



Dr Andrzej J. Kałka Uniwersytet Jagielloński

Rozmowa z autorem pracy:

„Development and Implementation of Computer-Aided Techniques of Data Modeling for Processing and Interpretation of Multicomponent Electronic Spectra Acquired for the Selected Organic Mixtures”

W jaki sposób zainteresował się Pan chemometrią?

Może zabrzmiało to dość szablono, ale głównie za sprawą mojego Promotora. Prawdę mówiąc, już na dość wczesnym etapie studiów zorientowałem się, że zagadnienia typowo kojarzone z chemią, takie jak synteza czy analiza konkretnych próbek, nie wzbudzają zbytnio mojego entuzjazmu. Zdecydowanie bardziej fascynujące wydały mi się (i nadal wydają) zagadnienia związane z odkrywaniem i opisem natury zjawisk, którym te próbki mogą ulegać, i dzięki którym powstały. To przekonanie rzuciło mnie wprost w ramiona chemii fizycznej, teoretycznej i komputerowej. W ścisłym sensie, w arkana chemometrii zostałem wprowadzony w ramach mojej pracy licencjackiej. Chociaż z początku ta tematyka wydawała mi się nieco abstrakcyjna, szybko przekonałem się, że odpowiednie wykorzystanie aparatu matematycznego oraz technologii komputerowej pozwala na wydobycie z danych, uzyskanych w pomiarach chemicznych, informacji niedostępnych „zwykłym śmiertelnikom”. Urzeczony tą perspektywą, w ramach swojej pracy naukowej postanowiłem poświęcić się rozwojowi tejże metodologii oraz jej praktycznemu zastosowaniu, które – jak już niestety wielokrotnie miałem okazję przekonać – nadal należy do rzadkości.

Co zadecydowało o wyborze pakietu MATLAB jako jednego z głównych wykorzystanych w badaniach środowisk obliczeniowych?

W pewnym sensie nie miałem w tej kwestii większego wyboru – środowisko chemometryczne już na dość wczesnym etapie rozwoju upodobało sobie właśnie ten konkretny pakiet, przez co MATLAB w naturalny sposób stał się (i nadal pozostaje) niejako domyślnym wyborem do prowadzenia obliczeń z zakresu chemometrii. Z perspektywy czasu utwierdzam się w przekonaniu, że wybór ten pozostaje całkiem słuszny. Trzon większości algorytmów stanowią bowiem obliczenia z zakresu algebry liniowej i macierzowej, do których MATLAB (skrót od MATrix LABoratory) został oryginalnie stworzony, tym samym oferując znaczną ilość funkcjonalnych dedykowanych rozwiązań. Poza konsolą, pakiet posiada także wbudowany interfejs graficzny (GUI), co – w porównaniu np. do pakietu R – znacząco ułatwia przetwarzanie oraz wizualizację danych, obejmujących chociażby tworzenie złożonych, wieloelementowych wykresów.

Jakie największe trudności napotkał Pan na drodze badawczej i w jaki sposób stawiał im Pan czoła?

Problemem, który chyba w największym stopniu pokrzyżował moje (i zapewne nie tylko moje) plany, była pandemia SARS-CoV-2, przypadająca na dwa pierwsze lata mojego doktoratu. Praktycznie całkowicie wstrzymana została wówczas krajowa i międzynarodowa wymiana badawcza, co skutecznie uniemożliwiło mi wyjazdy na staże, szkolenia czy konferencje. Na szczęście, wprowadzone ograniczenia nie wpłynęły w znaczący sposób na realizację prowadzonych przeze mnie badań – technologia komputerowa, na której w dużej mierze je oparłem, okazała się odporna na działanie koronawirusa i dobrze odnalazła się w nowej, „zdalnej” rzeczywistości. Drugi problem, o którym według mnie warto jest publicznie wspomnieć, stanowi trudność w pozyskaniu funduszy na prowadzenie własnych badań. Niestety, pula grantów przeznaczonych na sfinansowanie projektów autorstwa młodych naukowców – delikatnie rzecz ujmując – nie należy do największych. Tym samym uzyskanie środków na ich realizację jest często bardzo utrudnione lub wręcz niemożliwe. W moim przypadku bardzo pomocnym okazał się pro-

gram „Inicjatywa Doskonałości – Uniwersytet Badawczy” (IDUB), w ramach którego, w formie tzw. mini-grantów, wsparcie otrzymały projekty zgłoszone właśnie przez doktorantów.

