|  |  |
| --- | --- |
| cetlogo ***CHEMICAL ENGINEERING TRANSACTIONS*** ***VOL. xxx, 2025*** | A publication ofaidiclogo_grande |
| The Italian Associationof Chemical EngineeringOnline at www.cetjournal.it |
| Guest Editors: David Bogle, Flavio Manenti, Piero SalatinoCopyright © 2025, AIDIC Servizi S.r.l.**ISBN** 979-12-81206-21-2; **ISSN** 2283-9216 |

Integrative digital platform to enhance the value of data, knowledge, and scientific expertise for biorefinery platforms

Amaya Saint-Boisa, Jean-Pierre Belauda\*, Rachid Ouareta, Claire Vialleb, Caroline Sablayrollesb, Patrice Buchec, Adrien Reauc, Alrick Oudotc, Cédric Baudritd, Christophe Fernandezd, Claire Mayer-Laiglec, Julien Cufic, Leslie Lhomondc, Magalie Webere, Liliana Ibanescuf, Stéphane Dervauxf

aLaboratoire de Génie Chimique, Université de Toulouse, CNRS, Toulouse, France

bLaboratoire de Chimie Agro-industrielle, Université de Toulouse, INRAE, Toulouse, France

cUMR IATE - Ingénierie des Agro-polymères et Technologies Émergentes, Montpellier, France

dI2M - Institut de Mécanique et d'Ingénierie, Talence, France

eBIA - Unité de recherche sur les Biopolymères, Interactions Assemblages, Nantes, France

fMIA Paris-Saclay - Mathématiques et Informatique Appliquées, Palaiseau, France

jeanpierre.belaud@toulouse-inp.fr

The French 3BCAR ICAM 2024-2026 project aims at leveraging digitalization tools and sustainability principles to obtain an integrated digital platform for bioeconomy actors. ICAM, in French, stands for Knowledge Engineering and Multidimensional Analysis. The project’s goal is to develop an ontology based integrative digital platform that enhances data, knowledge, and scientific expertise around bio-refinery platforms. This paper introduces PowderLib, an instantiation of ICAM’s integrative digital platform for vegetable powders. The platform stores data and knowledge on the production and valorisation of plant powders. It connects four modules: PO2, an experimental database, Makebook, a tool to store and manage electronic knowledge books, CAPEX, a decision tree-reasoning engine and AdCVcalc&visu, an environmental assessment and visualization tool. The modules are interoperable because they adopt the same vocabulary from the domain ontology PO2/TransformON, designed to describe transformation processes and observations of associated characterization experiments. A use case shows how the platform assists users in the discovery of vegetal powders that match given end-product functional specifications in a reverse engineering approach.

* 1. Introduction

Increasing population, globalization and urbanization have led to consumption patterns that emit significant amounts of C02, raising the atmosphere, earth and oceans’ temperatures. Climate change originates extreme environmental phenomena such as heatwaves or sea level rise. It also deepens social and economic inequalities. To tackle global warming and transition towards sustainable human activities, members of the United Nations devised 17 Sustainable Development Goals (SDGs) that lie at the core of the 2030 Agenda for Sustainable Development. The United Nations describes sustainable activities as those who enable meeting present needs without compromising future generations’ abilities to meet theirs. Shifting towards bio-economical systems aligns with these goals. Bio-economical systems are those that produce and transform biomass to reduce waste, obtain bioenergy, biomolecules and biomaterials. Biomass can be from agricultural, forestry or aquaculture origin. Bio-economy aims at developing renewable sources and sustainable technologies to minimize environmental impacts (“Our manifesto,” 2023).

