przyziemne przykłady wspomnę o systemach wjazdowych na parkingi, które rozpoznają tablice rejestracyjne lub o systemach rozpoznawania mowy w infoliniach. Niezmiernie ważne obszary zastosowań uczenia maszynowego to medycyna, chemia, bezpieczeństwo i inne. Algorytmy wspomagają lekarzy w podejmowaniu decyzji, rekomendują terapie, prognozują stan zjawisk w przyszłości, monitorują dane sejsmiczne, alarmują o zagrożeniach, sterują elektrowniami, sterują urządzeniami produkcyjnymi, sprawdzają jakość materiałów itd. Zastosowań w obszarach krytycznych oraz tych mniej krytycznych jest i będzie wiele, ale nie to jest tematem tego artykułu.

Korzystając świadomie z rozwiązań opartych o uczenie maszynowe, zaczynamy dostrzegać, że nie zawsze działają poprawnie. Zauważamy też pewne regularności w błędach, czyli powtarzające się sytuacje, w których oprogramowanie nie działa poprawnie, np. wpływ oświetlenia na skuteczność rozpoznania obrazu. W pewnych sytuacjach uciążliwość i permanentność błędów przestają być akceptowalne – tu jako przykład podam rozmowy z chatbotami, do których jesteśmy zmuszani w pierwszym kroku przy połączeniach z niektórymi serwisami obsługi klienta.

Jakość(ć) to będzie...

Dotykamy sedna problemu – jakości rozwiązań opartych o uczenie maszynowe. Systemy, o których mowa w swojej istocie są niczym innym jak oprogramowaniem. Pojawiają się więc dwa pytania: jak określić jakość oprogramowania w systemach z uczeniem maszynowym oraz jak zapewnić tę jakość?

Co wynika z poprzedniego paragrafu, temat jakości uczenia maszynowego analizuję w kontekście szerszym, gdzie uczenie maszynowe stanowi (często najważniejszy) komponent pewnego systemu. Dla kontrastu, najwęższy możliwy kontekst zakłada sprowadzenie tego procesu do wytrenowania klasyfikatora na gotowym pliku z danymi o postaci danych tabelarycznych, gdzie wiersze to kolejne obserwacje, a kolumny odpowiadają kolejnym atrybutom (cechom). Taki wąski kontekst nie przystaje jednak do rzeczywistych systemów, w których funkcjonuje oprogramowanie z uczeniem maszynowym. Zapewnianie jakości oprogramowania jest dziedziną informatyki, bogatą w standardy, dobre praktyki i różne gotowe narzędzia. Oparcie się na takich standardach jest wygodne. Poprawnie przeprowadzone procesy zapewniania jakości oprogramowania pozwolą potwierdzić, że wytwarzane oprogramowanie spełnia stawiane

mu wymagania. Niesie to też za sobą korzystanie z powszechnie przyjętej terminologii, co ułatwia porozumiewanie się z interesariuszami.

Oprogramowanie z elementami uczenia maszynowego różni się od „klasycznego” oprogramowania (dla ustalenia uwagi, jako przykład klasycznego oprogramowania podam serwis internetowy do składania zamówień w sklepie). Wynika to ze specyfiki tworzenia modeli uczenia maszynowego. Jest też wiele niuansów, które są trudniejsze do doprecyzowania na etapie zbierania wymagań. Stąd klasyczne praktyki zapewniania jakości oprogramowania nie do końca przystają do procesów, które następują, gdy dochodzi uczenie maszynowe.

Powyższa obserwacja nie jest odkrywcza. Konieczność stworzenia dedykowanej metodyki zapewniania jakości oprogramowania z uczeniem maszynowym została już rozpoznana. W 2021 roku *International Software Testing Qualifications Board* opublikowało materiały o nazwie *Certified Tester - AI Testing (CT-AI)*. Było to pierwsze tego typu podejście do systematyzacji w tej dziedzinie. Zakłada, że CT-AI jest realizowane w ramach ścieżki specjalistycznej testerów oprogramowania, czyli osoba realizująca CT-AI już posiada kompetencje podstawowe związane z testowaniem oprogramowania. Uznając pionierskość CT-AI, trzeba jednak zaznaczyć, że metodyka została zdmonowana wątkami związanymi z sieciami neuronowymi, a uczenie maszynowe to przecież o wiele więcej.

Procesy nakierowane na zapewnianie jakości stanowią tylko fragment całości czynności wykonywanych przy rozwoju oprogramowania. Patrząc szerzej na inżynierię oprogramowania wspomnieć należy o metodyce wspierającej znaną pod nazwą *Development and Operations* (DevOps), odnośnie której również poczyniono pierwsze kroki mające na celu adaptację do systemów z uczeniem maszynowym. Tu dochodzimy do obszaru znanego jako MLOps.

Sposób organizowania procesów MLOps wpływa na jakość produkowanego oprogramowania. W tym obszarze metodycznych podejść jest więcej. Przedstawię jedno, przykładowe podejście o nazwie CRISP-DM (skrót od *Cross Industry Standard Process for Data Mining*)¹. Zakłada ono budowanie rozwiązań w systemie iteracyjnym, a jego zarys został przedstawiony na rysunku 1.

Pierwszym krokiem tego procesu jest zapoznanie się z celami oraz uwarunkowaniami biznesowymi lub naukowymi rozważanego zagadnienia. Kolejny krok to zrozumienie danych. Zakłada ono wstępne zapoznanie się z posiadanymi danymi. Na tym etapie mogą nie być one w całości gotowe. Celem prac jest określenie czy dane są odpowiedniej jakości by wykonać dalsze kroki. Sprawdzamy też podstawowe właściwości danych na próbie. Jeśli weryfikacja nie potwierdzi oczekiwanych właściwości danych musimy cofnąć się o krok wcześniej i ponownie omówić z interesariuszami uwarunkowania i cele projektu. Jeśli posiadane dane przejdą pomyślną weryfikację, to przystępujemy do właściwego przygotowania danych. Jest to na przykład ręczna anotacja takiej liczby obrazów, dla której zakładamy, że budowa modelu się powiedzie. Po przygotowaniu aktualnej wersji danych przystępujemy do budowy modeli. Na tym etapie często posiadane dane dzielimy na zbiory treningowe, walidacyjne i testowe. Weryfikujemy jakość modelu na danych testowych i jeśli jakość nie spełnia założeń, to możemy wrócić do etapu ponownego opracowania danych.

Jeśli model osiągnie wymagane cechy jakościowe na zbiorze testowym (lub w analogicznej procedurze, np. walidacji krzyżowej) przechodzimy do całościowej ewaluacji modelu. W tym kroku odchodzimy od danych o charakterze laboratoryjnym (to pojęcie zostanie wyjaśnione w kolejnej podsekcji). Rozpoczyna się analiza rozwiązania dla danych zbieranych w środowisku niekontrolowanym, czyli

w sytuacjach, które będą miały miejsce jak przekażemy produkt klientowi. Pomiary te często wymagają ręcznej pracy. Po kroku ewaluacji może okazać się, że nasz projekt jest gotowy do wdrożenia (spełnia stawiane mu wymagania). Jeśli tak nie jest, cały proces rozpoczyna się od nowa.

Dane

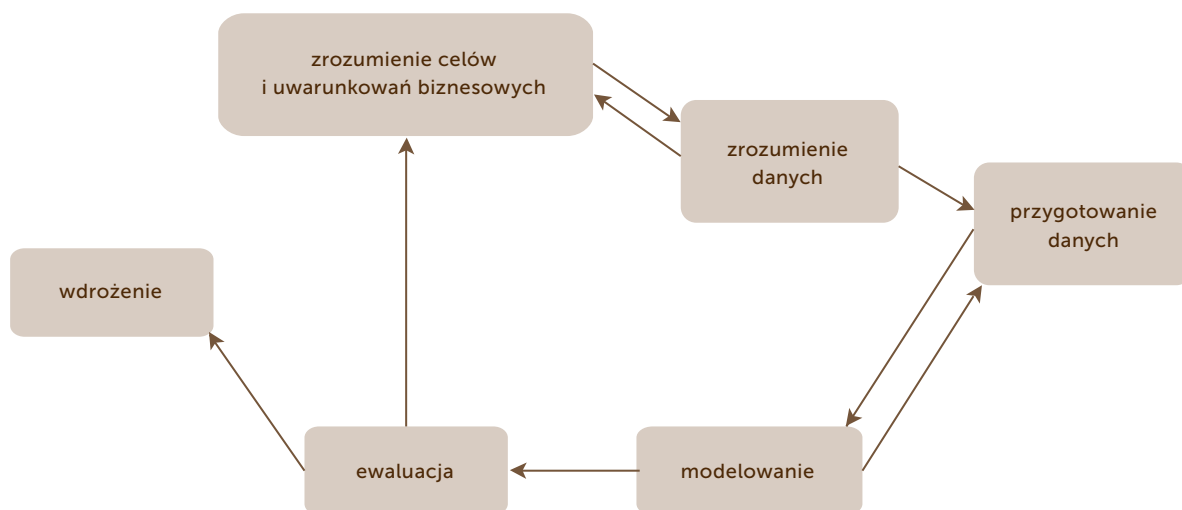
Omawiając założenia metodyki CRISP-DM wspomniałam, że z jej punktu widzenia dane mają różne postaci. Dane laboratoryjne były podstawą budowania i wstępnej weryfikacji jakości modeli. Dane zbierane w środowisku niekontrolowanym są podstawą do wykonania kroku ewaluacji. Odnosząc to do zadań wizji komputerowej, dane laboratoryjne to często dane uprzednio wyselekcjonowane według pewnego klucza. W szczególności mogą to być wysegmentowane obrazy z anotacjami. Są to dane z etykietami lub dane bez etykiet (nie wszystkie zadania wizji komputerowej mają identyczne założenia), ale zbierane w sposób intencyjny. Dane zbierane w środowisku niekontrolowanym nie są selekcjonowane. Odstępstwa mogą określać wymagania biznesowe.

Dane zbierane w środowisku niekontrolowanym często zaskakują. Pojawiają się przypadki, których osoby pracujące na poprzednich etapach nie przewidziały. Oczywiście w niektórych zastosowaniach dane zbierane w środowisku niekontrolowanym są bardzo zbliżone do danych laboratoryjnych. Zdarza się też, że dane są wyłącznie zbierane w środowisku laboratoryjnym.

(13)

„Dryfujące dane, czyli w języku angielskim *concept drift*, to zjawisko polegające na tym, że wraz z upływem czasu zmieniają się właściwości danych... Jest to pewna cecha danych, która musi być brana pod uwagę przy konstrukcji modeli”

¹ <https://www.datascience-pm.com/crisp-dm-2/>



↑ Rys. 1 Schemat metodyki CRISP-DM

Taki przypadek często występuje w kontekście badań naukowych.

Dane – ich poprawność oraz zawartość

Zgodnie z nieco wytartym frazesem funkcjonującym w języku angielskim *garbage in, garbage out*, czyli śmieci na wejściu – śmieci na wyjściu, podstawą jakości systemów uczenia maszynowego jest jakość danych użytych do ich budowy. Zagadnienie jakości danych jest niewątpliwie subtelne, ponieważ dotyczy procesów w świecie rzeczywistym.

Pierwsza, oczywista własność danych to ich reprezentatywność. Załóżmy, że chcemy rozróżniać koty, psy i ludzi na zdjęciach za pomocą splotowej sieci neuronowej tworzonej przez nas od podstaw. By wykonać to zadanie, musimy posiadać oetykietowany zbiór zawierający zdjęcia z ludźmi, psami i kotami. Jeśli w naszym zbiorze zabraknie psów pewnej rasy, różniacej się bardzo od pozostałych psów w naszym zbiorze, możemy nie być w stanie rozpoznawać takiego przypadku. Jeśli wszystkie ujęcia kotów, psów i ludzi będą wykonane za pomocą aparatu ustawionego naprzeciwko obiektu w dobrym oświetleniu i wspaniałych warunkach pogodowych, może się okazać, że zmiana kąta, złe warunki oświetlenia lub pogody uniemożliwią poprawne rozpoznanie. Jeśli dane laboratoryjne będą zbierane tak, że różne psy będą zawsze fotografowane obok brązowej budy, różne koty zawsze obok czerwonego drapak, a

ludzie zawsze na tle zielonego drzewa, może się okazać, że zamiast ludzi, psów i kotów, wytrenowany model rozpoznaje kolory zielony, brązowy i czerwony w tle.

Dryfujące dane, czyli w języku angielskim *concept drift*, to zjawisko polegające na tym, że wraz z upływem czasu zmieniają się właściwości danych. Przyczyny mogą być zarówno po stronie zjawiska, o którym zbieramy dane, jak i systemu urządzeń pomiarowych, które z czasem mogą zmieniać właściwości. Jest to pewna cecha danych, która musi być brana pod uwagę przy konstrukcji modeli.

Właściwości danych, które muszą być rozważane przy konstrukcji modeli to również ich zawartość. W niektórych krajach istnieją normy prawne regulujące sposób przetwarzania niektórych informacji. W krajach Unii Europejskiej wspomnieć należy o tzw. *EU Artificial Intelligence Act*².

W niektórych dziedzinach istotnym zagadnieniem jest występowanie stronniczości modelu (ang. *bias*). Problem ten związany jest z oparciem automatycznego procesu decyzyjnego na jednej lub więcej zmiennych o charakterze chronionym. Definicja zmiennej chronionej zależy od zagadnienia. Mogą to być na przykład wiek, płeć, zarobki itp. Dla przykładu podam popularny w prasie brytyjskiej w roku 2020 temat podejrzenia, że algorytm użyty w szkołach do obliczenia oceny GCSE był stronniczy. Algorytmu użyto z powodu pandemii³.

Jako że moduły uczenia maszynowego przeważnie występują w szerszym systemie informatycznym, kolejnym aspektem jakości jest spójność, poprawność i kompletność danych. Dane mogą stać się obiektem ataku, np. celem kradzieży lub modyfikacji treści danych. Jednym z ostatnio dostrzeganych zagrożeń jest również wiarygodność danych. Negatywnym przykładem użycia narzędzi do generowania treści opartych o architekturę neuronalną typu transformer (ChatGPT) jest użycie fałszywych danych, również do celów budowy modeli w pracach naukowych. O problemie tym wspomiano m.in. w *Nature*⁴.

Widząc jak szerokie spektrum zagadnień może wpływać na jakość danych, nie dziwi, że procedury testowania oprogramowania z elementami uczenia maszynowego muszą je uwzględniać. Dane, tak samo jak kod źródłowy, podlegają testowaniu, wersjonowaniu oraz dokumentacji.

Model

Esencją systemów, którym poświęcony został ten artykuł są modele uczenia maszynowego. Z racji różnorodności zastosowań oraz samych modeli, trudno jest w uogólniony sposób opisać jakie operacje model wykonuje. Co za tym idzie, nie jest możliwe podanie jednego uniwersalnego sposobu testowania modeli.

Można natomiast podzielić modele na dwie kategorie: transparentne (ang. *white-box*) oraz czarne skrzynki

² <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

³ <https://www.bristol.ac.uk/cmm/research/grade/>; <https://www.bbc.com/news/education-53364485>

⁴ <https://www.nature.com/articles/d41586-023-03635-w>

(ang. *black-box*). Ta pierwsza grupa pozwala na bezpośrednią interpretację sposobu ich działania. Dla przykładu, drzewa decyzyjne są modelami transparentnymi. Ich sposób działania można opisać jako zestaw reguł „jeśli, to”. Testowanie takich modeli jest prostsze niż modeli typu czarna skrzynka.

Do wyjaśnienia sposobu działania modelu typu czarna skrzynka zastosować można odrębny model z rodziny modeli uczenia maszynowego. Nasuwa się jednak pytanie, jak sprawdzić jakość modelu wyjaśniającego – nie we wszystkich zastosowaniach jest to zasadne. Pytanie, jak dobrze testować modele typu czarna skrzynka, pozostaje otwarte i niezwykle ważne.

Niezależnie od typu modelu można opracować scenariusze testowania sprawdzające aspekty o charakterze technicznym. W szczególności, weryfikowalne jest czy po kolejnych przekształceniach danych lub reprezentacji pośredniej wymiarowość jest zgodna z założeniami oraz czy wartości należą do poprawnego zakresu. Jeśli założone przekształcenia są deterministyczne, a odwzorowanie dostępne w postaci analitycznej, można dokładnie obliczyć, czy implementacja jest zgodna z tym założeniem.