Biorąc pod uwagę zamieszczenie opracowanych algorytmów w bazie RODBUK, jak ocenia Pan zasadność i potrzebę publikowania danych w otwartym dostępie?

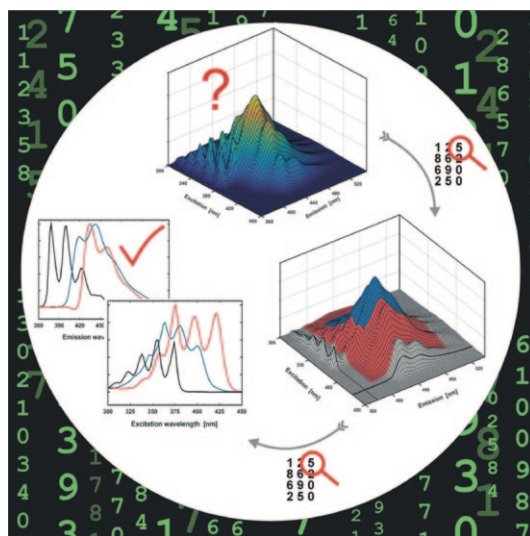
Inicjatywa tzw. „otwartej nauki” z założenia wydaje mi się bardzo słuszna – w mojej opinii wyniki rzetelnie prowadzonych badań w ogólności powinny być powszechnie dostępne dla wszystkich zainteresowanych. Niestety, problem stanowi praktyczna realizacja tego założenia. Publikowanie w formacie „Open Access” (OA) wiąże się z dość dużymi kosztami (rzędu kilku, a nawet kilkunastu tysięcy złotych), na pokrycie których spora część naukowców po prostu nie może sobie pozwolić. W tym aspekcie nie sposób zatem nie pochwalić idei otwartych repozytoriów danych (takich jak np. przywołany RODBUK), w których materiały związane z prowadzonymi badaniami mogą być bezpłatnie zdeponowane, a następnie udostępnione szerokiemu gremium potencjalnych odbiorców.

Na ile zasoby informatyczne udostępniane przez Cyfronet przysłużyły się prowadzonym badaniom?

O ile modelowanie danych eksperymentalnych przy użyciu „zwykłego” komputera nie powinno w praktyce nastęrczać większych problemów, o tyle nie mogę sobie wyobrazić prowadzenia z jego pomocą obliczeń kwantowo-chemicznych, stanowiących istotny element mojej pracy badawczej. Dla przykładu, na uzyskanie wyniku jednej z typowych symulacji (których w ramach doktoratu wykonałem dziesiątki, jeśli nie setki) na superkomputerze Ares potrzebowałem nieco ponad dwóch godzin – ta sama symulacja na przeciętnym PC zajęłaby mi szacunkowo pół doby. Wobec powyższego, mogę ze spokojnym sumieniem określić udostępnione zasoby obliczeniowe mianem nieocenionych. Przymuszczenie moją opinię podzielił również dziesiątki innych użytkowników infrastruktury PLGrid, z którymi – zwłaszcza w ostatnim czasie – dane mi było stawać w szranki o miejsce na podium w systemie kolejkowym slurm.

Co mógłby Pan doradzić początkującym naukowcom, dopiero rozpoczynającym lub planującym rozpoczęcie szkół doktorskich?

Przede wszystkim, aby szczerze przemyśleli swoją motywację i priorytety. Doktorat, szczególnie w ujęciu obowiązującej ustawy (szkoły doktorskie), oferuje sporo możliwości rozwoju naukowego oraz osobistego. Jednakże, aby w pełni z tych możliwości skorzystać, do czego oczywiście gorąco zachęcam, trzeba dysponować odpowiednią ilością czasu, a ten wygospodarować można w zasadzie tylko kosztem ograniczenia lub wręcz rezygnacji z regularnej pracy zarobkowej. To z kolei może odbić się negatywnie na doświadczeniu zawodowym, a co za tym idzie, także na pozycji na rynku pracy (poza sektorem B+R) zaraz po ukończeniu doktoratu. Trzeba zatem szczerze odpowiedzieć sobie na pytanie, czy w perspektywie kolejnych czterech lat ważniejszy jest dla mnie rozwój naukowy, czy zawodowy. Warto w tym miejscu hasłowo wspomnieć o programie tzw. doktoratów wdrożeniowych, oferujących w pewnym sensie opcję pośrednią pomiędzy dwiema powyższymi skrajnościami.



*Czym zajmuje się chemometria?
Ideogram zaczerpnięty z A.J. Kałka, A.M. Turek (2021),
J. Fluoresc. 31: 1599-1616*