ChemPharos: a comprehensive database of plastic-degrading enzymes, powered by Enalos Tools

Constantinos Papavasiliou,[1] Dimitris G. Mintis,[1,2] Andreas Tsoumanis,[1,3] Iseult Lynch,[2,4] Antreas Afantitis[1,2,3]

1. Introduction

The use of plastics, in everyday life, has contributed significantly to environmental pollution. In response to the escalating issue of plastic waste, natural degradation methods have emerged as potential waste management protocols. Notably, biodegradation by microorganisms has gained prominence as a preferred approach among environmentally conscious societies [1]. To address this environmental issue, we present ChemPharos, a database tailored for modelling-reading chemical datasets powered by the Enalos Cloud Platform [2] and demonstrate its use for rapid prediction of plastic degrading enzymes.

ChemPharos is a computational framework designed to store various data regarding chemicals, their properties and their functionalities, illustrated here via a dataset of enzymes that can degrade plastic polymers. The scope of the specific modelling-ready dataset is to harness cutting-edge Artificial Intelligence (AI) and Machine Learning (ML) tools for Micro and Nano-Plastics (MnNP) risk assessment as part of the EU-funded Marie Curie Training Network, Plastic Underground. The ultimate goal of the Plastic Underground instance of ChemPharos is to serve as a robust platform facilitating the seamless integration and automation of Molecular Dynamics (MD) simulations, enabling in-depth investigations into the impact of MnNP in soil and groundwater environments.

One of the key objectives of Plastic Underground is to fundamentally reshape MnNP risk assessment methodologies by integrating a comprehensive cloud-based platform with advanced data analytics and simulation functionalities. The deployment of ChemPharos aims to catalyse understanding of methods for remediation of plastic pollution in terrestrial and subterranean aquatic ecosystems. This platform is specifically designed to yield innovative insights into the behavioural dynamics of MnNPs and their subsequent environmental ramifications, thereby augmenting current strategies in environmental sustainability management.

[1] NovaMechanics Ltd, Nicosia 1070, Cyprus, info@novamechanics.com
[2] Entelos Institute Ltd, Larnaca 6059, Cyprus
[3] NovaMechanics MIKE, Piraeus 18545, Greece
[4] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham B15 2TT, United Kingdom

2. Data Collection

MySQL was selected as the Relational Database Management System (RDBMS) for ChemPharos due its known fast performance, scalability, reliability, and robust security features. MySQL, being open source, offers cost-effective management of large datasets and facilitates seamless integration with server-side java scripting [3]. Moreover, its compatibility with cloud environments ensures seamless integration for deployment and scaling within the cloud infrastructure. Additionally, MySQL provides tools for administration, monitoring, and optimization, aiding in efficient database management and performance monitoring [4].

Integrating data from ChEMBL, an existing open-source bioactivity database that is regularly updated [5], enriches ChemPharos by providing valuable insights for Micro and Nano Plastic (MnNP) risk assessment. By integrating ChEMBL into ChemPharos, the Plastic Underground project aims to enhance automated molecular simulations and enable explicit data insertion and modification with well-defined authorization mechanisms.

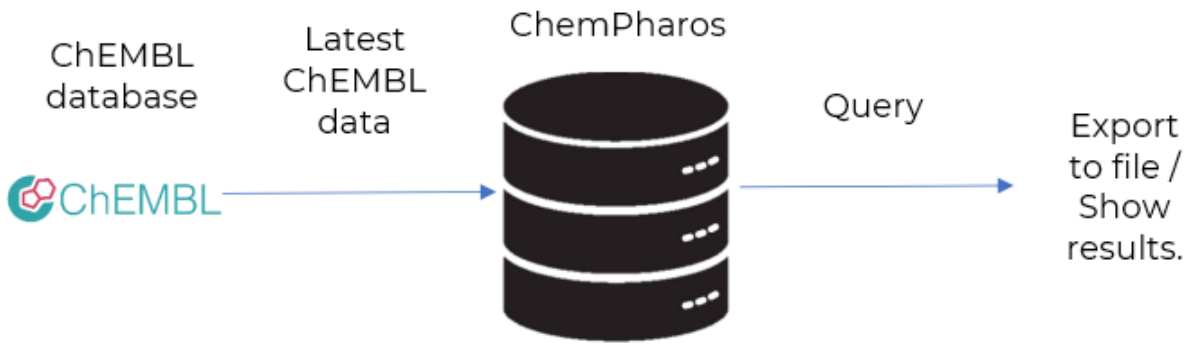

Figure 1: Proposed data workflow of ChEMBL and ChemPharos.