The ICAM project is financed by the Carnot consortium of research institutions, which comprises 39 French public research organizations recognized for their ability and dedication to meet private companies’ R&D needs. The project is led by the Carnot 3BCAR cluster, which incorporates 18 R&D entities, ranging from laboratories to pilot plants. The Carnot 3BCAR leverages biotechnologies and green chemistry to support business innovation. Biotechnologies rely on living organisms or biological systems to derive innovative products and services. The Carnot 3BCAR’s expertise covers the entire value chain, from biomass production, biorefinery and the study of bio-based products’ functional properties. Biorefineries integrate biomass conversion processes to produce fuels, power, heat and value-added chemicals from biomass. Bio-based products or bioproducts are materials, chemicals, and energy derived from renewable biological resources such as plants, animals, marine organisms, or microorganisms.

Lignocellulosic biomass from plant-based biomass constitutes one of the most promising and available renewable resources on Earth. Lignocellulosic biomass is mainly composed of cellulose, hemicellulose and lignin (Phanthong et al., 2018). Biomass can be processed into numerous products for a wide range of applications such as food packaging (Lammi et al., 2018) or 3D printing (Le Duigou et al., 2020). The variety of available biomass requires a diverse range of specific and technical treatment and conversion processes to generate the desired bio-products. Experimental data, expert and academic knowledge on biomass production and transformation processes is growing but widely dispersed and stored in different online data repositories or books. Some of the expert knowledge is not even archived, only held among experts. To tackle this, the ICAM project is developing an integrated digital platform that eases knowledge and innovation transfer and diffusion.

PowderLib (“PowderLib Portal,” 2025) is an instantiation of the integrated digital platform that results from the ICAM project. PowderLib gathers knowledge on vegetable powders. The platform embeds an open source data repository conceived with the PO2/TransformON domain ontology (Weber et al., 2023)*.* The PO2/TransformON domain ontology comprises the Process and Observation Ontology (PO2) at its core. The PO2 ontology has been specialized to describe biomass transformation processes and enable characterizations of wastes, food and bioproducts. The PO2/TransformON ontology is aligned with principles from the Semantic Web, endorsed by the World Wide Web Consortium (W3C). In line with this, the open source data repository is an RDF (Resource Description Framework) graph database. Data shared and stored in the repository is thus FAIR (Findable, Accessible, Interoperable, Reusable) by construction. The PO2/TransformON ontology lies at the heart of PowderLib, since it connects and makes modules interoperable. The first module is the PO2 open source RDF graph data repository dedicated to biomass transformation and experimental characterization data. The second one, named Makebook (Baudrit et al., 2024), is an electronic knowledge book software that stores knowledge in the form of concept maps (cmaps), knowledge sheets, graphs, downloadable documents or web pages. The third one is a reasoning engine named CAPEX, which enables navigating decision trees to find explanations and solutions to defaults or qualities. The fourth module, under development, is a visualization and environmental assessment module. It leverages machine-learning techniques to facilitate analysis and sustainability assessments. It is inspired from previous work from (Prioux et al., 2023) and (Munch et al., 2022b).

The following section illustrates PowderLib, the integrative digital platform that results from the ICAM project. Section 3 illustrates one use case of the platform and section 4 concludes the paper.

* 1. Integrative digital platform

The ICAM project has developed an integrative digital platform within an Open Science approach, aimed at various types of stakeholders who wish to access experimental data, find answers to scientific questions and obtain technical recommendations for the implementation of sustainable biorefineries. It facilitates knowledge and innovation transfer and diffusion. The platform is regularly updated to incorporate new knowledge. It advises industries through innovative technology-related decision-making methods. Powderlib is an instantiation of ICAM’s platform that is specialized in the production and valorization of plant-based vegetal powders. Its interoperable modules simplify the organization of academic knowledge, the structuring of operational expertise and grant access to FAIR experimental data. It will also provide sustainability analyses that take into account multi-dimensional environmental and potentially economic factors in order to aid and support users of the platform in their choice of vegetable powders and transformation technologies and itineraries.

The targeted users of the platform are:

1. R&D engineers from bioeconomy companies, who will eventually have access to a decision support system for explanation and technological reasoning, integrating sustainability indicators.
2. Data scientists interested in biorefinery applications within an Open Science framework, who will have access to a structured database for data mining and knowledge discovery.
3. Students specializing in bioeconomy and educators, who will benefit from a knowledge and expertise transfer tool.