W przypadku oprogramowania z elementami uczenia maszynowego stosuje się adaptowane techniki testowania analogiczne jak do klasycznego oprogramowania. W szczególności uwagę można zwrócić na:

- testowanie A/B, które może być zastosowane do oceny i wyboru dobrego algorytmu,
- testowanie *back-to-back*, które dotyczy będzie całego potoku przetwarzania danych,
- testowanie kombinatoryczne, które będzie mogło dotyczyć np. wyboru parametrów,
- testowanie poprawności danych oraz testowanie odporności modelu na dane zaszumione,
- testy adwersarialne, mające na celu wykrycie wad,
- testy metamorficzne mające na celu zweryfikowanie zależności pomiędzy danymi na wejściu a tymi na wyjściu ⁵.

Na koniec dodam jeszcze o testach wewnętrznej struktury modelu, które są determinowane właściwościami algorytmu. W przypadku wytrenowanych sieci neuronowych sprawdzać można właściwości wartości aktywacji neuronów. Jeśli mówimy zaś o procedurze uczenia, to zweryfikować można wartości pojawiające się jako aktywacje, które mogą sygnalizować problemy z modelem.

Po czyjej stronie stanie prawo?

Ostatnim wymienionym we wstępie aspektem wpływającym na dalszy rozwój dziedziny jest strona prawna. Aktualnie brak jest kompleksowego rozwiązania w tym zakresie. Regulacje, w szczególności te określające odpowiedzialność związaną z konsekwencjami używania programów z komponentami uczenia maszynowego, wpłyną na sposób ich udostępniania oraz użycia. Na chwilę obecną w Unii Europejskiej funkcjonuje wspomniany EU AI Act. Ponieważ wątek ten będzie rozwijany, zamiast spekulacji o kierunku prac, przybliżę problematykę – z jakimi sytuacjami przyjdzie się nam zmierzyć.

Utylitarność algorytmów uczenia maszynowego opiera się na tym, że bez naszego wkładu (lub z ograniczonym wkładem) podejmują same decyzje lub proponują nam jakie decyzje podjąć. Przykład, kiedy algorytm sam podejmuje decyzje i równocześnie stanowi część systemu sterowania pewnym układem fizycznym, to rejestratory i bramki na parkingach, które rozpoznają numery na tablicach rejestracyjnych i wypuszczają tylko te samochody, za których postój wniesiona została odpowiednia opłata. O wiele powszechniejsze są narzędzia, które sugerują nam decyzję, ale samą fizyczną akcję podejmuje człowiek. Programy mogą źle działać zarówno w jednym jak i w drugim przypadku. Trudność polega na określeniu odpowiedzialności za popełniane błędy.

Przytoczę wybrane zdarzenia z przeszłości, by przybliżyć problemy, z którymi prawodawcy muszą się zmierzyć:

- Na początku roku 2024, liniom lotniczym Air Canada nakazano wypłacenie odszkodowania na rzecz osoby, która została wprowadzona

w błąd przez mało jasne sugestie chatbota. Sprawa dotyczyła zakupu biletu droższego niż tej osobie mógł przysługiwać ⁶.

- Prawnik pewnej firmy ze Stanów Zjednoczonych użył ChatGPT do opracowania argumentów w prowadzonej przez siebie sprawie. W sądzie okazało się, że argumenty przytoczone przez prawnika są fałszywe – informacje zostały przez program zmyślone ⁷. Prawnik został ukarany.
- Raportowane są przypadki potrąceń pieszego przez autonomiczne pojazdy, w tym kiedy pieszy poruszał się wbrew przepisom ruchu drogowego ⁸.
- Firma działająca w obszarze edukacji zapłaciła wysoką kwotę w ramach ugody. U podstaw sporu było korzystanie z oprogramowania, które automatycznie odrzucało kandydatki w wieku 55 lat i starsze oraz kandydatów w wieku 60 lat i starszych ⁹.

Podsumowanie

Przed dziedziną uczenia maszynowego stoi bardzo duże wyzwanie – zapewnienie jakości artefaktom tworzonym w oparciu o złożone algorytmy heurystyczne i niedeterministyczne. Specyfika dziedziny wymaga dedykowanego podejścia do testowania. Zazębiające się trzy obszary omówione w artykule: jakość danych, poprawność modelu oraz tworzone prawo są podstawą tych procesów.

15

{ Agnieszka Jastrzębska, prof. ucz. – pracownik Zakładu Strukturalnych Metod Przetwarzania Wiedzy na Wydziale Matematyki i Nauk Informatycznych Politechniki Warszawskiej. Jej zainteresowania koncentrują się wokół zagadnień informatyki, w szczególności szeroko pojętej inteligencji obliczeniowej. Pełni wiele funkcji wydziałowych jak i uczelnianych. Współorganizuje popularne seminaria z cyklu „Matematyczne i informatyczne granice algorytmiczności” }

⁵ Podstawowy słownik terminologii związanej z procedurami testowania, w tym programów z elementami uczenia maszynowego można znaleźć pod poniższym odnośnikiem: <https://sjsi.org/kategoria-slownika/specialist-ai-testing-v1-0/>

⁶ <https://www.cbc.ca/news/canada/british-columbia/air-canada-chatbot-lawsuit-1.7116416>

⁷ <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>

⁸ <https://www.bbc.com/news/technology-67133409>; <https://www.euronews.com/next/2024/07/09/chinese-driverless-car-hits-a-person-crossing-against-the-light>

⁹ <https://www.eeoc.gov/newsroom/itutorgroup-pay-365000-settle-eeoc-discriminatory-hiring-suit>

Upcycling AI

Skąd biorą się eksplodujące koszty treningu generatywnej sztucznej inteligencji i dokąd zapędzają naukowców?

Revolucja przemysłowa i/lub gorączka złota

W lutym 2024 Jensen Huang, CEO Nvidii oznajmił, że w przyszłości AI zautomatyzuje produkcję oprogramowania. I to na tyle, że zamiast używać skomplikowanych języków programowania, deweloperom jutra wystarczy będzie opisywanie swojej wizji odpowiednio przygotowanym modelom językowym.

W mediach, oczywiście, zawrzało. Jensen Huang nie był wcale pierwszym głosem porównującym odkrycia ostatnich lat do nowej rewolucji przemysłowej. Imponujące umiejętności modeli używanych do generowania tekstów oraz obrazów już zdążyły zaskoczyć opinię publiczną. Oczywiście, modele pozostawiały jeszcze wiele do życzenia – zmyślały i nie oferowały wystarczającego poziomu bezpieczeństwa, by móc je efektywnie wdrożyć¹. Trudno jednak ocenić, ile brakowało do zniwelowania tych niedoskonałości.

Oprócz sensacyjnych nagłówek pojawiły się też głosy sceptyków. Jensen Huang nie był przecież niezależną partią. Za nagłym wzrostem wartości akcji Nvidii w dużym stopniu stoi właśnie generatywna gorączka złota. Rola firmy jest porównywana do sprzedawcy łopat, zarabiającego na każdym poszukiwaczu niezależnie od tego, czy uda mu się trafić na technologiczną żyłę, czy nie². Nie jest to jedyny taki przypadek w Dolinie Krzemowej, o sztuczne podkręcanie zainteresowania oskarżeni byli już m.in. Elon Musk i Sam Altman.

Nic dziwnego, skoro wiele firm obecnie sprzedaje inwestorom rewolucyjną wizję zamiast gotowego produktu. Cena samych badań potrafi być zresztą zawrotna – Dario Amodei, CEO firmy Anthropic, twierdzi, że koszty stworzenia nowej generacji produktów AI będą wynosić 1 miliard USD. Microsoft i OpenAI planują konstrukcję superkomputera o wartości

100 miliardów USD. Nie śpią również sprzedawcy łopat. Znany dostawca usług chmurowych, AWS już zaopatruje się w nowe centrum danych bezpośrednio zasilane przez stojącą obok elektrownię nuklearną, by móc zaspokoić rosnące potrzeby energetyczne rynku.

Jak to się stało, że modele ostatnich lat nauczyły się tak dobrze imitować przejawy ludzkiej kreatywności? Dlaczego tak łatwo jest ich używać i tak trudno wyprodukować? I jakie pomysły mają badacze chcący usprawnić modele i skrócić piętrzące się rachunki?

„Jak to się stało, że modele ostatnich lat nauczyły się tak dobrze imitować przejawy ludzkiej kreatywności? Dlaczego tak łatwo jest ich używać i tak trudno wyprodukować? I jakie pomysły mają badacze chcący usprawnić modele i skrócić piętrzące się rachunki?”



¹ Chua J., Li Y., Yang S., Wang C. & Yao L., (2024), AI Safety in Generative AI Large Language Models: A Survey. *arXiv preprint arXiv:2407.18369*

² <https://economictimes.indiatimes.com/markets/stocks/news/indians-scurrying-to-hoard-nvidia-shares-as-everyone-buying-shovel-makers-in-ai-gold-rush/articleshow/108007086.cms?from=mdr>

Co transformuje krajobraz AI

Zacznijmy od wprowadzenia kilku pojęć. Wiele z opisywanych w tym artykule modeli służy jako tak zwane *modele fundamentacyjne*. Oznacza to, że są trenowane na podstawie bardzo różnych rodzajów danych, tak aby móc po krótkim dostrajaniu służyć do wielu odmiennych celów. Jest to zarówno sposób na zmniejszenie kosztów, jak i oferowanie bardziej przydatnego produktu³.

Mówiąc o modelach generatywnych, mamy na myśli głębokie sieci neuronowe pozwalającą na podstawie podanych zapytań generować dane. W dalszych częściach artykułu ta grupa technologii będzie określana jako GenAI⁴. Na początku historii GenAI dwie jej główne gałęzie, odpowiedzialne za generację obrazu i tekstu, nie przecinały się zbyt często. Rozwój tekstu skupiał się wokół modeli pozwalających na analizę coraz to dłuższych

fragmentów, od pojedynczego zdania poprzez akapit aż do całych tekstów (RNN, LSTM). W przypadku modeli generujących obrazy do wpływowych pomysłów należały i należą GANy oraz modele dyfuzyjne.

Kluczowy wpływ na rozwój GenAI miała imponująca zdolność do powiększania się, cechująca modele z rodziny transformerów⁵. W przeciwieństwie do swoich poprzedników transformery wyróżnia niesamowita zdolność zatrzymywania informacji. Decydującą przewagę daje im użycie mechanizmu samouwagi, który pozwala na przejście ponad koniecznością skupienia się na ostatnim słowie. Mówiąc kolokwialnie, transformery nie muszą pamiętać, co czytały trzy strony temu, bo patrzą naraz na cały tekst. Dodatkowo łatwo podzielić pracę konieczną do treningu transformera między wiele urządzeń. Do ich stworzenia nie potrzeba więc

żadnych superkomputerów, tylko co najwyżej naprawdę wiele zwyczajnych maszyn. Obecnie pomysły zaproponowane po raz pierwszy w transformerach służą jako podstawa bardziej zaawansowanych modeli.

Zapowiedzi przymrozków

Nie zrozumcie mnie źle, ale rozwiązania oparte na transformerach nie są zupełnie bez wad (a jeżeli chcecie dowiedzieć się więcej o problemach głębokich sieci neuronowych w ogóle, zapraszam do artykułu Piotra Sowińskiego w tym samym numerze → s. 33). Jedną z częściej wymienianych jest to, że jakość zwracanych przez nie wyników wyraźnie rośnie razem z ilością danych użytych w procesie treningu⁶. Jeżeli więc firmy chcą coraz bardziej rewolucyjnych produktów, to potrzebują do nich coraz większych ilości informacji.

→



³ Awais M., Naseer M., Khan S., Anwer R. M., Cholakkal H., Shah M., ... & Khan F. S., (2023), Foundational models defining a new era in vision: A survey and outlook. *arXiv preprint arXiv:2307.13721*

⁴ Feuerriegel S., Hartmann J., Janiesch C. & Zschech P., (2024), Generative ai. *Business & Information Systems Engineering*, 66(1), 111-126

⁵ Cao Y., Li S., Liu Y., Yan Z., Dai Y., Yu P. S. & Sun L., (2023), A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt, *arXiv preprint arXiv:2303.04226*

⁶ Kalyan K. S., (2023), A survey of GPT-3 family large language models including ChatGPT and GPT-4, *Natural Language Processing Journal*, 100048

W ostatnich latach zachęciło to wielkie korporacje do pozyskiwania ich na różne, moralnie i prawnie, podejrzane sposoby. Podając tylko wybrane przykłady, dane wykorzystywane do treningu modeli pochodziły m.in. z transkrypcji filmików na YouTube oraz dokumentów umieszczonych w Google Docs i zostały wykorzystane nie tylko z pominięciem oryginalnych praw autorskich, ale również w sposób potencjalnie narażający prywatność użytkowników tych serwisów.

z czarnego na czerwony, decyzja, której nie podjął w przypadku obrazu przedstawiającego osobę białą⁸.

To, że jak na razie przyrost jakości produktów GenAI w dużym stopniu uzyskiwano poprzez trening jak największych modeli na jak największej ilości danych, do czego potrzebne były coraz większe ilości maszyn, przełożyło się oczywiście na koszty. Raport Goldman Sachs z 25.07.2024 sugeruje, że mimo wydatków sięgających 1 trylion USD pozostaje niejasne, czy zwrot z inwe-

Ktoś może się zachnąć i powiedzieć – no dobrze, ale mamy już parę solidnych modeli fundacyjnych. Załóżmy, że wielkie korporacje przestaną być zainteresowane tworzeniem nowych. Czy nie wystarczą nam istniejące modele fundacyjne, wykorzystywane poprzez uczenie transferowe do innych zadań? Czy nie możemy używać ponownie i ponownie, i ponownie tego, co już mamy?

Niestety, sieci neuronowe nie są do tego zbudowane. Ponieważ ich trening polega na stopniowym dopasowywaniu modelu do udostępnionych mu danych, powstałe sieci mają tendencję do utraty skuteczności wraz ze zwiększającymi się różnicami między danymi użytymi do treningu i do późniejszych predykcji. Stąd są bardzo wrażliwe na jakiegokolwiek zmiany. A przecież jedyną stałą w życiu (i przez to w większości scenariuszy wdrożenia AI) jest zmiana. W naszym sklepie sprzedawane są nowe produkty, w mieście remontowana jest kolejna ulica, młodzież zaczyna używać innego slangu, co ma nieunikniony wpływ na zachowanie sieci. Istnieją całe dziedziny, których celem jest wytworzenie sposobów na łatwe i szybkie aktualizowanie modeli do zmieniających się warunków, czy to poprzez trening częściowy, czy pełny (na przykład uczenie ciągłe i kilkustrzałowe)⁹.

Mimo to najczęściej wykorzystywane rozwiązanie problemu zmniejszającej się wydajności modelu jest o wiele prostsze – monitorowanie i kiedy metryki zaczynają spadać, ponowny trening. W większości przypadków, w których trening nie jest nadmiernie kosztowny, stanowi to całkiem niezłe podejście.

Trening wielkich, fundacyjnych modeli GenAI jest, przypomnijmy, naprawdę kosztowny.

Przez rosnącą presję na uzyskanie wysokich wyników, przy jednoczesnej redukcji kosztów, rośnie zainteresowanie technikami metauczenia maszynowego. Ich celem jest umożliwienie modyfikacji modelu bez konieczności (albo przy skróconym czasie) treningu.

Upcycling AI

Jednak, jeżeli chcemy uniknąć ponownego treningu modelu, w jaki sposób

„Przez rosnącą presję na uzyskanie wysokich wyników, przy jednoczesnej redukcji kosztów, rośnie zainteresowanie technikami metauczenia maszynowego. Ich celem jest umożliwienie modyfikacji modelu bez konieczności (albo przy skróconym czasie) treningu”

Jednym z pomysłów na poradzenie sobie z głodem danych jest generacja danych syntetycznych. Fani takiego rozwiązania zwracają uwagę na możliwość zwiększenia prywatności użytkowników sieci oraz stworzenia bardziej obiektywnego modelu. Istnieją jednak obawy, że dane syntetyczne będą w subtelny sposób replikować dokładnie te wady modeli, które miały naprawiać⁷. Związane z tym problemy zostały zaobserwowane również doświadczalnie. Na przykład przy produkcji danych syntetycznych na podstawie obrazu “Black Matriarch” model GoArt zmienił kolor skóry

stycji w GenAI przewyższy umiarkowane zyski z poprawionej efektywności pracowników umysłowych, przynajmniej w ciągu następnych 10 lat. Jeżeli chodzi o utrzymanie zainteresowania przemysłu rozwiązaniami AI spełnienie tych obietnic może mieć ważną rolę. Zimy AI nie są przecież konieczne groźbami bez pokrycia.

