

# MIE Prediction Tool Webinar

## Intro on Education within the EU NSC

- ▶ NanoCommons Knowledge Infrastructure
- ▶ Available material on AOP-related topics
- ▶ What's next?

## MIE Prediction Tool

- ▶ Webinar by Abhijit

## Take home

- ▶ Q&A



Martin Himly (PLUS)  
Chair WG-A Education,  
Training, Communication

[www.nanosafetycluster.eu](http://www.nanosafetycluster.eu)

*NanoCommons Webinar, 2021-05-18, #nanocommons*

by **Abhijit Dasgupta**, *University College Dublin, IE*

*This project has received funding from the European Union Horizon 2020 Programme (H2020) under grant agreement no. 731032.*

CC-BY 4.0 International

# The idea – problem & solution

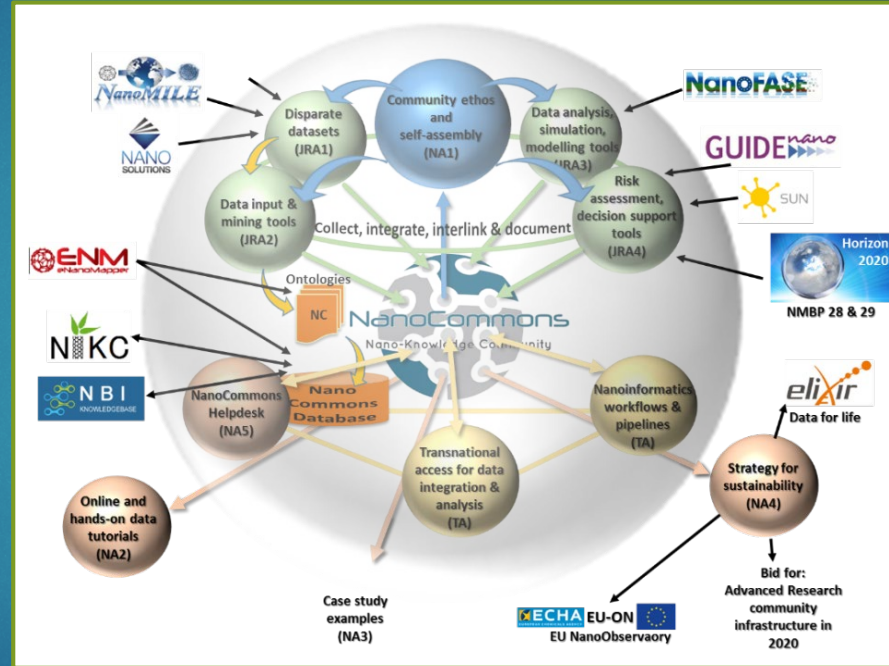


NanoCommons  
Nano-Knowledge Community

**Nanotechnologies** are a major area of investment & growth for the European economy

Knowledge and data remain fragmented and inaccessible hampering progress

Read-across approaches are currently absent for NMs, but would reduce the cost and time of nanosafety research and regulation



NanoCommons is creating an e-infrastructure platform for reproducible science, enhancing data integration & enabling nanoinformatics workflows to address these gaps



NanoCommons is **integrating and developing tools and services** for use by the nanomaterials communities and beyond

NanoCommons provides **Consultancy & Trainings** covering the entire nanosafety data life cycle:



Experimental Workflows  
Design & Implementation



Data Processing  
& Analysis



Data Visualisation  
& Predictive Toxicity



Data Storage  
& Online Accessibility

These tools & services can be **accessed** through the **NanoCommons Transnational Access (TA) scheme**

 [info@nanocommons.eu](mailto:info@nanocommons.eu)

# NanoCommons User Guidance Handbook

Home / e-Infrastructure / NanoCommons User Guidance Handbook



NanoCommons  
Nano-Knowledge Community

Overview Data management Nanoinformatics Workflows Electronic lab notebooks Ontologies

**NanoCommons**  
Nano-Knowledge Community

**The European Nanotechnology Community Informatics Platform: Bridging data and disciplinary gaps for industry and regulators**

The project has received funding from the European Union Horizon 2020 Programme (1010220) under grant agreement No. 721522