ICAM’s integrated digital platform relies on a common ontological resource, PO2/TransformON to make its modules interoperable. The ontology enables annotating and structuring knowledge with a single reference framework across its four modules or software components shown in Figure 1: CAPEX, Makebook, PO2 Manager and AdCVcalc&visu.



Figure 1: Digital components of PowderLib, an instantiation of ICAM’ integrated digital platform

PowderLib encompasses two knowledge compartments (generic and specific) for each module shown in Figure 1. The generic knowledge compartment of each module facilitates the creation of user-specific knowledge that is contained in specific knowledge compartments for each user. The generic knowledge compartments include an open science section (for which open access is granted under an accepted license) and a protected section. For example, generic decision trees can be used and specialized by companies. The following paragraphs detail each module’s aim, mechanism and functionalities.

* PO2/TransformON Ontology :

PO2/TransformON (Weber et al., 2023) is a domain ontology, based on the Process and Observation Ontology (PO2), designed to address the data harmonization challenge in the context of food, feed, bioproducts and biowaste engineering. PO2/TransformON facilitates data integration to move towards circular sustainable bioeconomy and nexus-oriented approaches. PO2 describes processes (step, input/output components) and observations associated with experiments of transformation processes. It has been specialized to contain vocabulary that structures and assists the description of biomass transformation processes and characterization processes of food, feed, bioproducts and wastes. PO2/TransformON’s vocabulary is inspired from existing food reference catalogues such as the European food classification system (FoodEx2) or the European Waste Catalogue. These have been used together with other international nomenclatures to define PO2/TransformON’s vocabulary and semantics in a World Wide Web consortium (W3C) format that provides system interoperability and software-driven intelligence.

* Makebook :

The Makebook module is a software tool that stores electronic knowledge books, which are structured hypertext networks with semantic capabilities that connect digital documents among them and facilitate navigation, knowledge representation and retrieval. These books enable users to organize knowledge in a given domain using concept maps (cmaps), process graphs, influence graphs, downloadable documents, web pages, and hypermedia knowledge sheets.

* CAPEX :

The CAPEX reasoning engine module is integrated into ICAM’s digital platform with the aim of providing technical recommendations for situations of interest (default or desired situations). This module implements a four-step iterative methodology. The first step lists default/desired qualities of vegetable powders and draws mind maps that identify potential causes and links between situations of interest associated to the functional specifications. The second step performs and computes individual efficiency evaluations for the technical solutions identified for the situations of interest. The third step automatically translates the mind maps into decision trees stored in an RDF knowledge graph. The fourth step relies on a multi-criteria decision-support system (MCDSS) that accounts with a graphical user interface that offers an explanatory and an action view. The explanatory view enables the users to navigate through decision trees to identify potential causes for situations of interest and technical actions that can help correct or reach the situations of interest. The action view enables multi-criteria filtering, action ranking and identification of possible side effects. The methodology is iterative because the users can update the mind maps of the first step to improve results of the MCDSS. See (Buche et al., 2023) for more details on the CAPEX reasoning engine.

* PO2Manager :

PO2Manager is the software tool that contains experimental data and knowledge associated to transformation processes. It is a data repository structured with vocabulary from the PO2/TransformON ontology. The present version of PowderLib’s PO2Manager data repository contains data on the production and transformation processes of vegetable powders (Munch et al., 2022a). PO2Manager software tool enables users to enrich the data repository by facilitating the annotation of data that is stored in the PO2 RDF (Resource Description Framework) database component. SPOOQ is a software web-based tool implemented to assist users in writing SPARQL queries to interrogate PO2Manager’s data repository. SPOOQ also allows users to publish and share data on INRAE’s, the French national institute for agricultural and environment, data warehouse, component of Research Data gouv, the French open data platform (<https://www.data.gouv.fr/fr/>).