3. Tools

Java (https://www.java.com/en/) has been chosen as the backend server-side programming language of ChemPharos for its versatility, portability and robust ecosystem [6].

Javascript (https://www.javascript.com/) is used as the scripting language for visualisations, empowering the user interface with dynamic and interactive content. [7].

ZK Framework (https://www.zkoss.org/product/zk) and the ZK User Interface Markup Language (ZUL) has been selected for developing the front end due to its component-based architecture which facilitates modular development of the graphical user interface [8].

Keycloak (https://www.keycloak.org/) has been employed as the Identity and Access Management System for ChemPharos to ensure that only authorised individuals (curators) have access to sensitive data and functionalities. [9].

Wildfly (https://www.wildfly.org/) has been selected as the runtime server for deploying and running the Java based backend of ChemPharos.

4. User Interface (UI)

Figures 2 and 3 provide a shapshot of the user interface, and how the user selects the features of interest.

4.1 Search Configuration

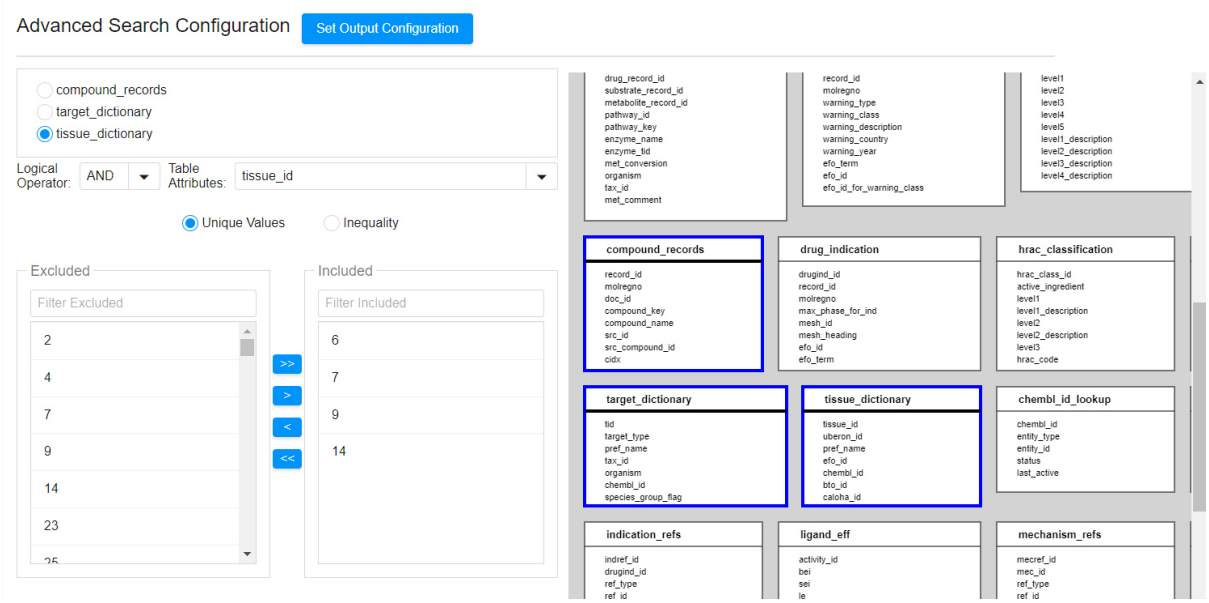

Figure 2: Advanced Search Configuration Page of ChemPharos.