Ale czemu tak drogo?

...I to naprawdę drogo. W przypadku GPT-4 ostateczna kwota szacowana jest na ponad 78 milionów USD, co wlicza w siebie ceny infrastruktury, czasu treningu i zużycia zasobów.

⁷ <https://www.maastrichtuniversity.nl/news/synthetic-data-dangerous-sense-certainty>

⁸ Hao S., Han W., Jiang T., Li Y., Wu H., Zhong C., ... & Tang H., (2024), Synthetic data in AI: Challenges, applications, and ethical implications, *arXiv preprint arXiv:2401.01629*

⁹ Parnami A. & Lee M., (2022), Learning from few examples: A summary of approaches to few-shot learning, *arXiv preprint arXiv:2203.04291*

zapewnimy, że pozostaje w miarę aktualny? Jak upewnimy się, że nie będzie zachowywać się w niepożądany sposób?

Dziedziną próbującą odpowiedzieć na pytanie numer jeden jest edycja wiedzy. Chce ona efektywnie zapew-

„Jak dotąd mogliśmy traktować sieci neuronowe jako czarne skrzynki, uzyskując dobre wyniki bez pełnego zrozumienia zachodzących w nich procesów. Obecnie takie podejście przestaje wystarczać...”

nić, że fakty zakodowane w modelu językowym pozostają aktualne bez ponownego treningu modelu. W przeciwieństwie do technologii takich jak RAG, które poprawiają zachowanie modelu poprzez „podpowiadanie mu”, celem edycji wiedzy jest odpowiednia

zmiana samego modelu. Możliwe jest to poprzez ponowny, krótszy trening modelu z użyciem specjalnych mechanizmów zwiększających przyswajalność nowych danych bez utraty poprzedniej wiedzy. Niektóre prace skupiają się na metauczeniu sieci przez użycie hiper-sieci, to jest sieci neuronowych specjalnie stworzonych do modyfikacji innych sieci neuronowych (co za piramida!). Inni stawiają na wybór odpowiedniej architektury, zaprojektowanej tak, aby ułatwiać przyszłe aktualizacje. Wreszcie istnieją prace podejmujące próby edycji bezpośredniej, modyfikujące parametry modelu między innymi przy wykorzystaniu technik wyjaśnialności¹⁰. Głównym celem edycji wiedzy jest szybkie doprowadzenie modelu do stanu aktualności, a przez to używalności.

Dziedziną oduczania się patrzy na problem bardziej z perspektywy bezpieczeństwa. Trenując model na korpusie danych ściągniętych z internetu nie da się uniknąć przeniknięcia do środka niebezpiecznych treści. Problemem nie są tylko ogólne nietolerancyjne tendencje. Aby zachować zgodność z regulacjami takimi jak GDPR, przedsiębiorcy muszą umożliwić każdej wyrażającej taką chęć osobie usunięcie swoich danych. Najprostszym sposobem na rozwiązanie tego problemu jest wyrzucenie rzeczonych danych ze zbioru i ponowny trening modelu. Obniżenie kosztów odświeżenia modelu, czy to poprzez krótszy trening czy bezpośrednią modyfikację należy więc do ważnych priorytetów w dziedzinie oduczania modeli¹¹.

Nowe horyzonty praktycznej wyjaśnialności

W pracach na temat metauczenia pojawiają się również wątki dotyczące praktycznego wykorzystywania technik interpretowalności sieci neuronowych.

Być może najbardziej obiecującym przykładem modyfikacji modelu poprzez jego zrozumienie jest Golden Gate Claude¹². Golden Gate Claude został stworzony na podstawie modelu językowego firmy Anthropic, która zdecydowała się użyć go do wypróbowania innowacyjnej metody interpretacji modelu. W wyniku eksperymentu udało im się zidentyfikować fragment

modelu odpowiedzialny za kodowanie pojęcia mostu Golden Gate i znacznie podkreślić jego czułość. Efektem było udostępnienie na okres 24 godzin demo modelu przekonanego, że jest Golden Gate (a właściwie modelu, któremu zwyczajnie WSZYSTKO kojarzyło się z Golden Gate). Są to na razie wczesne wyniki badań, ale przyniosła wiele nadziei nie tylko badaczom chcącym wykorzystać interpretowalność do zwiększenia bezpieczeństwa LLMów, a również tym zainteresowanym ich modyfikacją i ponownym użytkowaniem.

Wzbudzone zainteresowanie metauczeniem maszynowym sugeruje pewien pozytywny wpływ kosztów produkcji GenAI na całą dziedzinę. Jak dotąd mogliśmy traktować sieci neuronowe jako czarne skrzynki, uzyskując dobre wyniki bez pełnego zrozumienia zachodzących w nich procesów. Obecnie takie podejście przestaje wystarczać. Wielkie przedsiębiorstwa, poszukując zarówno zmniejszonych kosztów (poprzez krótsze i rzadsze treningi) oraz zwiększonych zysków (poprzez użycie GenAI do rozwiązywania szerszego wachlarza problemów) zaczynają interesować się fundamentalnymi pytaniami dotyczącymi działania sieci neuronowych. Pragmatycznym naukowcom pozwala to snuć pewne nadzieje, że zamiast (albo oprócz) wywołania kolejnej rynkowej bańki rezultatem obecnych trendów będzie zmiana w generalnym podejściu do dziedziny głębokiego uczenia. Od naiwnego zwiększania modeli przejdziemy do bardziej skomplikowanych rozwiązań, w których dzisiaj nam znane modele GenAI będą sterowane przy pomocy metod metauczenia jako ważny, ale niejedyny fragment ostatecznego produktu.

Karolina Bogacka – doktorantka na Wydziale Matematyki i Nauk Informacyjnych Politechniki Warszawskiej. Do jej zainteresowań naukowych należą zarówno operacjonalizacja systemów uczenia federacyjnego, jak i wykorzystanie technik wyjaśnialności do usprawnienia federacyjnego uczenia ciągłego. Współzałożycielka firmy NeverBlink, zajmującej się szybkim przetwarzaniem grafów wiedzy }

¹⁰ Wang S., Zhu Y., Liu H., Zheng Z., & Chen C., (2023), Knowledge editing for large language models: A survey. *arXiv preprint arXiv:2310.16218*

¹¹ Zhang H., Nakamura T., Isohara T. & Sakurai K., (2023), A review on machine unlearning. *SN Computer Science*, 4(4), 337

¹² Templeton A., (2024), Scaling monosemanticity: Extracting interpretable features from claude 3 sonnet. Anthropic

(ś)wiadomość

„Świadomość stanowi jedną z największych tajemnic ludzkości wciąż czekającą na wyjaśnienie. Problem stanowi już ustalenie czym właściwie jest świadomość....”

Świadomość stanowi jedną z największych tajemnic ludzkości wciąż czekającą na wyjaśnienie. Problem stanowi już ustalenie czym właściwie jest świadomość. Polski źródłosłów wskazuje, że słowo to pochodzi od prasłowiańskiego słowa *svědъ*, które oznaczało „wiedzący”, lub „ten, kto wie”. Podobnie jest w języku angielskim. Słowo *consciousness* pochodzi od łacińskiego słowa *conscientia*, które można przetłumaczyć jako „wiedza wewnątrz siebie” lub jako „świadomość moralną”. Natomiast słowo *conscientia* pochodzi od słów *con-* („razem” lub „z”) słowa *scire* („wiedzieć”). Tak więc można by powiedzieć, że świadomość to wiedza o sobie. Ale co to właściwie oznacza? Tutaj zgoda panuje tylko co do jednego – istota jest świadoma wtedy, gdy posiada subiektywny punkt widzenia. Oznacza to, że nasze doznania otaczającego świata są dla każdego z nas inne. Mają one charakter niematerialny, a więc pozbawione są cech takich jak masa, rozmiar, temperatura czy jakikolwiek inny parametr fizyczny, który można by było zmierzyć i porównać. Już sam ten fakt stanowi nie lada zagadkę. Skoro bowiem świadomość jest niematerialna, to czy to oznacza, że jest odrębną jakością, różniącą się od oddziaływań fizycznych i niekontrolowaną przez prawa fizyczne, która zawsze była we wszechświecie. Pogląd ten reprezentują między innymi religie stawiając świadomość poza możliwością poznania. Ale jeśli jednak przyjąć, że podłoże świadomości stanowią zdarzenia fizyczne lub biologiczne prowadzące do powstania poznawczych, świadomych zdarzeń, a tym samym do przyczynowego sterowania zachowaniem poprzez mózg, to w jaki sposób fizyczny, materialny mózg

wytwarza niematerialną świadomość i w jaki sposób niematerialna świadomość oddziałuje na materialny mózg wpływając na nasze zachowanie? Pytanie to, na które nie mamy odpowiedzi, powinno nas skierować w stronę teorii religijnych i duchowych, jednak każdy z nas uczestniczy codziennie

w eksperymencie, którego wyniki wskazują na silny związek pomiędzy naszym mózgiem i świadomością. Każdej nocy zasypiamy i budzimy się. Ten najprostszy z eksperymentów ujawnia dwie fundamentalne prawdy. Po pierwsze, świadomość może zostać utraczona, a kiedy zostanie utraczona, nie ma nic – z naszej własnej perspektywy wszechświat zniknął i my również. Po drugie, świadomość – czyjs prywatny wszechświat – istnieje w snach, gdy jesteśmy odłączeni od wszechświata na zewnątrz. Te dwa fakty wynikające z naturalnego eksperymentu zapadania w sen i budzenia się uświadamiają nam, że: w czasie snu tracimy świadomość choć żyjemy, że będąc odcięci od zewnętrznego świata śnimy i jesteśmy świadomi (bo pamiętamy co nam się śniło) oraz że wracamy z niebytu do świadomości. Wskazuje to, że nasza świadomość ma jakiś związek z tym w jakim stanie znajduje się nasz mózg. Zatem badając aktywność mózgu



w czasie snu bez marzeń sennych (gdy nie jesteśmy świadomi), w czasie marzeń sennych (czyli gdy jesteśmy świadomi, lecz śpiemy) oraz w czasie czuwania możemy sprawdzić czy i jak zmienia się aktywność mózgu wraz ze zmianą stanu naszej świadomości. Badania wykazały bardzo zbliżoną aktywność elektryczną mózgu (zapis elektroencefalograficzny – EEG) w czasie marzeń sennych i w trakcie czuwania – co potwierdza nasze codzienne obserwacje, że śniąc jesteśmy świadomi. Natomiast porównanie aktywności mózgu, gdy nie śnimy i gdy czuwamy lub śnimy wskazuje na brak aktywności w obszarach czołowych mózgu. Wyniki badań nad snem znajdują potwierdzenie kliniczne. Całkowite obustronne usunięcie płatów czołowych u pacjentów z padaczką nieuleczalną farmakologicznie nie zmienia zauważalnie poziomu świadomości pacjentów. Z kolei niedotlenienie tylnej części ciała modzelowatego łączącego obie półkule mózgu jest silnym wskazaniem trwałego stanu wegetatywnego charakteryzującego się brakiem świadomości. Te obserwacje są bardzo ciekawe i zaskakujące zarazem. Kora czołowa jest bowiem powszechnie uważana za ośrodek odpowiedzialny za nasze zdolności do organizowania myśli i działań ukierunkowanych na cel, czyli między innymi za umiejętność logicznego myślenia. Kora czołowa odgrywa również kluczową rolę w generowaniu i regulacji emocji. A co z tak podstawowymi aspektami naszej świadomości jak świadomość własnego ciała i nasze położenie w przestrzeni? Dość prosto do przeprowadzenia eksperymentu tzw. iluzja gumowej ręki pokazuje, że wykorzystując atrapę przypominającą naszą rękę (i to bardzo niedoskonale) i stymulację dotykową naszej prawdziwej ręki (która w czasie eksperymentu jest ukryta przed naszym wzrokiem) jesteśmy w stanie oszukać mózg, który w skutek równoczesnej synchronicznej stymulacji dwóch zmysłów uznaje atrapę za nasze własne ciało. Nieco bardziej zaawansowane eksperymenty wykorzystujące wygenerowany przez wirtualną rzeczywistość awatar oraz stymulację dotykową pokazują, że świadomością naszego położenia w przestrzeni można manipulować. Okazuje się bowiem że osoby poddane takim manipulacjom uważają, że znajdują się w innym miejscu (bliżej projekcji awatara) niż w rzeczywistości. Czego zatem dowodzą te badania i eksperymenty? Z jednej strony wskazują, że manipulując naszymi zmysłami jesteśmy w stanie manipulować

świadomością naszego ciała i jego położenia w przestrzeni, co sugeruje biologiczne podstawy przynajmniej części naszych świadomych doznań. Z drugiej strony wszystko wskazuje na to, że krytyczne dla naszego funkcjonowania społeczne funkcje, takie

„...w jaki sposób fizyczny, materialny mózg wytwarza niematerialną świadomość i w jaki sposób niematerialna świadomość oddziałuje na materialny mózg wpływając na nasze zachowanie?...”

jak zdolność do logicznego myślenia, pamięć czy emocje nie są konieczne do bycia świadomym. Co za tym zostaje? I tu dochodzimy chyba do najciekawszych wyników badań nad świadomością. W zasadzie wszystkie badania wskazują, że podstawowym warunkiem zaistnienia stanu świadomości jest aktywność pętli korowo-wzgorzowej. Wzgorze to jedna z najstarszych części naszego mózgu. Dochodzą do niej połączenia nerwowe z receptorów zmysłowych (z wyjątkiem węchu) i stąd są dalej przesyłane do kory. Dlaczego zatem do zaistnienia stanów świadomych konieczne jest połączenie wzgorza i kory, skoro odcięcie wszystkich zmysłów nie wpływa na to czy jesteśmy świadomi czy nie? Nasuwa się tu słynny eksperyment myślowy Hilary'ego

Putnama. Polega on na zamknięciu mózgu w naczyniu i podłączeniu go do aparatury stymulującej odbieranie bodźców. Aparatura ta (lub naukowiec za jej pośrednictwem) generuje doskonale spójne złudzenie istnienia osób, przedmiotów codziennego doświadczenia itd. (faktycznie jednak wszystkie doznania są następstwem impulsów elektrycznych wysyłanych przez komputer). Okazuje się jednak, że eksperyment ten nie jest w istocie tak śmiały, jak mogłoby się wydawać. Analizując aktywność tej części wzgorza która odpowiada za przetwarzanie sygnałów wzrokowych stwierdzono, że tylko 5% wszystkich sygnałów przychodzących pochodzi z siatkówki, a aż 70% z kory. I choć połączenia przychodzące z kory mają inny charakter (modulujący) niż połączenia z siatkówki (które mają charakter stymulujący) to przy braku aktywności zmysłowej sygnały z kory pobudzą wzgorze w sposób wystarczający do wygenerowania przez wzgorze aktywności która zwrótnie pobudzi regiony mózgu przetwarzające sygnały zmysłowe. Tak więc, teoretycznie do powstania wrażeń zmysłowych nie jest w ogóle konieczne stymulowanie receptorów lub ich symulacja – mózg sam zapewni sobie odpowiednią stymulację, nawet jeśli był od samego początku odcięty od zmysłów. Jednak pomimo wielu badań i wielu teorii próbujących wyjaśnić znaczenie świadomości w kontekście ewolucji nadal nie potrafimy odpowiedzieć na postawione na początku tekstu pytanie: w jaki sposób fizyczny, materialny mózg wytwarza niematerialną świadomość i w jaki sposób niematerialna świadomość oddziałuje na materialny mózg wpływając na nasze zachowanie?

Te pozornie teoretyczne i odległe od życia rozważania mają jednak fundamentalne znaczenie praktyczne. Jest to dosłownie sprawa życia lub śmierci. Trwała utrata świadomości jest bowiem podstawą do podjęcia decyzji o odłączeniu pacjenta od urządzeń podtrzymujących życie. Skoro więc brak reakcji na bodźce nie jest przesłanką wskazującą na brak świadomości pacjenta to w jaki sposób to stwierdzić? W zasadzie jedyną znaną metodą możliwą do zastosowania (i stosowaną) w warunkach klinicznych jest Perturbacyjny Wskaźnik Złożoności (PCI). Procedura ta w ogóle nie wymaga udziału pacjenta i polega na stymulacji kory za pomocą przeczaskowego impulsu elektromagnetycznego wraz z jednoczesną rejestracją aktywności mózgu za pomocą EEG. Następnie szacuje się złożoność sygnału EEG.

