

Supplementary

Maarten Langen¹ and Vera van Noort^{1,2}

¹KU Leuven, Department of Microbial and Molecular Systems, Computational Systems Biology, Leuven 3001, Belgium

²Leiden University, Institute of Biology Leiden (IBL), Leiden 2333 BE, The Netherlands

August 2026

Output schema

Table S1: Output schema generated by Metacontextify. For each output column, the table lists the source of the data, a description of the field, and a representative example value.

column name	data source	description	example value
protein_id	MGnify	Protein identifier for MGnify	MGYP000261684433
assembly_id	MGnify	Assembly identifier for MGnify	ERZ534239
sample_id	MGnify	Sample identifier for MGnify	SAMN06268884
ena_sample_id	ENA	Protein identifier for ENA	SAMN06268884
is_marine	Metacontextify	Computed by Metacontextify based on the Biome property associated with MGnify Sample	True
lat	MGnify, ENA	Latitude of a sample	37.474067
lon	MGnify, ENA	Longitude of a sample	-121.973033
sample_date	MGnify, ENA	Sampling date of a sample	2011-12-09
depth_raw	MGnify, ENA	Raw value for depth from MGnify or ENA.	50-100 m
temperature_metadata	MGnify, ENA	Temperature retrieved from MGnify or ENA	11.7
salinity_metadata	ENA	Salinity retrieved from the ENA	23.8
ph_metadata	MGnify, ENA	pH retrieved from MGnify or ENA	7.6
depth	Metacontextify	Parsed depth. Tries to parse units (m and km), depth ranges and decimal separators. If empty, Metacontextify used surface for Copernicus property retrieval.	75

temperature_copernicus [C]	Copernicus	Temperature retrieved from Copernicus in degrees Celsius	11.250434897840023
salinity [psu] or salinity_copernicus [psu]	Copernicus	Salinity retrieved from Copernicus in degrees Celsius. ‘_copernicus’ suffix when ‘salinity_metadata’ column is present.	19.30829796474427
is_grid_mean_physical	Metacontextify	Boolean flag representing whether Metacontextify had to apply 3x3 grid averaging	False
is_depth_approx_physical	Metacontextify	Boolean flag representing whether Metacontextify had to use the value at bathymetry	False
phosphate [mmol.m-3]	Copernicus	Phosphate concentration retrieved from Copernicus in $\text{mmol}\cdot\text{m}^{-3}$	0.23928465694189072
nitrate [mmol.m-3]	Copernicus	Nitrate concentration retrieved from Copernicus in $\text{mmol}\cdot\text{m}^{-3}$	2.0118019580841064
oxygen [mmol.m-3]	Copernicus	Phosphate concentration retrieved from Copernicus in $\text{mmol}\cdot\text{m}^{-3}$	236.86504364013672
is_grid_mean_daily_chemical	Metacontextify	Boolean flag representing whether Metacontextify had to apply 3x3 grid averaging	False
is_depth_approx_daily_chemical	Metacontextify	Boolean flag representing whether Metacontextify had to use the value at bathymetry	False
ph_copernicus [1]	Copernicus	pH retrieved from Copernicus	8.039353370666504
phytoplankton [mmol.m-3]	Copernicus	Phytoplankton concentration in carbon in $\text{mmol}\cdot\text{m}^{-3}$	1.0517327785491943
is_grid_mean_monthly_chemical	Metacontextify	Boolean flag representing whether Metacontextify had to apply 3x3 grid averaging	False
is_depth_approx_monthly_chemical	Metacontextify	Boolean flag representing whether Metacontextify had to use the value at bathymetry	False