**Overview**

- H2020 research infrastructure project
- Call topic: NPM-NI-02-2019-2021: Integrating Activities for Advanced Communities

Download [2.19 MB]

**US - EU**  
Collaborative Community Forum

2020 U.S.-EU NanoEHS COR Workshop: Bridging Insights and Perspectives  
November 16-17, Virtual Meeting

2020 U.S.-EU NanoEHS COR Workshop: Bridging Insights and Perspectives  
September 16-17, Virtual Meeting

**NanoCommons Desktop Hackathon - M...**

**NanoCommons**  
Knowledge Community

For access to Services, Resources & Events  
→ [Visit our NanoCommons Knowledge Community](#)

NanoCommons – How can we help you?

[www.nanocommons.eu](http://www.nanocommons.eu)

Working Towards a Harmonized Nanosafety E-Infrastructure for Data and In Silico Tools, 2020 U.S.-EU NanoEHS COR Workshop: Bridging Insights and Perspectives, September 16-17, Virtual Meeting

You can still register to see the recordings.

Intro to NanoCommons by Martin Himly, PLUS (at the Jaqpot Hackathon)

zenodo

Online training tools for nanosafety assessment – NanoCommons for



# Events



NanoCommons  
Nano-Knowledge Community

Category



Organisations involved



Role



Filter

Reset

Submit an event



Services

Library

Events

About

Conference

22 – 26 Aug 2021 / Maastricht, NL

**Online MIE Prediction Tool Webinar**

Webinar, Training, workshop (organized by NanoCommons)

18 May 2021 / online event

United Kingdom Research and Innovation (UKRI)

National University of Ireland Dublin - University College Dublin (NUID-UCD)

BioNanoNet (BIONANONET)

Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)

Online QSAR Modelling Hackathon by Easy Access to Jaqpot: Deploy your model as a web service in a few minutes

Webinar, Hackathon, Training, Workshop (organized by NanoCommons)

13 Apr 2021 / online event

The University of Birmingham (UoB)

United Kingdom Research and Innovation (UKRI)

National Technical University of Athens (NTUA)

BioNanoNet (BIONANONET)

Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)

**Online SPARQL Access to WikiPathway and AOP-Wiki Webinar**

Webinar, Hackathon, Training, workshop (organized by NanoCommons)

23 Mar 2021

United Kingdom Research and Innovation (UKRI)

BioNanoNet (BIONANONET)

Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)

NanoSafety  
Cluster



# Events



NanoCommons  
Nano-Knowledge Community

## EU Nanosafety Cluster Education Day @ NanoSAFE 2020 Digital Conference

Webinar, Training, Workshop (organized by NanoCommons)  
16 Nov 2020

The University of Birmingham (UoB)  
Edelweiss Connect GmbH (EwC)  
United Kingdom Research and Innovation (UKRI)  
National Technical University of Athens (NTUA)  
LEITAT - ACONDICIONAMIENTO TARRASENSE ASSOCIACION (LEITAT)  
Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)  
NovaMechanics Ltd (NovaM)  
Universiteit Maastricht (UM)

## 2020 U.S.-EU NanoEHS COR Workshop: Bridging Insights and Perspectives

Conference, Workshop  
16 - 17 Sep 2020

The University of Birmingham (UoB)  
Edelweiss Connect GmbH (EwC)  
National University of Ireland Dublin - University College Dublin (NUID-UCD)  
Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)  
Universiteit Maastricht (UM)

## Online Electronic Lab Notebook basics - Hackathon

Webinar, Hackathon, Training, Workshop (organized by NanoCommons)  
14 Jul 2020

The University of Birmingham (UoB)  
United Kingdom Research and Innovation (UKRI)  
BioNanoNet (BIONANONET)  
Universitat Salzburg / Paris Lodron University of Salzburg (PLUS)

## SmartNanoTox International Online Conference

24 - 24 Jun 2020 / Online

The University of Birmingham (UoB)  
BioNanoNet (BIONANONET)

## Online Jqpot Hackathon - Take your research from the bench to the

The University of Birmingham (UoB)





Home / Events

## Date Added

Anytime

+ Register event

Subscribe ▾

? What are events in TeSS?