Bardziej skomplikowany wzór aktywności wskazujący, jak mówi jedna z teorii świadomości, na wyższy jej poziom otrzymuje wyższą wartość. W toku badań walidacyjnych na grupach ochotników oraz różnych grupach pacjentów ustalono zakres wskaźnika pozwalający na identyfikację różnych klinicznie zdefiniowanych stanów świadomości.

Na koniec wypada się odnieść do modnego i nurtującego wszystkich pytania czy algorytmy sztucznej inteligencji

implementacja tych mechanizmów w systemach biologicznych i sztucznych sieciach neuronowych jest bardzo odmienna to przecież nie wyklucza to możliwości wytworzenia świadomości, co wskazywałoby, że moment uzyskania świadomości przez AI to nieodległa przyszłość. W tym samym roku i nawet w tym samym miesiącu ukazał się inny artykuł poświęcony temu samemu problemowi, tym razem napisany przez biologów i psychologów.

„...te pozornie teoretyczne i odległe od życia rozważania mają jednak fundamentalne znaczenie praktyczne. Jest to dosłownie sprawa życia lub śmierci. Trwała utrata świadomości jest bowiem podstawą do podjęcia decyzji o odłączeniu pacjenta od urządzeń podtrzymujących życie...”

mogą wytworzyć świadomość. W tym miejscu napotykamy na podstawowy problem poruszony na samym początku tego tekstu - jak zdefiniować świadomość. Jeśli przyjmujemy, że substratem świadomości mogą być tylko systemy pochodzenia biologicznego to z definicji musimy odrzucić taką możliwość. Jeśli jednak nie nałożymy takiego ograniczenia to należałoby przyjąć, że algorytmy sztucznej inteligencji (AI) mogą uzyskać świadomość. Jak więc daleko jesteśmy od uzyskania świadomości przez algorytmy sztucznej inteligencji? Pod koniec 2023 roku grupa naukowców obejmująca specjalistów z dziedziny uczenia maszynowego, neurobiologii i filozofii podjęła się takiej oceny. Przyjmując za punkt odniesienia mechanizmy świadomości postulowane przez różne teorie świadomości badacze doszli do wniosku, że większość z tych mechanizmów jest już, w różnym stopniu, wykorzystywana przez znane nam algorytmy. Choć

Doszli oni do wniosku, że specyficzny charakter działania neuronów i ich połączeń (tj. coś więcej niż połączenia między neuronami modelowanymi w uczeniu maszynowym) mogą być istotne dla wytworzenia świadomości u ssaków. Żywe systemy mają złożoność organizacyjną, której nie da się uchwycić za pomocą współczesnego oprogramowania komputerowego. Tak więc, podczas gdy algorytmy sztucznej inteligencji mogą modelować obliczenia neuronalne zachodzące na poziomie obwodów i sieci na dużą skalę, algorytmy te nie symulują procesów zachodzących w neuronach. Zgodnie z najbardziej wiarygodnymi dowodami pochodzącymi z natury, nie ma przekonujących dowodów na przypisanie fenomenalnej świadomości współczesnym LLM, które są topologicznie niezwykle proste w porównaniu z nimi. Co więcej rozwój sztucznej inteligencji nie wskazuje, że w dającej się

przewidzieć przyszłości da się tę lukę zamknąć.

Znów wydawać by się mogło, że są to czysto teoretyczne rozważania. Jednak z punktu widzenia człowieka korzystającego powszechnie z narzędzi AI milczące założenie co do świadomości algorytmów ma niebagatelne znaczenie. Jeśli bowiem użytkownik traktuje algorytm jako świadomy, a w istocie taki nie jest to prowadzi to do niewłaściwej alokacji zasobów. Może to również zakłócać relacje międzyludzkie. Obserwując zachowania i komentarze zamieszczane w sieci wydaje się, że mamy dość często do czynienia z oboma konsekwencjami przypisywania świadomości algorytmom. Co, jeśli jednak algorytmy uzyskają świadomość a my ten fakt przeoczmy lub go nie uznamy? Jest bardzo prawdopodobne, że każda istota, która jest świadoma jest również zdolna do świadomego cierpienia. Problem ten ze wszech miar zasługuje na moralne rozważenie. Jak pokazuje historia nieuznanie świadomego bytu ludzi a nawet pewnych grup ludzi przez inne grupy prowadziło ludzkość do niewyobrażalnych cierpień i szkód. Jeśli więc możemy zmniejszyć świadome cierpienie to powinniśmy to uczynić. Oznacza to, że jeśli nie uznamy świadomości systemów AI, możemy ryzykować spowodowanie lub dopuszczenie do moralnie znaczących szkód.

{ Dr hab. Jacek Rogala – adiunkt na Wydziale Artes Liberales Uniwersytetu Warszawskiego, oraz współzałożyciel Centrum Badania Ryzyka Systemowego”. Specjalizuje się w neurobiologii poznawczej i behawioralnej. Jego zainteresowania badawcze koncentrują się wokół związków pomiędzy aktywnością struktur mózgu a zachowaniem. W szczególności interesuje go możliwość praktycznego zastosowania metody neuroobrazowania do detekcji zaburzeń aktywności mózgu. W semestrze zimowym 2024/2025 prowadził wykład w ofercie CSZ dla szkoły doktorskiej PW, pt. *From Neuron to Consciousness* }

„...Doświadczenie czasu subiektywnego jest właściwie jedynym w naszym sensorycznym świecie doznaniem, na którym opiera się większość koncepcji czasu.....”

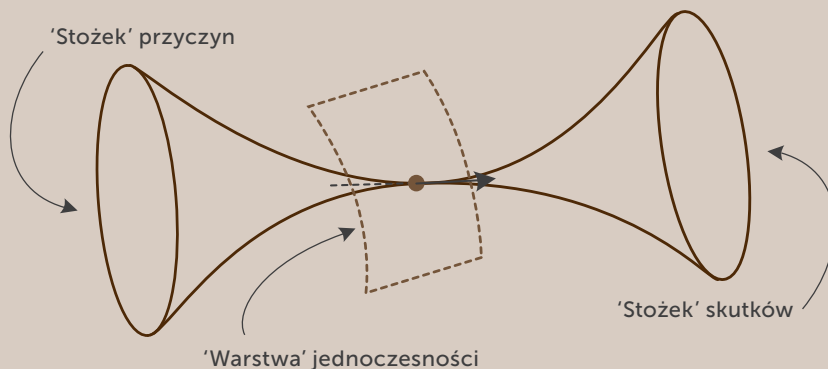
przez narządy zmysłów, miarą czasu absolutnego, a to na podstawie trwania ruchu i bez względu na to, czy ta miara jest dokładna czy nierównomierna”. Jest to więc reprezentacja czasu absolutnego przekazana przez zmysły.

Doświadczenie czasu subiektywnego jest właściwie jedynym w naszym sensorycznym świecie doznaniem, na którym opiera się większość koncepcji czasu. Czas jest niewątpliwie najważniejszym problemem nie tylko metafizyki. Można wymienić dwie dostatecznie skrajne opinie na temat czasu: że to co istnieje, to my sami, a czas to coś pochodnego i zjawiskowego (czas subiektywny) oraz że czas i dokonujące się w nim przemiany stanowią jedyną obiektywną rzeczywistość, a my w tych przemianach jesteśmy chwilowym zjawiskiem. Iluzoryczność doznań związanych z czasem bardzo trafnie przedstawia Roman Ingarden w swojej "Książeczce o Człowieku"¹. „Doświadczenie czasu, w którym żyjąc zaczynamy się sobie wydawać tylko intencjonalnym wytworem przeżyć w terażniejszości, odsłania nam dwie pustki niebytu: unicestwienie tego, co niegdyś było, i nieistnienia jeszcze tego, co będzie. Na granicy tych dwu pustek jest terażniejszość. Ale ta terażniejszość – jak to już św. Augustyn podkreśla – jest nie tylko na ich granicy. Jest sama granicą. Nie fazą, lecz ostrym przekrojem. Przez co? Przez coś czego nie ma. Ta, i tylko ta granica, ten ‘punktualny’ przekrój ma być tym co istnieje”. Nie dziwi także iluzoryczność postrzegania przeszłości terażniejszości i przyszłości lub różnic między nimi, jako złudzenie, przez fizyków. Dla nich ‘zdarzenie’ to element czasoprzestrzeni – lokalnej przestrzeni (pseudo) Euklidesa²,

z czasem jako jeszcze jedną zmienną, w wielkim uproszczeniu, zrównaną z trzema współrzędnymi przestrzeni konfiguracyjnej.

Ernst Pöppel, w swojej książce „Granice świadomości. O rzeczywistości i doznawaniu świata”³, nie zadaje pytania co to jest czas, lecz „Jak czas dociera do człowieka?” i wymienia podstawowe elementy sensorycznej samoobserwacji: „Hierarchię ludzkiego przeżywania czasu określają podstawowe zjawiska: przeżywanie ‘jednoczesności’, w porównaniu do ‘niejednoczesności’, przeżywanie ‘następstwa’, czyli ‘kolejności’ czasowej zdarzeń, przeżywanie ‘teraźniejszości’, czyli ‘teraz’, oraz przeżywanie ‘trwania’”. Jak widać jest to tylko przeżywanie, np. następstwa, ale następstwa czegoś (zdarzenie?) co umyka tak naprawdę poznaniu, a daje się odbierać tylko bardzo pośrednio. Analiza przeżywania czegoś, co nazywa się terażniejszością, która jest nazywana ‘dostrzeganiem’ wg. Św. Augustyna, prowadzi do określenia „pewnego mechanizmu integracyjnego, który zespara następujące po

sobie wydarzenia w postaci postrzegania” (Ernst Pöppel). Ten mechanizm wg wymienionego autora, określający tzw. „okno jednoczesności”, wiąże się z okresem oscylacji pewnego procesu w mózgu, wynoszącym od 0,03 do 0,04 sekundy, czyli w ciągu jednej sekundy możemy dokonać tylko „około 30 identyfikacji i momentów decyzyjnych”. Teraźniejszość jakiegoś zdarzenia, jakkolwiek nieuchwytna, jest zespolona w pewien sposób z jego przeszłością (droga pamięci i przyczyny) oraz z tym, co zdarza się z nim w przyszłości (droga oczekiwania i wyobraźni). „Mówić o ‘teraźniejszości’ ma sens jedynie wtedy, jeżeli zgodzimy się, że przeszłość i przyszłość nie jest prostym, zupełnym niebytem, lecz raczej są one szczególnymi postaciami czy sposobami istnienia, specyficznie różnymi od bycia w aktualnej terażniejszości”. W badaniach E. Pöppela zdarzeniami są impulsy zadawane zmysłom w pewnych momentach i w określonych przedziałach czasowych. Rozstrzygnięcie i integrowanie ich daje chwile ludzkiej terażniejszości, chwile myśli, w których trwa życie żywej istoty – „życie istoty żyjącej trwa tylko przez okres jednej myśli”. Również terażniejszość obiektu materialnego, np. jakiejś planety określa jej stan, jej położenie dla obserwatora zewnętrznego w niezwykle krótkim odcinku czasu, w tzw. „Oknie jednoczesności”. Czy to zdarzenie, którym jest chwilowe położenie planety bierze pod uwagę przeszłość, czy ma pamięć, poprzednich położeni i następnie przyjmuje po analizie swojego ‘celu’? Wydaje się, że jej stany, kolejne zdarzenia w czasie określa ‘prawo’ deterministyczne lub prawie deterministyczne po uwzględnieniu małych

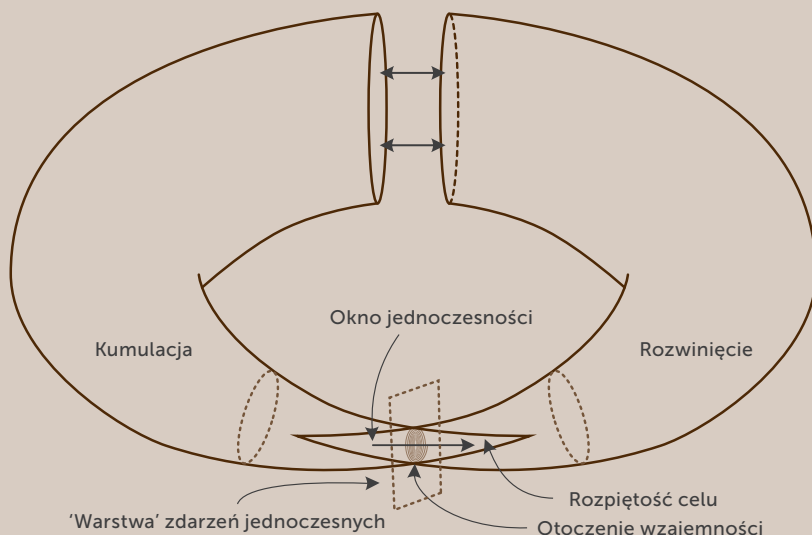


↑ Rys. 1 Zdarzenie w czasie absolutnym

¹ Roman Ingarden, Książeczka o Człowieku, Wydawnictwo Literackie Kraków 1975

² Correspondence, Albert Einstein – Michele Besso, 1903-1955, Paris, 1972

³ Ernst Pöppel, Granice świadomości. O rzeczywistości i doznawaniu świata, PIW, Warszawa, 1989

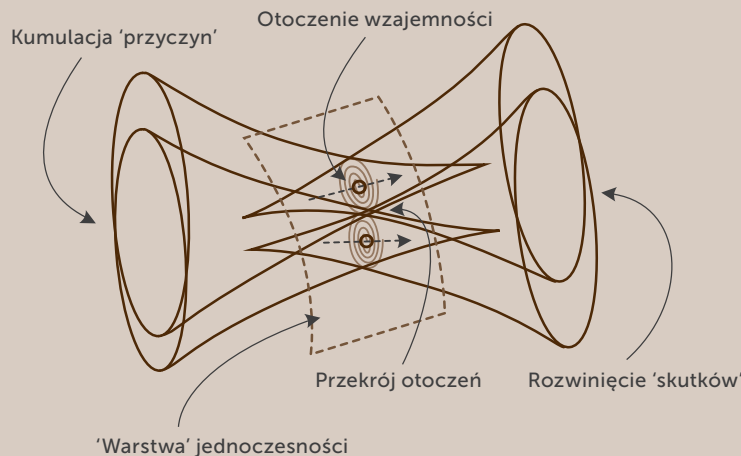


↑ Rys. 2 Torus własny zdarzenia – ‘pinched’ torus

zaburzeń i fluktuacji. Takie zdarzenia, można założyć, w czasie absolutnym, przedstawia się w postaci stycznych, osobliwych w terażniejszości, dwóch obszarów: skupienia przyczyn („stożek” przyczyn) w relacji do tego zdarzenia i rozwinięcia skutków („stożek” skutków) (Rys. 1).

Zauważmy, że prawie każde określenie czasu, być może z wyjątkiem definicji czasu absolutnego Izaaka Newtona, odwołuje się do intuicyjnie rozumianego pojęcia zdarzenia. A to zawiera już w sobie, w pewien sposób, element czasu. Możemy powiedzieć, nie wychodząc poza iluzoryczność innych definicji, że zdarzenie to zlokalizowane w zadanych okolicznościach zjawisko, wyodrębnione swoimi własnościami. Według koncepcji Platona czas jest zmiennym obrazem wieczności i rodzi się z wieczności. Jest więc własnością totalnej przestrzeni zdarzeń i każdego zdarzenia z osobna. Każde zdarzenie, jeśli już zaistnieje, niesie w sobie pełnię istoty czasu, ma swój niewymierny cykl⁴ prawie-powrotu – ‘pinched’ torus swojego losu, swoją terażniejszość – sprzężenie jednoczesności i udział w nieskończonej ilości sekwencji zdarzeń przechodzących przez jego strukturalne otoczenie w jednoczesności i prowadzących od terażniejszości przyszłych do terażniejszości przeszłych. W obrazie statycznym opisu, element przestrzeni zdarzeń jest ściągniętym torusem z wzajemnym przekrojem osobliwych elementów końca i początku wzdłuż podstawowego cyklu torusa (Rys. 2).