* AdCVcalc&visu :

The AdCVcalc&visu module is being developed to enable users to assess and compare the sustainability of different biomass transformation itineraries. The module is inspired from previous work coupling big data, artificial intelligence and Life Cycle Assessment (LCA) described in (Prioux Nancy et al., 2023) and (Prioux et al., 2023). The approach described in (Prioux Nancy et al., 2023) consists of five main steps: objective and boundary definition, data architecture and enrichment, sustainability assessment, result visualization and analysis and decision-making. The approach implements neural networks to complete missing data. Partial dependence plots (PDP) and individual conditional expectation plots (ICE) are leveraged to examine and improve the neural networks. Bayesian networks could also be implemented to augment missing data. They could be combined with knowledge and data structured with the PO2 Ontology to represent transformation processes and account for uncertainty and causality in transformation processes to answer expert’s queries (Munch et al., 2022b). The analysis and visualization part of the AdCVcalc&visu module could harness traditional dimensionality reduction techniques (Multi-Dimensional Scaling) and unsupervised clustering methods (k-means) as described in (Prioux et al., 2023) to help users understand the complexity of different processing trajectories and derive links between processes, inputs and impacts. The final prototype of the model should adopt vocabulary and semantics from the PO2/TransformON ontology.

Section 3 depicts one use case that illustrates the functionalities and novelty of the integrated digital platform.

* 1. Use case : finding vegetable powders with given specifications

This section details one use case of PowderLib (<https://ico.iate.inrae.fr/powderlib_portail/>), the instantiation of ICAM’s integrative digital platform for vegetable powders. The use case shows how the Makebook module and the SPOOQ software tool associated to the PO2Manager experimental database are used together to identify vegetable powders with given characteristics. The use case puts into play a data scientist that wants to discover plant biomasses with the following input variable profile: cellulose content between 25% and 33%. The data scientist has established this specification with a machine learning model trained with data available in the platform to discover relationships between input and output variables of composite manufacturing process.

To perform the search the user starts by navigating the Makebook module. Figure 2 shows the conceptual mind map that the user encounters at the homepage of PowderLib’s Makebook. Since the user is looking for vegetable powders that match a given specification, the user opens the conceptual mind map associated to the Data Papers concept linked to the vegetal powders concept in Figure 1. The conceptual mind map of the Data Papers concept enables the discovery of the existence of a dataset named ““Planet-Milling\_itineraries\_for\_a\_collection\_of\_crop\_byproducts”, which is available on a public repository and can be queried through the SPOOQ software tool. Figure 3 shows the query expressed through SPOOQ’s graphical user interface to find biomasses with a cellulose content between 25% and 33%. Figure 4 shows the result of the query. Two milling itineraries are obtained by interrogating PO2Manager’s experimental database through SPOOQ. The milling itineraries are associated to hemp core (HC) and rice husk (RH) biomasses.s



Figure 2: Conceptual mind map of the Makebook's homepage



Figure 3: The user's query performed through SPOOQ's graphical user interface

Figure 4 shows that two milling itineraries applied to a Hemp Core or Rice Husk enable obtaining vegetable powders with the required cellulose content specified in the users’ query.



Figure 4: Results of the user's query that looks for vegetable powders with a given cellulose content

* 1. Conclusions

To conclude, this paper has presented and illustrated Powderlib, an instantiation of ICAM’s integrative digital platform designed to support innovation transfer around vegetal powders and aid in the eco-design of sustainable transformation processes. Powderlib adopts vocabulary and semantics from the PO2/TransformOn ontology, which has been designed to address the data harmonization challenge around feed, food, bioproducts and waste. Four modules are linked together using vocabulary and semantics from the PO2/TransformOn ontology. The CAPEX module is a reasoning engine designed to provide technological recommendations. The Makebook software tool stores electronic knowledge books that structure information as RDF graphs. PO2Manager is a software application that enables storing data in data repositories structured with the PO2/TransformOn ontology. SPOOQ is a web-based tool conceived to ease queries of repositories contained in PO2Manager. AdCVcalc&visu is a visualization and environmental assessment model under development that leverages AI methodologies to augment missing data, assess environmental impacts of different transformation itineraries and recommend pathways accounting for uncertainty and casual relations. Present work is focused on addressing the challenge of collecting and obtaining the necessary data to compute multi-criteria LCA environmental assessment indicators. Present work is also focused on the development of the AdCVcalc&visu module to enable the sustainability assessment and comparison of vegetal powders’ transformation pathways.