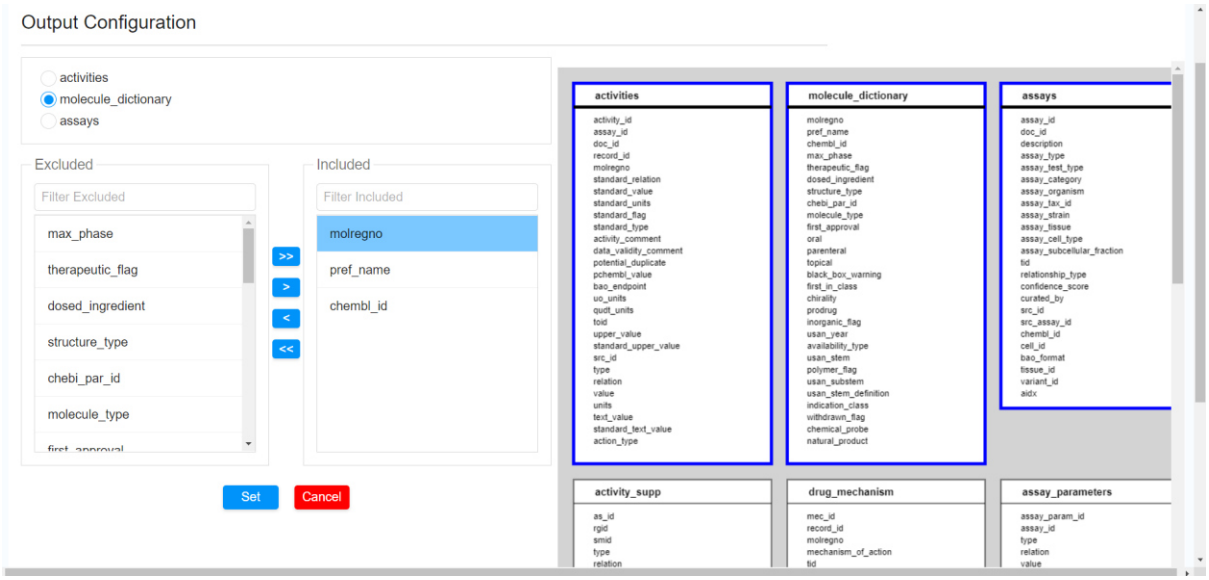

Figure 3: Output Configuration Page of ChemPharos.

5. Future Work

While ChemPharos is set for beta deployment on the Enalos Cloud Platform, upcoming enhancements are in the pipeline to elevate its capabilities, including:

- Optimisation of querying and export functionalities
- Paging of results

- AI and ML tools for MnNP risk assessments
- Automation of molecular dynamics simulations

These envisioned developments showcase ChemPharos' dedication to provide an adaptable, user-friendly platform that remains at the forefront of molecular data analysis and visualisation.

6. Conclusions

The ChemPharos platform provides an integrated and user-centric solution for molecular simulations and MnNP risk assessments. Fundamentally, the Plastic Underground instance of ChemPharos is designed not only to refine the process of MnNP risk evaluation but also to exemplify the progressive and collaborative ethos inherent in software development. This approach ensures that the platform remains adaptive and responsive, effectively addressing the dynamic needs and challenges faced by researchers and decision-makers in the scientific domain.

7. Acknowledgments

This work received funding from the European Union's Horizon Europe research and innovation program under the Marie Sklodowska-Curie Doctoral Network PLASTICUNDERGROUND (Grant Agreement No. 101072777)

8. References

[1] Bahl, S., Dolma, J., Singh, J. J., and Sehgal, S., "Biodegradation of plastics: A state of the art review.," *Materials Today: Proceedings,* vol. 39, pp. 31-34., 2021.

[2] Melagraki, G. and Afantitis, A., "Enalos InSilicoNano platform: an online decision support tool for the design and virtual screening of nanoparticles". RSC advances, 4(92), pp.50713-50725, 2014.

[3] Palanisamy, S., and SuvithaVani, P., "A survey on RDBMS and NoSQL Databases MySQL vs MongoDB," in *2020 International Conference on Computer Communication and Informatics (ICCCI).*, Coimbatore, 2020.

[4] MetricFire, "A Modern Guide to MySQL Performance Monitoring," 11 10 2023. [Online]. Available: https://www.metricfire.com/blog/a-modern-guide-to-mysql-performance-monitoring/. [Accessed 15 03 2024].

[5] Gaulton A., et. al., "The ChEMBL database in 2017," *Nucleic Acids Res,* pp. 45(D1):D945-D954, 2017.

[6] Baddam, P. R., Vadiyala, V. R., & Thaduri, U. R. , "Unraveling Java's Prowess and Adaptable Architecture in Modern Software Development.," *Global Disclosure of Economics and Business,* vol. 7(2), pp. 97-108, 2018.

[7] AIContentfy, “Unleash the Power of JavaScript: Build Interactive Websites,” 14 03 2024. [Online]. Available: https://aicontentfy.com/en/blog/unleash-power-of-javascript-build-interactive-websites. [Accessed 17 03 2024].

[8] Ochoa-Zezzatti, A., et al., “Improve Decision Support Using Adaptive Data Mining.,” in *2009 International Conference on Electrical, Communications, and Computers. IEEE*, Dubai, 2009.

[9] Chatterjee, A. and Prinz, A, “Applying Spring Security Framework with KeyCloak-Based OAuth2 to Protect Microservice Architecture APIs: A Case Study.,” *Sensors,* vol. 22, p. 1703, 2022.