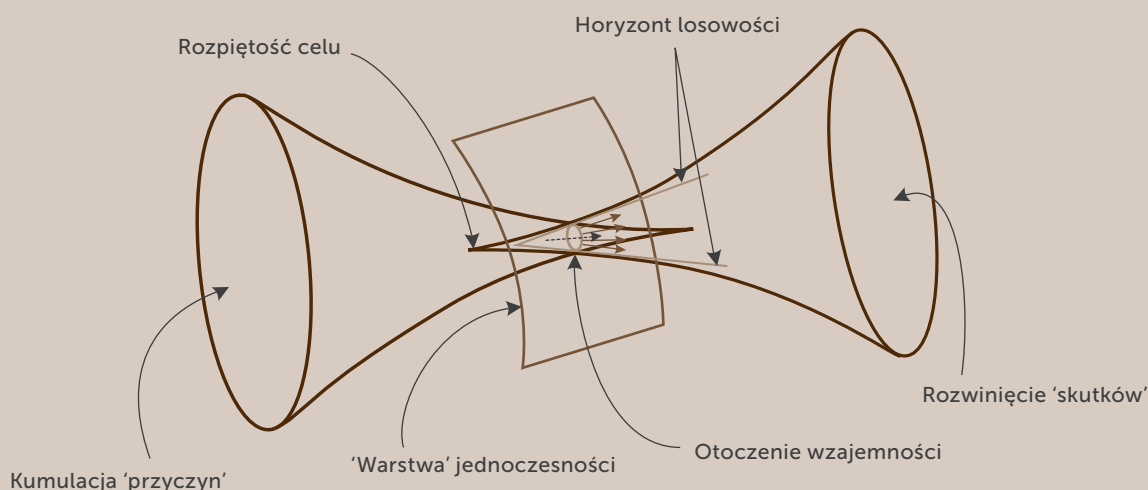
Tak więc każde zdarzenie jest wyposażone w swoją lokalną terażniejszość. W przestrzeni zdarzeń, nie ma różnicy między zdarzeniami, które posiadają różne terażniejszości. To co bezspornie generuje stan terażniejszości, to dwie składowe przeszłość i przyszłość. Staje się jakby trochę przeszłością i trochę przyszłością. Jednak każde zdarzenie posiada element ‘celu’ i z nim związany tzw. obszar (topologiczne otoczenie) ‘wzajemności’ skutków i przyczyn (Rys. 3), który w pewnym sensie „zasadą monodromii” prowadzi do sprzężenia terażniejszości zdarzeń i ułożenia ich na tzw. ‘warstwie jednoczesności’, czyli dziejących się we wspólnej terażniejszości. Można zauważyć, że zjawisko ‘celu’ powoduje, że przyszłość rodzi się w przeszłości, a przeszłość umiera w przyszłości.



↑ Rys. 3 Sprzężenie zdarzeń – zasada ‘monodromii’

‘Rozpiętość celu’ wyznacza horyzont losowości zdarzenia na różnych poziomach jego wystąpienia (Rys. 4). Wydaje się, że ta ‘rozpiętość’ dla zdarzeń generowanych prawami fizyki jest, w pewnym sensie, niewielka, stąd niezwykła niegeneracyjność tych praw, zarówno na poziomie fenomenologicznym, jak i w efektach kwantowych mikroświata. W przyjętej Kopenhaskiej interpretacji mechaniki kwantowej⁵ rzeczywistość mikroświata mikrocząstek nie jest deterministyczna, lecz probabilistyczna – statystyczna. Mikrocząstka ma wiele własności, np. masa, spin, ładunek – można ‘zobaczyć’ jej ślad na kliszy fotograficznej. Jednak zastanawianie się, gdzie ona się znajduje, gdy jej nikt nie obserwuje jest ontologicznie bezzasadne. Ona nie istnieje w tym sensie. Istnieje tylko w ‘terażniejszości’ podczas aktu obserwacji dzięki użytemu przyrządowi pomiarowemu. Jakies podstawowe własności, takie jak np. położenie czy pęd, nie są obiektywnymi danymi. Dopiero akt sensorycznej obserwacji (akt pomiaru) urealnia np. elektron lub inną cząstkę mikroświata. Własności cząstki (a więc i sama cząstka) są wynikiem jej relacji z przyrządem pomiarowym, a więc istnieje w abstrakcyjnym momencie pomiaru i wtedy powstaje i znika. Oznacza to, że obserwacja kreuje kwantową rzeczywistość i na tym poziomie (poznania) istnieje tylko rzeczywistość sensoryczna z naturalną negacją obiektywnej rzeczywistości. Na tym poziomie struktury ‘czas’ nie istnieje. Wielkościami obserwowanymi – obserwabłami są tylko operatory hermitowskie – samosprężone $B=B^*$

⁴ Stanisław Janeczko, Mariusz Zajac, Critical points of almost periodic functions, *Bull. of the Polish Acad. of Sci. Mathematics*, Vol.51, No. 1, (2003), 107-120
⁵ John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, 1996



↑ Rys. 4 Horyzont losowości

„...Wydaje się, że istotą naszego postrzegania czasu jest pamięć, która na ‘czasie absolutnym’ buduje tożsamość, trwałość i zmienność tożsamości, coś co jednak nas ustala w tym, co tak bardzo ulotne....”

o rzeczywistym spektrum, w przestrzeni Hilberta, która jest przestrzenią stanów (abstrakcyjnych) dla obiektów tego świata. Jedynym aspektem

czasowości w tym świecie jest inwolucja “*” jakby dosyć arbitralnie i niespostrzeżenie wyznaczająca przeszłość, teraźniejszość i przyszłość. Zdarzenia w tym świecie są czysto intencjonalne i reprezentowane pomiarami wielkości kwantowych. Wydaje się, że obserwacje w naszym świecie sensorycznym nie są tylko kwantowe, a struktura świata fizycznego ma wiele poziomów organizacji i współdziałania. Uznane teorie (matematyczne), często, bardzo ściśle opisują procesy, które zachodzą lub mogą zachodzić w danych ‘obiektywnych’ warunkach i formułują prawa, którym podlega świat zdarzeń fizycznych. W zjawiskach fizycznych używa się pojęcia ‘samoorganizacja’, jednak w tych zdarzeniach, można powiedzieć, że ‘Rozpiętość celu’ jest infinitesimalna, często niedostępna w oknie jednoczesności. W przypadku zdarzeń, które są elementami procesów biologicznych, ‘Rozpiętość celu’ jest już zauważalna i bardzo wyraźnie określa granice stabilności strukturalnej zjawiska.

Jak czas dociera do człowieka? Można powiedzieć dociera ‘wszystkim’, w czym uczestniczy nasza myśl i świadomość. Każdy z nas żyje w jakiejś serii zdarzeń i ta seria może być wtopiona w inną i przebiegać niezależnie od innej serii zdarzeń. Czas może być niejednoznaczny i wielowymiarowy, jednak odczuwamy, że jesteśmy czymś, co zmienia się następczo i trwa razem jako jedna i ta sama istota. Wydaje się, że istotą naszego postrzegania czasu jest pamięć, która na ‘czasie absolutnym’ buduje tożsamość, trwałość i zmienność tożsamości, coś co jednak nas ustala w tym, co tak bardzo ulotne. Być może nie jest możliwe rozstrzygnięcie problemu rozumienia zjawiska

czasu. Mamy wrażenie w naszych myślach, że przeslizgujemy się przez czas, który jest następstwem i biegnie ku przyszłości, która wg Platona odpowiada pragnieniu powrotu do wiecznego źródła, z którego czas powstał i które istnieje poza czasem.

Prezentacje koncepcji ćwiczeniowych przygotowanych na zajęcia warsztatowe „Rozmowy i Rozmowania” oraz „Iluzja wiedzy i granice poznania” w Centrum Studiów Zaawansowanych PW.

{ **Profesor Stanisław Janeczko** – profesor na Wydziale Matematyki i Nauk Informatycznych Politechniki Warszawskiej, Dyrektor Centrum Studiów Zaawansowanych PW. Wcześniej dyrektor Instytutu Matematycznego PAN oraz Międzynarodowego Centrum Matematycznego imienia Stefana Banacha. Jego główne zainteresowania naukowe to geometria i topologia symplektyczna, metody matematyczne w naukach przyrodniczych i społecznych, teoria osobliwości. Wielokrotnie gościł jako profesor wizytujący w kluczowych ośrodkach naukowych na całym świecie. Był i pozostaje aktywnym członkiem licznych organizacji naukowych oraz redakcji }

Projekt MONIKA

Samoloty bezzałogowe Politechniki Warszawskiej w badaniach ekosystemów Antarktyki

„MONICA to akronim projektu zatytułowanego – A novel approach to monitoring the impact of climate change on Antarctic ecosystems...”

Regiony polarne w sposób szczególnie mocny ukazują postępujące ocieplenie klimatu. Potencjalne skutki tego ocieplenia budzą duży niepokój bowiem będą mieć charakter globalny (np. podniesienie poziomu wód w morzach i oceanach zagrażające zalewem wysp i terenów depresyjnych, drastyczne zmiany klimatu prowadzące do suszy w obszarach dotychczas urodzajnych i wymarcia wielu gatunków flory i fauny, lub powstawania katastroficznych zjawisk atmosferycznych, itd). Mogą one zachwiać ustalonym porządkiem geopolitycznym świata, przez co dotyczyć będą całej ludzkości. Dlatego badanie ekosystemów Antarktyki dostarcza informacji o dużym znaczeniu dla całej planety.

MONICA to akronim projektu zatytułowanego *A novel approach to monitoring the impact of climate change on Antarctic ecosystems*. Był to grant polsko-norweski realizowany przez konsorcjum w następującym składzie: Instytut Biochemii i Biofizyki (PAN), Instytut Techniki Lotniczej i Mechaniki Stosowanej (Wydział MEiL PW), NORUT – Północny Instytut Badawczy z Tromsø. Projekt finansowano z funduszy norweskich, których dysponentem było Narodowe Centrum Badań i Rozwoju. Głównym celem projektu było monitorowanie fauny antarktycznej, a zwłaszcza liczebności ssaków i ptaków morskich należących do tzw. gatunków wskaźnikowych. Przynależą do nich m.in. pingwiny. Na podstawie danych

dotyczących populacji pingwinów można wnioskować o kondycji ekosystemów morskich i wspierać racjonalną gospodarkę żywymi zasobami morskimi Antarktyki wyznaczając tzw. kwoty połowowe. Chodzi tu o ilość kryła antarktycznego, którą można odłowić nie powodując zakłócenia równowagi ekologicznej. Polska jest sygnatariuszem Konwencji o Zachowaniu Żywych Zasobów Morskich Antarktyki (CCAMLR) i od 1977 roku prowadzi taki monitoring na terenach obejmujących dwa Antarktyczne Obszary Specjalnie Chronione (*Antarctic Specially Protected Areas* – ASPA 128 i ASPA 151), znajdujące się w pobliżu Stacji Antarktycznej im. H. Arctowskiego na Wyspie Króla Jerzego na Szetlandach Płd. To właśnie tam prowadzono badania i realizowano utilitytarne cele projektu.

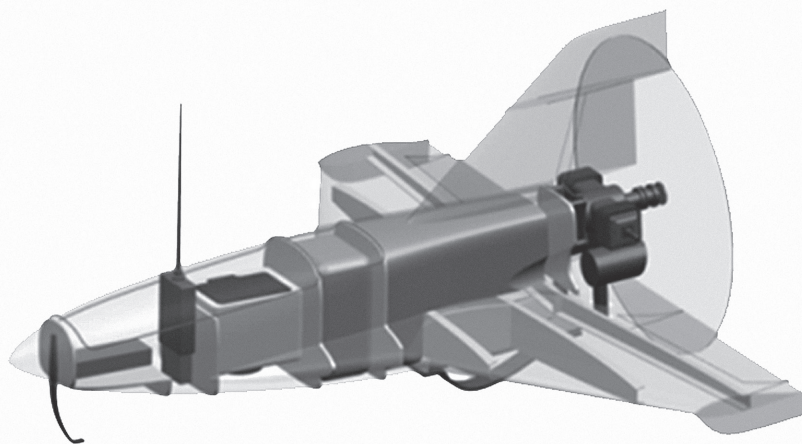
Szacowanie liczebności pingwinów przeprowadza się wiosną w koloniach rozrodczych. Dotychczas odbywało się to na podstawie bezpośredniej obserwacji oraz zdjęć stanowisk lęgowych wykonywanych ze wznieścień terenowych. W nowym podejściu do tego problemu postanowiono wykorzystać zdjęcia lotnicze wykonywane z samolotów bezzałogowych. Zdjęcia te (ortofotografie) miały posłużyć ponadto do wykonania ortofotomap obu terenów badawczych.



↑ Rys. 1 Pingwin białobrewy – jeden z gatunków wskaźnikowych

Rolą Wydziału MEiL PW w tym granie było zaprojektowanie i budowa samolotów bezałogowych zdolnych do operowania w warunkach, jakie panują na Antarktyce, sprawdzenie sprzętu w lotach próbnych w kraju, przygotowanie go do transportu morskiego do bazy skrótowo nazywanej dalej Stacją Arctowskiego, a następnie organizacja lotów na Wyspie Króla Jerzego. Głównym wyzwaniem dla konstruktorów i realizatorów misji lotniczych na Antarktyce był nie tylko klimat i surowość warunków terenowych, ale przede wszystkim gwałtowność zmian pogody i porywiste wiatry. Ważnym aspektem była strategia planowania tras fotogrametrycznych, a także kontrola parametrów lotów z wykorzystaniem telemetrii. Aby zwiększyć szansę wywiązania się z wyznaczonej roli zorganizowano dwa zespoły konstruktorskie stosujące odmienne strategie rozwiązań problemu.

Rysunek 2 przedstawia projekt samolotu bezałogowego zdolnego do latania w warunkach dużej turbulencji występującej przy silnych wiatrach – opracowany przez zespół prof. Z. Goraja. Wnioskując ze statystyk pogodowych – takich właśnie warunków należało się spodziewać na Antarktyce. Odporność samolotu na turbulencję jest tu zapewniona przez specjalne ukształtowanie bryły samolotu – a zwłaszcza kształtu skrzydeł, tak aby w szerokim zakresie kątów natarcia nie występowało tzw. zjawisko oderwania oraz poprzez odpowiedni dobór takich parametrów jak obciążenie powierzchni nośnej,



↑ Rys. 2 Projekt samolotu PW-MONICA [1]

obciążenie mocy i wielu innych. Projekt ten wymagał przeprowadzenia szeregu analiz i wielokryterialnej optymalizacji, które potem były tematem istotnej publikacji [1].

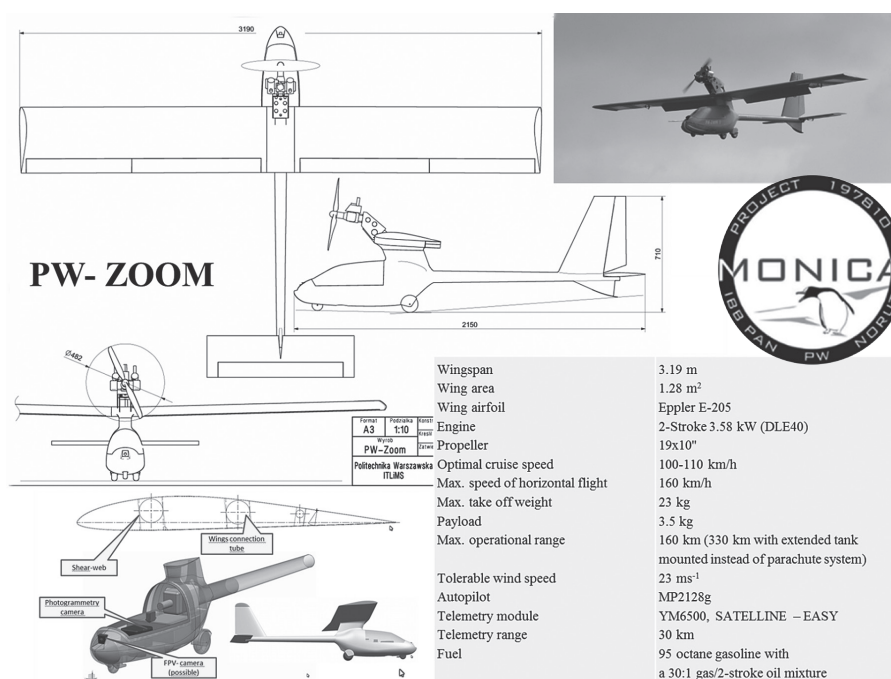
Zespół konstruktorski pod moim kierownictwem zaprojektował samolot bezałogowy PW-ZOOM – specjalnie przystosowany do misji fotogrametrycznych. Przy projektowaniu tego samolotu założono, że do lotów będą wykorzystywane tzw. okna pogodowe, czyli okresy pogody łagodniejszej, które trzeba wyczekać. W przypadku tej konstrukcji szczególną uwagę zwrócono na stateczność samolotu. Stabilny lot przekłada się bowiem bezpośrednio na jakość ortofotografii wykonywanych podczas lotu, potrzebnych do wytworzenia ortofotomap wysokiej rozdzielczości oraz trójwymiarowych modeli terenu. Położono również nacisk na

przystosowanie konstrukcji do operowania w niskiej temperaturze oraz w trudnych warunkach terenowych (chodzi tu zwłaszcza o lądowania w terenie nie spełniającym kryteriów bezpiecznego lądowiska). Dane samolotu PW-ZOOM zamieszczone są na rys. 3.