Perspectives are to extend the PowderLib platform with data and knowledge related to vegetal powders and transformation processes for the production of biocomposites used for food packaging and 3D printing. A potential limitation of the current platform is the introduction of new data in the Makebook and PO2/Manager modules. Since the introduction of new data is currently done manually. A perspective is to create an interface so that researchers can input new data as they produce it and publish articles.

Acknowledgments

This work acknowledges funding from the Carnot 3BCAR and ANR PEPR FAIRCARBON SLAM’B research institutions.

References

Baudrit, C., Fernandez, C., Couteaux, J., Buche, P., Bel, N., Charles, C., Notz, E., 2024. Cheese eK-Book: a new web-based medium for capitalising on, structuring and transferring cheesemaking knowledge and know-how. Journal of Documentation.

Buche, P., Couteaux, J., Cufi, J., Destercke, S., Oudot, A., 2023. Integrating collective know-how for multicriteria decision support in agrifood chains—application to cheesemaking. Front. Artif. Intell. 6.

Lammi, S., Le Moigne, N., Djenane, D., Gontard, N., Angellier-Coussy, H., 2018. Dry fractionation of olive pomace for the development of food packaging biocomposites. Industrial Crops and Products 120, 250–261.

Le Duigou, A., Correa, D., Ueda, M., Matsuzaki, R., Castro, M., 2020. A review of 3D and 4D printing of natural fibre biocomposites. Materials & Design 194, 108911.

Munch, M., Buche, P., Dervaux, S., Breysse, A., Berthet, M.-A., David, G., Lammi, S., Rol, F., Viretto, A., Angellier-Coussy, H., 2022a. Biocomposites from poly(3-hydroxybutyrate-co-3-hydroxyvalerate) and lignocellulosic fillers: Processes stored in data warehouse structured by an ontology. Data in Brief 42, 108191.

Munch, M., Buche, P., Dervaux, S., Dibie, J., Ibanescu, L., Manfredotti, C., Wuillemin, P.-H., Angellier-Coussy, H., 2022b. Combining ontology and probabilistic models for the design of bio-based product transformation processes. Expert Systems with Applications 203, 117406.

Our manifesto, 2023. . Bioeconomy For Change. URL https://www.bioeconomyforchange.eu/en/our-manifesto/ (accessed 2.14.25).

Phanthong, P., Reubroycharoen, P., Hao, X., Xu, G., Abudula, A., Guan, G., 2018. Nanocellulose: Extraction and application. Carbon Resources Conversion 1, 32–43.

PowderLib Portal [WWW Document], 2025. URL https://icotest.iate.inrae.fr/powderlib\_portail/ (accessed 2.14.25).

Prioux, N., Ouaret, R., Hetreux, G., Belaud, J.-P., 2023. Environmental assessment coupled with machine learning for circular economy. Clean Techn Environ Policy 25, 689–702.

Prioux N., Ouaret, R., Belaud, J.-P., 2023. Environmental Assessment of Glucose Production Using Neuronal Networks. Chemical Engineering Transactions 105, 301–306.

Weber, M., Buche, P., Ibanescu, L., Dervaux, S., Guillemin, H., Cufi, J., Visalli, M., Guichard, E., Pénicaud, C., 2023. PO2/TransformON, an ontology for data integration on food, feed, bioproducts and biowaste engineering. npj Sci Food 7, 47.