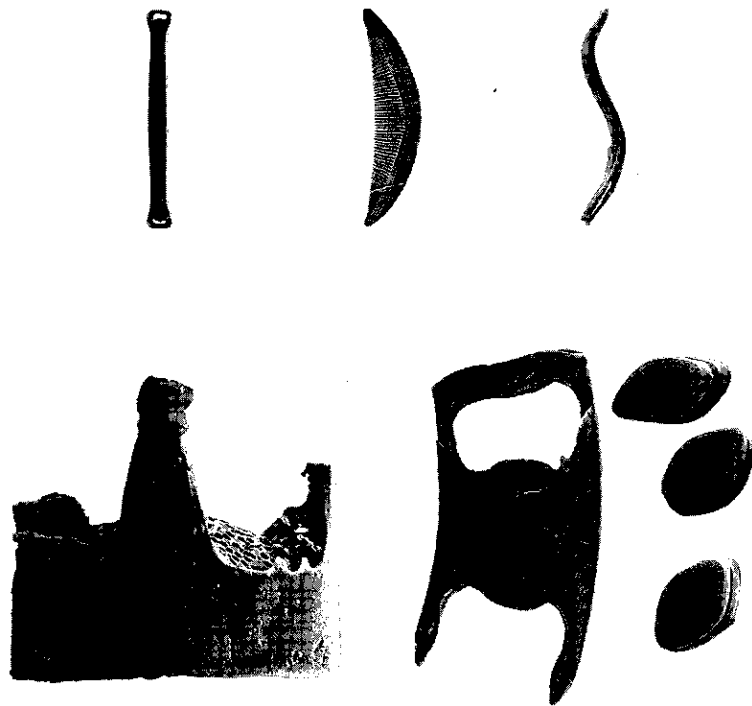


# Abstract Book



26<sup>th</sup> INTERNATIONAL  
DIATOM SYMPOSIUM

Yamagata, Japan 28 Aug.-2 Sept. 2023

LOCAL ORGANISING COMMITTEE  
(Yamagata)

Richard Jordan  
Kenta Abe  
Motohisa Fukuda  
Yangguang Gu  
Hisako Kubota-Kawai  
Mami Nomura  
Jun Yokoyama

with help from

Takuto Ando  
Andrea Burfeid-Castellanos  
Shijun Jiang  
Shinya Sato

# The new University of Duisburg-Essen freshwater diatom image data set (UDE Diatoms I)

\*Michael Kloster<sup>1</sup>, Andrea Burfeid-Castellanos<sup>1</sup>, Danijela Vidacović<sup>1,2</sup>, Ntambwe Albert Serge Mayombo<sup>1</sup>, Mimoza Dani<sup>1</sup>, Aishwarya Venkataramanan<sup>3,4</sup>, Cédric Pradalier<sup>4</sup>, Martin Laviale<sup>3</sup> & Bánk Beszteri<sup>1</sup>

<sup>1</sup>University of Duisburg-Essen, Faculty of Biology, Phycology Group, Essen, Germany

<sup>2</sup>University of Belgrade, Institute of Chemistry, Technology and Metallurgy, National Institute of the Republic of Serbia, Belgrade, Serbia

<sup>3</sup>Université de Lorraine, CNRS, LIEC, Metz, France

<sup>4</sup>Georgia Tech Lorraine, CNRS IRL 2958, Metz, France

\* michael.kloster@uni-due.de

Here, we introduce a collection of more than 83,000 light microscopy images of individual diatom valves or frustules, covering a broad range of taxa and morphology in about 320 samples from 15 different river and lake ecotypes. The diatoms were imaged at high resolution ( $< 0.1 \mu\text{m}/\text{pixel}$ ) by transmitted light bright-field microscopy, with focus stacking to artificially increase focal depth up to  $25 \mu\text{m}$ , allowing simultaneous observation of valve ornamentation and shape. Taken from a real-world setting, the images partly also include debris, mineral particles or other diatoms. Four experts identified over 500 diatom species; more than 100 species are represented by at least 100 specimens and nearly 150 by at least 50 specimens each. This data set is about one order of magnitude larger than previously published diatom data sets, and its high interspecies similarity makes it a valuable resource e.g. for benchmarking fine-grained out-of-distribution (OOD) detection, on which we present a new approach.