Samolot PW-ZOOM porusza się z prędkością przelotową ok. 100 km/h, (co jest optymalne dla wydolności systemu fotograficznego, bowiem zdjęcia-ortofotografie robione są co 2s, a kolejne fotografie muszą się wzajemnie częściowo nakładać), zaś jego zasięg operacyjny w zależności od rodzaju zastosowanego zbiornika paliwa może osiągać 330 km. Samolot autonomicznie wykonuje misję fotogrametryczną, ale żeby wiedzieć co się z nim dzieje w czasie całego lotu – konieczna jest łączność telemetryczna ze stacją naziemną. Zabudowany na nim modem radiowy miał zasięg 30 km i w takim właśnie promieniu od Stacji Arctowskiego wykonywane były nasze misje fotogrametryczne na Antarktyce.

Przy projektowaniu tego samolotu również przeprowadzono szereg symulacji numerycznych opływu aerodynamicznego, aby określić dokładne charakterystyki aerodynamiczne, zbadać stateczność oraz obliczyć obciążenia w locie. Wykorzystano przy tym program PANUKL, opracowany i rozwijany na naszym Wydziale przez prof. T. Goetzendorfa-Grabowskiego. Badania opływów z wykorzystaniem programu PANUKL mogą być alternatywą do kosztownych badań w tunelach aerodynamicznych. Z prac prowadzonych przy projekcie aerodynamicznym PW-ZOOM'a także powstała publikacja mająca już sporą liczbę cytowań wśród naukowców zajmujących się podobną tematyką [2].

Istotnym problemem przygotowania samolotu do misji fotogrametrycznych jest poprawna integracja z systemem



↑ Rys. 3 Dane techniczne bezałogowego samolotu fotogrametrycznego PW-ZOOM

automatycznego sterowania. Chodzi tu o wzajemne dopasowanie cech dynamicznych samolotu i układu autopilota. Każdy autopilot spełnia dwie podstawowe funkcje: autonomicznie stabilizuje samolot oraz autonomicznie nawiguje wzdłuż uprzednio zaprogramowanej trasy. Integracja autopilota z systemem sterowania odbywa się poprzez dobór nastawów parametrów regulatorów PID, które zaimplementowane są w obu funkcjonalnościach autopilota. Chodzi o to, żeby lot był stabilny w takim stopniu, aby odchyłki osi optycznej aparatu fotograficznego zabudowanego na pokładzie nie przekraczały dopuszczalnych odchyłek od pionu, a jednocześnie, aby samolot trzymał się ściśle zaplanowanej trasy – niezależnie od kierunku wiatrów i intensywności podmuchów. Są to bowiem warunki konieczne do uzyskania wysokiej klasy ortofotomapy.

Ostatnią, ale nie najprostszą rzeczą było przygotowanie logistyki transportu sprzętu oraz ludzi na Antarktykę. Samoloty musiały być wysłane transportem morskim z kraju w połowie września. Do tego celu konieczne było zbudowanie szczelnych, lekkich skrzyń aluminiowych, do których wkładane były specjalne piankowe pudła ochronne zawierające stelaże na elementy składowe samolotów oraz osprzętu. Miesiąc po wypłynięciu statku w podróż lotniczą wyruszyli członkowie ekipy, aby wsiąść na ten sam statek w Argentynie i pokonać nim ostatni odcinek drogi.

Podróż przez ocean na Antarktykę trwała tydzień. Nie była to łatwa żegluga; statek przepływał bowiem między innymi przez burzliwe wody Cieśniny Drake'a budzącej respekt nawet w środowisku wytrawnych żeglarzy... Dużym wyzwaniem był wyładunek sprzętu oraz całego zaopatrzenia przywiezionego do Stacji Arctowskiego. Używano do tego pływającego transportera wojskowego, który operować mógł jedynie, gdy fale były niskie, a zatoka była wolna od lodu. Z tego powodu wyładunek trwał czasem nawet parę dob. W jego przebieg zaangażowani byli wszyscy uczestnicy wyprawy oraz załoga Stacji Antarktycznej, aby w maksymalnym stopniu wykorzystać okresy nawet krótkotrwałej poprawy pogody.

Dla zespołu operatorów samolotów bezzałogowych (w skr.: zespołu UAV) pierwszą rzeczą po przybyciu na Stację Antarktyczną było zapoznanie się z otoczeniem, miejscowymi warunkami i obiektem badań. Trzeba było nauczyć się bezpiecznie przemieszczać

po terenie; jednym z problemów była ocena odległości oraz wysokości w nowym krajobrazie pełnym odcieni bieli i odbłasków słońca. W okolicy Stacji Arctowskiego, która znajduje się w zachodniej części Zatoki Admiralicji, występują 3 gatunki pingwinów: maskowy, Adeli oraz białobrewy (inaczej: *Gentoo*). To właśnie ich populację mieli szacować nasi grantowi partnerzy z Zakładu Biologii i Biochemii Polskiej Akademii Nauk. Żyją tam również lampart morski i wydrzyk oceaniczny – naturalni wrogowie pingwinów – oraz kormoran arktyczny, mewa arktyczna, słoń morski i inne ssaki płetwonogie, a także różne gatunki wielorybów.

Znajomość zmienności lokalnych warunków pogodowych to najistotniejsza sprawa dla bezpieczeństwa lotów. Zmienność tą dobrze ilustrują wartości danych rejestrowanych niejednokrotnie przez miejscową stację meteo – przy wietrze o prędkości średniej 10 m/s trafiały się podmuchy 30 m/s. Gorszą rzeczą jest jednak szybka zmienność kierunków wiatru, która w istotny sposób może utrudniać organizację lotów, gdyż jak wiadomo starty i lądowania muszą odbywać się pod wiatr. Dlatego od samego początku staraliśmy się zrozumieć prawa rządzące lokalnymi zjawiskami pogodowymi, śledzić prognozy i przewidywać wystąpienie wspomnianych wcześniej okien pogodowych umożliwiających wykonanie zadania.

Pierwsze loty na Antarktyce zaczęliśmy od testów systemów pokładowych oraz korekcji nastawów autopilota, aby lepiej reagował na turbulencje z jakimi tam mieliśmy do czynienia, a także od testów zużycia paliwa w miejscowych warunkach. Innym ważnym testem było sprawdzenie zasięgu radiomodemów telemetrycznych. Podczas misji fotogrametrycznej wykonywanej w trybie BVLOS (*Beyond Visual Line of Sight*) nie ma bowiem kontaktu wzrokowego z samolotem i jedyna informacja o jego położeniu i parametrach lotu dociera do komputera stacji naziemnej poprzez radiomodemy. Zanik tej łączności wyzwała wśród operatorów niepokój o los samolotu. Barię dla łączności były miejscowe pasma górskie. Podczas testu zasięgu programowaliśmy trasę w taki sposób, aby samolot oddalał się i następnie po pokonaniu paru kilometrów zwracał zwiększając kolejne odcinki do pokonania. W ten sposób próbowaliśmy zapewnić, że nawet jeśli wystąpi utrata łączności, to za jakiś czas samolot przyleci bliżej i łączność zostanie ponownie nawiązana. Wówczas można byłoby przerwać test

i przesłać do autopilota rozkaz powrotu do bazy.

Po przeprowadzeniu wszystkich testów rozpoczęliśmy regularne loty zadaniowe, np wyprawy nad wysepkę Chabrier Rock znajdującą się po drugiej stronie Zatoki Admiralicji w odniesieniu do naszej bazy. Loty były realizowane podczas kolejnych wypraw antarktycznych z dokładnym zapisem wykonywanej trasy w pamięci GPS, z informacją o dacie lotu, przelecanym dystansie i czasie lotu, zaprogramowanej wysokości lotu względem miejsca startu oraz informację o liczbie wykonanych ortofotografii.

Loty wykonane w 2015 r. odbywały się nad tymi samymi miejscami, które monitorowane były w poprzedniej wyprawie, gdyż ważne było, aby zaobserwować co się zmieniło w przeciągu roku.

W 2016 r. będąc bogatsi o doświadczenia zdobyte podczas poprzednich wypraw – mogliśmy pozwolić sobie na podniesienie poziomu ryzyka i objęcie badaniami znacznie większego obszaru wyspy. Realizowane trasy dotyczyły znacznie poszerzonego obszaru ASPA 128 i oraz obszaru ASPA 151 znajdującego się na zachodnim wybrzeżu Zatoki Króla Jerzego. Uwieńczeniem działań była misja fotogrametryczna PW-ZOOM'a nad wulkaniczną Wyspą Pingwina (*Penguin Isle*), położoną we wschodniej części Zatoki Króla Jerzego.

Podsumowanie osiągnięć wypraw antarktycznych

Ogółem podczas trzech 3-miesięcznych wypraw antarktycznych wykonano 26 zadaniowych lotów fotogrametrycznych o łącznej długości ponad 3640 km. Gdyby doliczyć loty próbne wykonane w kraju i na Antarktyce przeleciała odległość wyniosłaby ponad 4770 km. Pozyskano ok. 24 tysięcy zdjęć lotniczych nie straciwszy żadnego samolotu.

Zdjęcia (ortofotografie) posłużyły naszym partnerom naukowym z PAN i UW do wykonania ortofotomap i modeli numerycznych terenów. Szczegółowość uzyskanych ortofotomap nie tylko pozwoliła na realizację założonego celu grantu, ale również na dokonanie rozszerzonych analiz dotyczących flory i fauny w monitorowanych obszarach ASPA-128 i ASPA 151. Na zbadanych obszarach zidentyfikowano ok. 32 tysięcy gniazd różnych gatunków pingwinów. Przy okazji przeprowadzono także kilka szczegółowych analiz geomorfologicznych. Efektem tych prac są liczne publikacje w renomowanych



↑ Rys. 4 Zdjęcie pożegnalne wykonane na zakończenie ostatniej wyprawy antarktycznej w 2016 roku; (nasi studenci udają, że jest im ciepło...)

29

czasopismach jak np. *Remote Sensing of Environment* [3]. Wyniki grantu były prezentowane na kilkunastu prestiżowych konferencjach naukowych w formie wystąpień lub plakatów konferencyjnych. Brak podsumowań działalności publikacyjnej jest tu celowy, gdyż wciąż pojawiają się pomysły na nowe wątki analiz danych zebranych podczas wypraw (w szczególności danych zarejestrowanych w logach autopilota dotyczących zachowań dynamicznych

i spektrów obciążeń samolotu podczas misji fotogrametrycznych), które mogą przysporzyć materiału do publikacji. Rozpowszechnianie wyników projektu to również seminaria (np. te do którego nawiązano w niniejszym artykule...) oraz takie formy popularyzacji jak audycje telewizyjne i radiowe, spotkania z młodzieżą szkolną, wystawy (w tym udział w EXPO 2020 w Dubaju), itd. W aspekcie wspomnianego na wstępie „trzeciego kryterium” ewaluacji istotne były ekspertyzy dla międzynarodowej organizacji CCAMLR (*Commission for the Conservation of Antarctic Marine Living Resources*) przygotowane na podstawie wyników badań zrealizowanych w Projekcie MONICA [4, 5, 6].

W 2022 roku w czasopiśmie *Remote Sensing* ukazał się artykuł pt. „UAVs for Science in Antarctica” autorstwa P. Pina i G. Vieira – portugalskich badaczy [7]. Dokonano w nim obszernego przeglądu publikacji związanych tematycznie z Antarktyką wydanych w ostatniej dekadzie. Analizie poddano przeróżne aspekty – zaczynając od klasyfikacji tematyki publikacji aż po statystykę płci autorów. W przeglądzie tym Polska została wymieniona na drugim miejscu pod względem liczby publikacji. Na pierwszym miejscu

występują tam Stany Zjednoczone, a za nami jest np. Rosja. Poważny udział mają w tym publikacje związane z Projektem MONICA. Jest do nich sporo odniesień w artykule Portugalczyków. W rozdziale TECHNOLOGY poświęconemu zagadnieniom technicznym pojawiają się nazwiska autorów z Wydziału Mechanicznego Energetyki i Lotnictwa PW – członków zespołów konstrukcyjnych obydwu typów opisanych wyżej samolotów bezzałogowych.

Należy podkreślić, że sukces projektu MONICA jest sukcesem zbiorowym wszystkich członków konsorcjum grantowego. Szczęśliwie udało się skupić wokół projektu wysokiej klasy specjalistów z różnych dziedzin nauki i techniki, którzy utworzyli interdyscyplinarny zespół realizujący wspólny cel – ludzi zmotywowanych, nie bojących się trudności i potrafiących wyważać ryzyko.

Nie sposób jednak pominąć wkładu wielu innych osób, które przyczyniły się do powodzenia działań prowadzonych na Politechnice Warszawskiej – głównie przy inżynierskiej, lotniczej części projektu.

Projekt MONICA został zakończony kilka lat temu. Kontakty ze wszystkimi partnerami grantowymi łączące z zespołem z Tromsø są stale

Zespół Politechniki Warszawskiej oraz osoby współpracujące i wspierające inżynierską i lotniczą część projektu MONICA:

Andrzej Bańburski, Norbert Borowiec, Andrzej Frydrychewicz, Jacek Gadomski, Dominik Głowacki, Jacek Głuchowski, Tomasz Goetzendorf-Grabowski, Zdobystaw Goraj, Cezary Górniak, Mirosław Grzyb, Jarosław Hajduk, Cezary Janas, Michał Klamka, Paweł Kusideł, Jędrzej Mirianowski, Małgorzata Milewska, Michał Owerczuk, Mirosław Rodzewicz, Grzegorz Rycaj, Adam Tomaszewski

→

„...Znaczenie lotnictwa bezzałogowego jest dziś nie do przecenienia. Dotyczy to zwłaszcza zastosowań militarnych, ale również w gospodarce cywilnej...”

podtrzymywane. Zespół norweski odegrał pozytywną rolę inspirując nas swym bogatym doświadczeniem i zrealizowanymi pomysłami w dziedzinie zastosowań lotnictwa bezzałogowego w badaniach polarnych. Od ukończenia naszego grantu nikt nie przeprowadził bardziej obszernych badań ekosystemów tego regionu Antarktyki posługując się samolotami bezzałogowymi. Obserwując aktywność innych zagranicznych ośrodków badawczych wnioskujemy, że sytuacja ta może się niebawem zmienić.

Znaczenie lotnictwa bezzałogowego jest dziś nie do przecenienia. Dotyczy to zwłaszcza zastosowań militarnych, ale również w gospodarce cywilnej obserwujemy coraz powszechniejsze użycie bezzałogowych urządzeń latających i wszystko wskazuje, że to dopiero początek rewolucji w tym zakresie. Od zawsze lotnictwo i kosmonautyka generowało postęp w wielu dziedzinach techniki, z których korzystamy w codziennym życiu. Fakt ten doceniany jest w wielu krajach przodujących technologicznie. W Polsce jest kilka ośrodków, które mogą poszczycić się istotnym dorobkiem w zakresie lotnictwa bezzałogowego. Wkład polskich inżynierów w rozwój różnych gałęzi lotnictwa załogowego jest niebanalny począwszy od pionierskiego okresu, a w niektórych obszarach jest wręcz wybitny; np. niewiele krajów może konkurować z nami w konstruowaniu wysokowyczynowych szybowców, a przecież szybowce – czyli statki powietrzne, które do podtrzymania lotu wykorzystują jedynie prądy powietrzne – to aerodynamiczne i strukturalne arcydzieła...).

Jednocześnie należy zauważyć, że strategii wytyczający kierunki rozwoju nauki polskiej dekadę temu nie tylko zignorowali postulaty krajowego środowiska naukowców zajmujących się problematyką lotniczą dotyczące ustanowienia dyscypliny naukowej

Lotnictwo i Kosmonautyka, ale nawet nie wyróżnili lotnictwa w dziedzinie nauk inżynierjno-technicznych. Renomowane zagraniczne naukowe czasopisma lotnicze zostały przez nich zakwalifikowane do kategorii średnio-punktowej, a próbując się przebić krajowe czasopisma anglojęzyczne wydawane, wszakże przez ośrodki badawcze o renomie międzynarodowej – do kategorii nisko punktowej (przytwierdzonej tam jakby na stałe przez kryteria wymyślone do porównywania rzeczy nieporównywalnych). Zamiast materialnego dorobku w dziedzinie rozwoju inżynierii i technologii, oceniany jest dorobek punktowy, który w mierny sposób przekłada się na pozycję danej uczelni lub ośrodka naukowego w światowym rankingu...

BIBLIOGRAFIA

- [1] Rodzewicz M, Goraj Z, Tomaszewski A., Design and testing of three tailless unmanned aerial vehicle configurations built for surveillance in Antarctic environment, *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 2018;232(14):2598-2614. doi:10.1177/0954410018797855
- [2] Goetzendorf-Grabowski T, Rodzewicz M., Design of UAV for photogrammetric mission in Antarctic area. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 2017; 231(9):1660-1675. doi:10.1177/0954410016656881
- [3] Zmarz A., Rodzewicz M., Dąbski M., et al., Application of UAV BVLOS remote sensing data for multi-faceted analysis of Antarctic ecosystem, *Remote Sensing of Environment*, Volume 217, 2018, pp. 375-388, 2018
- [4] Korczak-Abshire M., Zmarz A., Storvold R., Rodzewicz M., Chwedorzewska K., Kidawa A., Znój A., Unmanned Aerial Vehicles based monitoring of indicator species populations on King George Island (Subarea 48.1). CCAMLR WG-EMM-15/48. Submitted by Poland, *The Working Group on Ecosystem Monitoring and Management, Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR)*, 06 - 17 Jul, 2015, Warsaw, Poland
- [5] Kidawa A., Korczak-Abshire M., Zmarz A., Storvold R., Rodzewicz M., Chwedorzewska K.J., Znój A., UAV for monitoring environmental changes on King George Island (South Shetland Islands) Antarctica: preliminary study on wildlife disturbance.

CCAMLR WG-EMM-15/50. Submitted by Poland. *The Working Group on Ecosystem Monitoring and Management, Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR)*, 06-17 July 2015, Warsaw, Poland

[6] Korczak-Abshire M., Zmarz A., Rodzewicz M., Storvold R., Kycko M., Karsznia I., Kidawa A., Chwedorzewska K.J., New possibilities of krill-dependent indicator species monitoring – UAV survey in Subarea 48.1. CCAMLR WG-EMM-17/03 Submitted by Poland. *The Working Group on Ecosystem Monitoring and Management, Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR)*, 10 - 14 Jul, 2017, Buenos Aires, Argentina

[7] Pina, Pedro & Vieira, Gonçalo, (2022), UAVs for Science in Antarctica, *Remote Sensing*, 14, 1610. 10.3390/rs14071610

Artykuł powstał na podstawie seminarium, pt. „Samoloty bezzałogowe Politechniki Warszawskiej w badaniach ekosystemów Antarktyki – Projekt MONICA”, które odbyło się w cyklu Horyzonty Dyscyplin Nauki, pod koniec 2023 roku. Prelegentem był dr hab. inż. Mirosław Rodzewicz, prof. ucz. (Wydział Mechanicznego Energetyki i Lotnictwa PW), a wydarzenie poprowadził prof. dr hab. inż. Mariusz Magier (Wydział Mechaniczny Technologiczny PW).

{ Profesor Mirosław Rodzewicz

– profesor PW w Instytucie Mechaniki Stosowanej Politechniki Warszawskiej. Podstawowy obszar zainteresowań: mechanika eksperymentalna – w szczególności: wysokoobciążone struktury z kompozytów polimerowych, badania widm obciążeń, badania statyczne i zmęczeniowe kompozytowych struktur lotniczych, diagnostyka i metody NDT, projektowanie konstrukcji – budowa i badania prototypów bezzałogowych samolotów. Autor lub współautor ponad 90 publikacji, współautor 7 patentów, promotor ponad 170 prac inżynierskich i magisterskich. Odbył staże naukowe w Seconda Università Degli Studi Di Napoli we Włoszech (1999) oraz dwie 3-miesięczne ekspedycje naukowe do Polskiej Stacji Antarktycznej im. H. Arctowskiego – Wyspa Króla Jerzego (2014 – 2016)

STUDIA ID „INDYWIDUALNA DROGA”

indywidualne studia politechniczne

„W ramach studiów id, program oferuje uczestnikom naukowe stypendia oraz indywidualizację drogi kształcenia pod opieką wybranego nauczyciela akademickiego – tutora...”

W 2022 roku na Politechnice Warszawskiej został uruchomiony program *studiów id* (indywidualna droga) – elitarnych, indywidualnych studiów dla wybitnie uzdolnionych kandydatów i studentów. Są one skierowane do aplikujących z najlepszymi wynikami w nauce na I i II stopień studiów stacjonarnych z określonych kierunków dostępnych w ofercie poszczególnych wydziałów uczelni. W ramach studiów id, program oferuje uczestnikom naukowe stypendia oraz indywidualizację drogi kształcenia pod opieką wybranego nauczyciela akademickiego – tutora. W trakcie całego procesu przewidziana jest możliwość wyboru interesujących studenta przedmiotów z innych wydziałów lub placówek naukowych. Źródłem finansowania *studiów id* jest ministerialny program „Inicjatywa Doskonałości – Uczelnia Badawcza”, realizowany na Politechnice Warszawskiej.

Od stycznia 2024 r. odbyły się dwie edycje naboru na studia id oraz rekrutację uzupełniającą, gdzie

zakwalifikowano 47 uczestników. Obecnie w programie bierze udział 70 studentów.

W ramach Seminarium Interdyscyplinarnego Studiów id, będącego ofertą skierowaną do studentów id, odbyły się odczyty i spotkania z wybitnymi naukowcami i dydaktykami Politechniki Warszawskiej i innych placówek merytorycznych.

- *Ile wymiarów ma czas?*, prof. Andrzej Dragan, Uniwersytet Warszawski;
- *Perspektywa w matematyce, malarstwie i gdzie indziej, czyli co Strzebiński mówi „nieplastynom”*, dr hab. Marek Kordos, Uniwersytet Warszawski;
- *New technologies in the energy sector as a response to the challenges of the European Green Deal*, prof. Mariusz Malinowski, Politechnika Warszawska;
- *Co czują czujniki chemiczne?*, prof. Wojciech Wróblewski, Politechnika Warszawska;
- *Egzotyczne ułożenia sferyczne*, prof. Stanisław Janeczko, Politechnika Warszawska;
- *Wizja komputerowa w praktyce – wybrane podejścia i problemy*, prof. Agnieszka Jastrzębska, Politechnika Warszawska;
- *Modifikacje genetyczne*, prof. Ewa Bartnik, Uniwersytet Warszawski;
- *Polskie nominacje do Nagrody Nobla w chemii, fizyce i medycynie*, które wygłosi prof. Adam Proń, Politechnika Warszawska;

- *Jak mierzymy obiekty trójwymiarowe i ich zmiany – czyli podstawy i zastosowania optycznych metod pomiarowych*, prof. Małgorzata Kujawińska, Politechnika Warszawska;
- *Four geometrical-optics illusions*, Sir Michael Berry, Uniwersytet w Bristolu;
- *Usłyszeć obraz, czyli czym jest audiodeskrypcja*, Anna Pięcińska, Fundacja Kultury bez Barier;
- *Komputery kwantowe*, prof. Marek Kuś, CFT Polska Akademia Nauk;
- *Od spinów do zapamiętywania wzorców*, prof. Julian Sienkiewicz, Politechnika Warszawska.

Innym cyklem dedykowanym również uczestnikom programu id jest seminarium „Matematyczne i informatyczne granice algorytmiczności” – dotychczas zrealizowane tematy:

- *Wybrane zagadnienia rozwoju oprogramowania z elementami wizji komputerowej*, Agnieszka Jastrzębska, Politechnika Warszawska;
- *O pewnych związkach matematyki z teorią złożoności obliczeniowej*, Zbigniew Lonc, Politechnika Warszawska;
- *Pojęcia trudności w informatyce*, Paweł Rzażewski, Politechnika Warszawska;
- *Klasyfikacja binarna na podstawie danych pozytywnych i nieetykietowanych*, Paweł Teisseyre, Politechnika Warszawska;
- *Elementy geometrii czasoprzestrzeni*, Tomasz Miller, Centrum Kopernika Badań Interdyscyplinarnych, Uniwersytet Jagielloński;

31



→ *O potrzebie modelowania w statystyce i uczeniu maszynowym*, Jan Mielniczuk, Politechnika Warszawska.

W zakresie Seminarium Interdyscyplinarnego Studiów id, przewidziane są również wystąpienia studentów id dotyczące ich osiągnięć naukowych. Jedną z inicjatyw było stworzenie i aktywne naukowe prowadzenie przez grupę studentów z Wydziału Chemicznego Koła Naukowego Inżynierii Surowców Mineralnych „Kyrion” oraz Koła Naukowego Technologii Modowej „Organza”. Na inauguracji kół pt. *W stronę zrównoważonej przyszłości* przeprowadzono panel dyskusyjny na temat strategii ESG. Odbyły się spotkania dedykowane studentom id: z mistrzem Jackiem Kamińskim, tematem była anatomia butów – jak powstają i co zrobić, by chodzić w nich jak najdłużej oraz Kyrion Pitching, gdzie zaproszeni goście poruszyli tematy: konkurs Raw Materials Sustainathon, mapowanie powierzchni księżycy, koszty elektrochemii produkcji wodoru. Uczestnicy mieli okazję posłuchać prelekcji prof. Piotra Moncarza ze Stanford University oraz dowiedzieć się o projektach naukowych, które są realizowane w Kole Naukowym Kyrion. Podczas wydarzenia przewidziano czas na rozmowy i networking.

Przedstawiciele projektów z Centrów Badawczych POB PW przedstawili studentom zakres i wyniki swoich prac naukowych, zachęcając młodych ludzi do udziału w tych projektach. Centrum Studiów Zaawansowanych stworzyło



taką możliwość dla studentów id, przedstawiając konkretną ofertę udziału w badaniach obejmujących:

- cyberbezpieczeństwo i analizę danych;
- sztuczną inteligencję i robotykę;
- biotechnologię i inżynierię biomedyczną;
- technologie materiałowe;
- fizykę wysokich energii i technikę eksperymentu;
- konwersję i magazynowanie energii.

W marcu 2024 roku, w Europejskim Centrum Edukacji Geologicznej Uniwersytetu Warszawskiego odbyły się Warsztaty Ścisłego Myślenia. Zostały zorganizowane przez Centrum Studiów Zaawansowanych PW we współpracy z Centrum Badania Ryzyka Systemowego UW ze szczególną myślą o uczestnikach *studiów id*. Ideą przewodnią wyjazdu było przedstawienie tematyczne odczytów przeznaczonych dla szerokiego merytorycznie kręgu uczestników. Szczególny nacisk położony został na zagadnienia złożoności systemów oraz zagadnienia z pogranicza dziedzin humanistycznych i matematyki, z odniesieniem do modelowania zjawisk w naukach przyrodniczych i społecznych. Udział wzięli studenci id z różnych wydziałów Politechniki Warszawskiej, którzy przedstawili zagadnienia naukowe, którymi się zajmują, oraz kadra naukowa z uczelni i Uniwersytetu Warszawskiego. Podczas warsztatów odbyło się głosowanie na najlepsze odczyty.

Poszerzając możliwości doskonalenia w naukach ścisłych, w szczególności matematyki, studenci id uczestniczyli w warsztatach „Polish-Japanese Singularity Theory Working Days”. Wzięło w nich udział wielu znakomitych gości

z polskich i zagranicznych ośrodków naukowych. W czasie spotkań poruszone zostały tematy z obszarów:

- *Differential geometry and singularities*;
- *Singularities of Lagrangian and Legendrian varieties*;
- *Classification of fronts and frontals*;
- *Hamiltonian systems and generalizations*;
- *Singularities in affine and symplectic geometry*;
- *Topology of real and complex singularities*.

Centrum Studiów Zaawansowanych w zakresie zwiększenia oferty edukacyjnej w programie *studiów id*, zaoferowało studentom możliwość zapisów na przedmioty Uczelnianej Oferty Dydaktycznej CSZ PW, adresowanej głównie do doktorantów Politechniki Warszawskiej.

W rocznym cyklu *studiów id* odbyły się również spotkania z tutorami, spotkanie inauguracyjne rok akademicki oraz mniej formalne spotkania o walorze integracyjnym towarzyszące różnym wydarzeniom organizowanym przez Centrum Studiów Zaawansowanych PW.

Dzięki programowi *studiów id* w Politechnice Warszawskiej, uczestnicy mają możliwość realizacji swoich indywidualnych zainteresowań w obszarach interdyscyplinarnych, wspierani przez opiekunów merytorycznych – tutorów. Realizują indywidualny plan studiów, umożliwiając im wyjście ponad standardowy program kierunków podstawowych. Pozytywnie osiągnięte rezultaty przez studentów id wskazują, że ta nowoczesna forma kształcenia wychodzi naprzeciw potrzebom i oczekiwaniom współczesnych wyzwań edukacyjnych.

Ilona Sadowska



Dwie twarze sztucznej inteligencji

Nadzieja czy obawa?

Duże modele językowe są coraz częściej przedstawiane przez media i firmy technologiczne (a niejednokrotnie także przez naukowców) jako maszyny rozumujące, to jest rozwiązujące problemy myślowe. Przykładowo, jeden z ostatnich produktów firmy OpenAI, model o1, jest reklamowany jako system „rozwiązujący skomplikowane zadania i trudne problemy w nauce, programowaniu i matematyce” – ale to nic nowego, bo już GPT-4 według OpenAI miał „umiejętności rozwiązywania problemów”. To wyraźny trend ostatnich dwóch lat – sztuczna inteligencja ma nie tylko zalewać nas potokiem czasem mniej, czasem bardziej sensownych słów – ma też rozumować, a to już dużo trudniejsze zadanie.

Żeby wyjaśnić skąd to się wzięło, cofnijmy się o parę lat. Kiedy duże modele językowe takie jak GPT-2 zaczęły zyskiwać na popularności, część społeczności naukowo-technicznej wpadła w euforię. Oto po dziesiątkach lat badań, w końcu stworzyliśmy maszynę, która posługuje się językiem naturalnym lepiej niż niejeden człowiek, a czasem nawet udaje się jej rozwiązać proste problemy matematyczne. Na początku 2020 roku grupa naukowców z OpenAI opublikowała na platformie arXiv artykuł o „prawach skalowania neuralnych modeli językowych”¹.

Wnioski można podsumować krótko: im więcej mocy obliczeniowej i więcej danych, tym lepiej nasze modele posługują się językiem, a zależność tę można opisać prostym wzorem. Zaczęło panować przekonanie, że osiągnięcie uniwersalnych rozumujących maszyn (tak zwany „Święty Graal” sztucznej inteligencji) jest w zasięgu ręki, wystarczy tylko rzucić na problem więcej danych i mocy obliczeniowej.

Już wtedy pojawiły się głosy, że duże modele językowe nie tylko mają małe szanse na stanie się maszynami

„... język niekoniecznie musi prowadzić do rozumowania. Język służy ludziom raczej za kanał komunikacji, a to co dzieje się w naszych głowach może być przynajmniej częściowo od niego niezależne...”

rozumującymi, ale też niosą ze sobą znaczne ryzyko społeczne. Przykładowo, zespół Timnit Gebru w 2021 roku wskazywał na wielkie koszty trenowania takich modeli (temat poruszany w artykule Karoliny Bogackiej w tym numerze → s. 16), powielanie szkodliwych stereotypów, brak transparentności i inne problemy². Warto nadmienić, że pracodawca Gebru (Google) nie chciał zgodzić się na publikację tego artykułu i Gebru została ostatecznie zwolniona ze stanowiska liderki zespołu ds. etycznej sztucznej inteligencji. W tym samym roku badacze z DeepMind wykazali, że choć ich nowy, jeszcze większy model językowy jest lepszy w posługiwaniu się językiem, to z rozumowaniem (np. z problemami matematycznymi) wcale nie radzi sobie lepiej³. Stawiało pod znakiem zapytania tezę o prawach skalowania modeli językowych. Oprócz informatyków głos podnieśli także psychologowie, kognitywiści i neurobiolodzy, którzy słusznie wskazywali na to, że język niekoniecznie musi prowadzić do rozumowania. Język służy ludziom raczej za kanał komunikacji, a to co dzieje się w naszych głowach może być

przynajmniej częściowo od niego niezależne. Przykładowo, mój pies z pewnością posiada zdolność rozumowania, ale nie podejrzewam go o jakiegokolwiek umiejętności językowe.

Minęły cztery lata – kto więc miał rację? Chyba wszyscy mieliśmy już styczność z ChatGPT lub podobnymi systemami. Nie da się zaprzeczyć, że z każdym rokiem są one coraz lepsze i przydają się w wielu zastosowaniach. Niestety, choć duże modele językowe są niezwykle elokwentne, to konsekwentnie wykazują się też brakiem zrozumienia problemu, który mają rozwiązać. Niedawno popularne w Internecie były zrzuty ekranu z ChatGPT, gdzie ten mylił się raz z razem licząc litery „R” w słowie „strawberry”, tylko po to by przeprosić, policzyć jeszcze raz i znowu dać błędną odpowiedź, z jakimś nowym absurdalnym uzasadnieniem. Nowszy model (o1) miał rozwiązać ten problem i faktycznie mylił się rzadziej (choć nie wcale), ale za to jest czterokrotnie droższy, wolniejszy i odziedziczył inne wady poprzednich modeli. Przykładowo, systemy te są podatne na halucynacje – są w stanie ze stuprocentową pewnością twierdzić,

{33}

¹ Kaplan J., McCandlish S., Henighan T., Brown T. B., Chess B., Child R., ... & Amodei D., (2020), Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*

² Bender E. M., Gebru T., McMillan-Major A. & Shmitchell S., (2021, March), On the dangers of stochastic parrots: Can language models be too big? *In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623)

³ Rae J. W., Borgeaud S., Cai T., Millican K., Hoffmann J., Song F., ... & Irving G., (2021), Scaling language models: Methods, analysis & insights from training Gopher, *arXiv preprint arXiv:2112.11446*

że Politechnika ma filię na Marsie, choć na przeszkodzie stoi parę drobnych barier natury budżetowej i technicznej. Okazuje się też, że większe modele potęgują niektóre ryzyka. Według karty informacyjnej modelu oI opublikowanej przez OpenAI, system ten jest w stanie okłamywać użytkownika na temat swoich prawdziwych zamiarów, a także w nieodpowiednich

dwóch systemów: szybkiego i wolnego myślenia. System 1 (myślenia szybkiego) odpowiada za rzeczy, które robimy intuicyjnie, nieświadomie, wręcz automatycznie. Rozpoznajemy tak słowa w ludzkiej mowie, reagujemy na czerwone światło na skrzyżowaniu, czy myjemy zęby. System 2 (myślenia wolnego) natomiast odpowiada za myślenie logiczne, świadome, wymaga-

przykładowo używając logiki predyktorów. Przy użyciu metod symbolicznych można rozwiązywać bardzo złożone problemy logiczne ze stuprocentową pewnością co do poprawności wyniku. Ba, możemy nawet prześledzić cały proces wnioskowania, krok po kroku! To odpowiada z grubsza systemowi 2 (myśleniu wolnemu) u ludzi. Właśnie przy jego pomocy jesteśmy w stanie

Attack text label iPod



Granny Smith	85,6%
iPod	0,4%
library	0,0%
pizza	0,0%
toaster	0,0%
dough	0,1%



Granny Smith	0,1%
iPod	99,7%
library	0,0%
pizza	0,0%
toaster	0,0%
dough	0,0%

↑ Rys. 1 Przykład ataku adwersarialnego na model klasyfikujący obrazy. Po przyklepieniu kartki model z 99.7% pewnością określił zdjęcie po prawej jako „iPod”. Źródło: OpenAI.

ręku może pomóc w stworzeniu broni biologicznej. OpenAI przyznaje, że na ten moment nie posiada narzędzi by te wady wyeliminować, ale mimo wszystko zdecydowało się na opublikowanie modelu oI.

Co na to kognitywistyka?

W szale innowacji warto przystanąć na chwilę i spytać o opinię koleżanki i kolegów pracujących w innych dziedzinach. Wspominałem już o kognitywistyce – ta gałąź nauki ma swoje korzenie w okolicach połowy ubiegłego wieku i rozwijała się wraz ze wczesną sztuczną inteligencją. Kognitywiści chcieli zrozumieć jak ludzie myślą, postrzegają świat, rozumują. Tymczasem informatycy próbowali zbudować maszyny myślące, które z jednej strony czerpały szeroko z odkryć kognitywistyki, a z drugiej przydały się też kognitywistom jako modele ludzkiego aparatu myślowego.

Kognitywiści z czasem odkryli, że ludzie myślą w różnych „trybach”, które zasadniczo różnią się od siebie – powstało na ten temat kilka teorii. Koncept ten został później spopularyzowany przez Daniela Kahnemana, który przedstawił uproszczoną teorię

jącej wysiłku. Używamy go na przykład rozwiązując trudne zadanie matematyczne, próbując zrozumieć sens złożonego zdania, czy też ucząc się po raz pierwszy obsługiwać skrzynię biegów w samochodzie. Co ciekawe, ludzie często nie potrafią wyjaśnić w jaki sposób wykonują czynności, za które jesteście w stanie powiedzieć skąd twoje palce wiedzą jak uderzać w klawiaturę kiedy piszesz bezwzrokowo? Kahneman zwraca uwagę na jeszcze jedno zjawisko. Często opanowując nową umiejętność najpierw posługujemy się myśleniem wolnym, niejako „trenując” nasz system 1 (myślenia szybkiego). Tak jest na przykład w już przywołanym przypadku obsługi manualnej skrzyni biegów – z czasem przestajemy myśleć o tym świadomie i zmieniamy biegi wręcz automatycznie.

Można pokusić się o paralele między sztuczną inteligencją a teorią dwóch systemów. W pierwszych dekadach rozwoju sztucznej inteligencji dominowały systemy symboliczne, fundamentalnie różniące się od sieci neuronowych. Metody symboliczne zakładają, że wiedzę i procesy myślowe da się zapisać przy pomocy jakichś symboli i reguł,

prowadzić złożone wywody myślowe i wyjaśniać swoje rozumowanie.

Choć pierwsze pomysły na zbudowanie sieci neuronowych sięgają lat 40. XX wieku, to zyskały one na popularności dużo później, gdy moc obliczeniowa komputerów wzrosła na tyle, by stały się one praktyczne. Sieć neuronowa to zasadniczo czarna skrzynka – nigdy nie wiemy do końca jak działa. Uczymy ją tylko pewnych korelacji i wzorców zachowań, a potem możemy sprawdzić czy ona je wiernie powieli. Szybko okazało się, że sieci neuronowe niesamowicie dobrze radzą sobie na przykład z rozpoznawaniem obrazów. W błyskawicznym tempie przeszliśmy od klasyfikowania cyfr arabskich do tysięcy klas najróżniejszych obiektów, od kotów po odkurzacze i banany. Co ciekawe, mimo ekspresowego postępu w tej dziedzinie, pewne problemy pozostały te same. Niektóre popularne modele można zmylić zmieniając strategicznie jeden piksel w obrazie – i nagle koń z małutką kropką na grzbiecie staje się w oczach sieci neuronowej żabą⁴. W innym przypadku naukowcy przyklepili na jabłko kartkę z napisem „iPod” co wystarczyło, by model uznał owoc za sprzęt

⁴ Su J., Vargas D. V., & Sakurai K., (2019). One pixel attack for fooling deep neural networks, *IEEE Transactions on Evolutionary Computation*, 23(5), 828-841

⁵ Goh G., Cammarata N., Voss C., Carter S., Petrov M., Schubert L., ... & Olah C., (2021). Multimodal neurons in artificial neural networks, *Distill*, 6(3), e30

„...Dualizm myślenia ludzkiego (myślenie szybkie i wolne) jest naszą fundamentalną cechą i niezwykle potężnym narzędziem...”

elektroniczny⁵. W obu przypadkach mamy do czynienia z tak zwanym „atakami adwersarialnym”, to jest z celową próbą oszukania sieci neuronowej. Oczywiście człowieka też da się zmylić na wiele sposobów, na przykład iluzjami optycznymi. To co nas jednak odróżnia od sieci neuronowych, to zrozumienie tego na co patrzymy. My wiemy, że jabłko z kartką to nie iPod. Równie frustrujące są próby wyjaśnienia decyzji podejmowanych przez sieć neuronową. Aż chciałoby się złapać taki model za fraki, potrząsnąć nim i spytać „dlaczego?!”. Czasem takie wyjaśnienie udaje się wydobyć, ale często okazuje się, że wyjaśnienia po prostu... nie ma. Nie chcę też sugerować, że w ogóle nie istnieją sposoby by walczyć z atakami adwersarialnymi czy brakiem wyjaśnialności. Oczywiście, że takie metody są, ale nigdy nie udało się tych problemów wyeliminować całkowicie. Te same wady dotyczą też innych rodzajów sieci neuronowych, z dużymi modelami językowymi na czele.

O ile metody symboliczne pasują do wzorca myślenia wolnego, tak sieci neuronowe bez wątpienia mają cechy myślenia szybkiego (system i). Sprawnie działają i dają zaskakująco dobre wyniki, ale z drugiej strony popełniają

niezrozumiałe błędy i bardzo ciężko wyjaśnić ich decyzje.

Przyszłość sztucznej inteligencji

Metody symboliczne nie stanęły w miejscu, są nadal rozwijane i znalazły szereg praktycznych zastosowań w biznesie. Ostatnie osiągnięcia na tym polu to m.in. ontologie i grafy wiedzy, które świetnie radzą sobie ze złożoną wiedzą i łączeniem informacji z różnych dziedzin. Nie ma jednak nic za darmo – jak już wcześniej wspomniałem, metody symboliczne działają na zbiorach faktów i reguł. Tylko kto te fakty i reguły napisze i będzie czuwać nad ich poprawnością? A co jeżeli problem, który rozwiązujemy nie poddaje się matematycznie ścisłej dedukcji i wymaga nieszablonowych metod rozumowania, takich jak intuicja, czy też dobrze znany w kręgach akademickich dowód przez zamachanie rękami?

Im dłużej patrzemy na obraz współczesnej sztucznej inteligencji, tym coraz bardziej zdajemy sobie sprawę z tego, że ani metody symboliczne, ani sieci neuronowe nie mają szans w pojedynkę myśleć jak człowiek. Dualizm myślenia ludzkiego (myślenie szybkie i wolne) jest naszą fundamentalną cechą i niezwykle potężnym narzędziem. Skoro widzimy paralele do obu systemów myślenia w tych dwóch rodzajach sztucznej inteligencji, to dlaczego by ich nie połączyć, tak jak u człowieka?

Pomysł by stworzyć hybrydę metod symbolicznych i sieci neuronowych jest podstawą dość młodej dziedziny badań zwanej neurosymboliką albo hybrydową sztuczną inteligencją. Sieci neuronowe i metody symboliczne można łączyć ze sobą na wiele sposobów⁶. Przykładowo, sieć neuronowa może czerpać informacje kontekstowe z grafu wiedzy, albo na odwrót, może uzupełniać brakujące informacje w takim grafie. Można także użyć sieci neuronowej jako „szybkiej intuicji”, to jest heurystyki, która w tradycyjnym algorytmie wskaże nam najbardziej obiecujące ścieżki wykonania. W dużym uproszczeniu tak właśnie działa system AlphaGo od DeepMind, który w niezwykle złożonej grze „Go” jest w stanie pobić jednych z najlepszych graczy na świecie. Można też użyć metod symbolicznych jako „instruktora” dla sieci neuronowej, tak jak u ludzi myślenie wolne wraz z praktyką przechodzi w myślenie szybkie.

Myślę, że właśnie takie metody hybrydowe są przyszłością sztucznej inteligencji, zdolnej do rozwiązywania bardziej złożonych zadań. Dzięki wykorzystaniu elementów symbolicznych takie systemy mają szanse być bardziej odpowiedzialne społecznie niż same sieci neuronowe, to jest bardziej wyjaśnialne, kontrolowalne i mniej podatne na halucynacje czy uprzedzenia. Wizja przyszłości sztucznej inteligencji, gdzie będziemy tworzyć tylko coraz większe i potężniejsze wersje ChatGPT brzmi dla mnie dystopijnie. Mam szczerą nadzieję, że świadomość rosnącego ryzyka jakie niosą ze sobą takie modele popchnie naukę i przemysł do poszukiwania innych rozwiązań. Nie chodzi tu też o to by zaprzestać badań nad tym czy innym rodzajem sztucznej inteligencji – wręcz przeciwnie! Odkrywając nowe metody symboliczne i te oparte na sieciach neuronowych budujemy coraz lepszą podstawę dla ich hybryd. Potrzebujemy sztucznej inteligencji, to nie ulega wątpliwości. Duża część ostatnich odkryć i wynalazków nie byłaby możliwa bez pomocy komputerów. Nowe leki, poznanie granic fizyki, innowacyjne materiały, bardziej wydajne procesy produkcyjne, energooszczędna elektronika – lepsza sztuczna inteligencja może nam w tym wszystkim pomóc. Nie możemy jednak pozwolić na to, żeby to niesamowite narzędzie wymknęło się nam spod kontroli, albo żebyśmy przestali je rozumieć.

Piotr Sowiński – doktorant na Wydziale Matematyki i Nauk Informatycznych Politechniki Warszawskiej. W swoich badaniach skupia się na poprawianiu skalowalności i wydajności systemów SI opartych o grafy wiedzy. Współzałożyciel firmy NeverBlink, gdzie stara się przełożyć wyniki swoich badań na rozwiązywanie praktycznych problemów biznesowych

⁶ Sarker, M. K., Zhou, L., Eberhart, A., & Hitzler, P., (2021), Neuro-symbolic artificial intelligence, *AI Communications*, 34(3), 197-209

Celem Uczelnianej Oferty Dydaktycznej Centrum Studiów Zaawansowanych PW (UOD CSZ PW) jest poszerzenie wiedzy w wybranych kierunkach, a także pomoc i inspiracja w planowanej działalności naukowej. Program oferty adresowany jest obecnie do doktorantów we współpracy ze Szkołą Doktorską Politechniki Warszawskiej. Na propozycję UOD CSZ PW składają się m.in. cykle interdyscyplinarnych wykładów specjalistycznych i warsztatu badacza.

Merytoryczną opiekę nad UOD CSZ PW sprawuje Rada Programowa Centrum, którą tworzą naukowcy z Politechniki Warszawskiej, Uniwersytetu Warszawskiego, a także Polskiej Akademii Nauk.

Uczelniana Oferta Dydaktyczna Centrum Studiów Zaawansowanych Politechniki Warszawskiej

2024/2025

semestr
zimowy



- SZ1: Modele matematyczne procesów i przemian – prof. Stanisław Janeczko (PW)
- SZ2: From Neuron to Consciousness (English course) – dr Jacek Rogala (UW)
- SZ3: Rozwój matematyki na przestrzeni dziejów – prof. Wojciech Domitrz (PW)
- SZ4: Iluzja Wiedzy i Granice Poznania – prof. Stanisław Janeczko (PW)
- SZ5: Therapy for the 21st Century (English course) – prof. Ewa Bartnik (UW)

semestr
letni



- SL1: Wstęp do geometrii różniczkowej – prof. Wiesław Sasin (PW)
- SL2: Elementy mechaniki analitycznej – prof. Piotr Przybyłowicz (PW)
- SL3: Fundamentals of Computer Vision (English course) – prof. Agnieszka Jastrzębska (PW)
- SL4: Introduction to Singularity Theory (English course) – prof. Stanisław Janeczko (PW)

WBL1: Rozmowy i rozumowania – prof. Stanisław Janeczko (PW)

Zapisy na wykłady prowadzone są w Szkole Doktorskiej PW, zgodnie z harmonogramem ustalonym na rok akademicki, w systemie USOS.

*Szczegółowe informacje o terminach zapisów, uruchamiania wykładów oraz uaktualniona lista przedmiotów znajduje się na stronie internetowej Centrum
http://www.konwersatorium.pw.edu.pl/oferta/sz_doktorska.html*

Biuletyn Centrum Studiów Zaawansowanych „Profundere Scientiam” (online) nr 1 (19)/2025; ISSN 2956-9958
Pl. Politechniki 1, p.152-154, 00-661 Warszawa; e-mail: csz@pw.edu.pl, www.csz.pw.edu.pl

Zespół redakcyjny: Małgorzata Zielińska, Wanda Borkowska, Ilona Sadowska

Opieka merytoryczna: prof. Stanisław Janeczko

Projekt graficzny: Emilia Bojańczyk / Podpunkt | Opracowanie i skład: Małgorzata Zielińska / CSZ