⏸ Hide past events

Search events...



## Scientific topic

Ontology and terminology

1

## Event type

Workshops and courses

6

Meetings and conferences

1

## Online

 Only show online events

## Country

Greece

1

## Organizer

www.nanocommons.eu

3

NanoCommons

2

Martin Himly, PLUS, Chair o...

1

Martin Himly, PLUS, Chair o...

1

## City

Athens

1

## Sponsor

EU H2020

4

EU H2020 NanoCommons

1

## Target audience

## 8 events found

Content provider: NanoCommons X

Grid

Map



Earliest



Ontological Annotation of Datasets: "1st NanoCommons Hackathon"

9 October 2018

Athens, Greece



Annotating Your Experimental Data workshop

28 April 2020



Online Jaqpot Hackathon

3 June 2020



Getting into using ELNs for experimental and computational workflows

14 July 2020



EU Nanosafety Cluster Training Day @ NanoSAFE 2020 Digital Conference

23 November 2020



EU Nanosafety Cluster Education Day @ NanoSAFE 2020 Digital Conference

16 November 2020



Online QSAR Modelling Hackathon by Easy Access to Jaqpot

13 April 2021



Online SPARQL Access to WikiPathways and AOP-Wiki Webinar

23 March 2021



December 10, 2020

Presentation Open Access

Edit

New version

# EU NanoSafety Cluster Education Day @ NanoSafe 2020

Dr Martin Himly

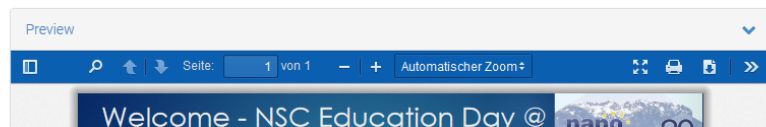
The EU NanoSafety Cluster WG-A on Education, Training, and Communication organized this Education Day on Nov 16 as satellite event of the NanoSafe Digital Congress 2020. This day acted as guidance for the entire NanoSafety community, including young researchers, to highlight how individual research projects fit as a puzzle piece into the wider picture. It was an orientation-giving and educational event depicting the overall strategy behind NanoSafe(ty).

## Means to achieve this included:

- to offer a **WG-overarching education/communication/discussion** event involving the audience *via* interactive sessions;
- to layout ways to **go beyond** with anything the nanosafety community have learned/developed to serve the emerging topics of Horizon Europe (emerging contaminants incl. microplastics, nanomedicine, safety assessment of novel/innovative/advanced materials for tomorrow along their entire life cycle);
- to foster participation in creating better sustainable materials (than e.g. nanosilver in socks), technologies, medical approaches, etc.
- to exhibit the perspectives of the NSC Working Groups and the different currently ongoing projects;
- to be **as interactive as possible** using **hands-on** activities, e.g., by showing **how-to** operate e-tools, upload/retrieve data to/from repositories & perform models;
- to facilitate vivid contribution to discussions using survey tools such as Mentimeter, WooClap, VoxVote, etc.
- to offer different perspectives in **pro/contra discussions**, e.g., by defining **challenger - defender roles taken by experts on specific topics** or evtl. by dividing attendees into zoom breakout rooms
- to show application of emerging NRG frameworks or SbD tools;
- to **request feedback** to user interfaces and enable stakeholder involvement

We herewith share and document the educational materials for later use by the scientific community.

Martin Himly, Chair of the EU NSC Working Group A with gratitude to all presenters for their contributions!



- ### Communities
- EU NanoSafety Cluster [Remove](#)
  - H2020 Infrastructure Project NanoCommons [Remove](#)
  - H2020 NMBP-15 Project SbD4Nano [Remove](#)
  - H2020 Project SABYDOMA [Remove](#)
  - Horizon2020 GRACIOUS [Remove](#)
  - Nano Risk Governance [Remove](#)
  - NanoSolveIT H2020 [Remove](#)
  - Nanoinformatics Project [Remove](#)

35

views

60

downloads

[See more details...](#)

Indexed in

Publication date:

December 10, 2020

DOI:



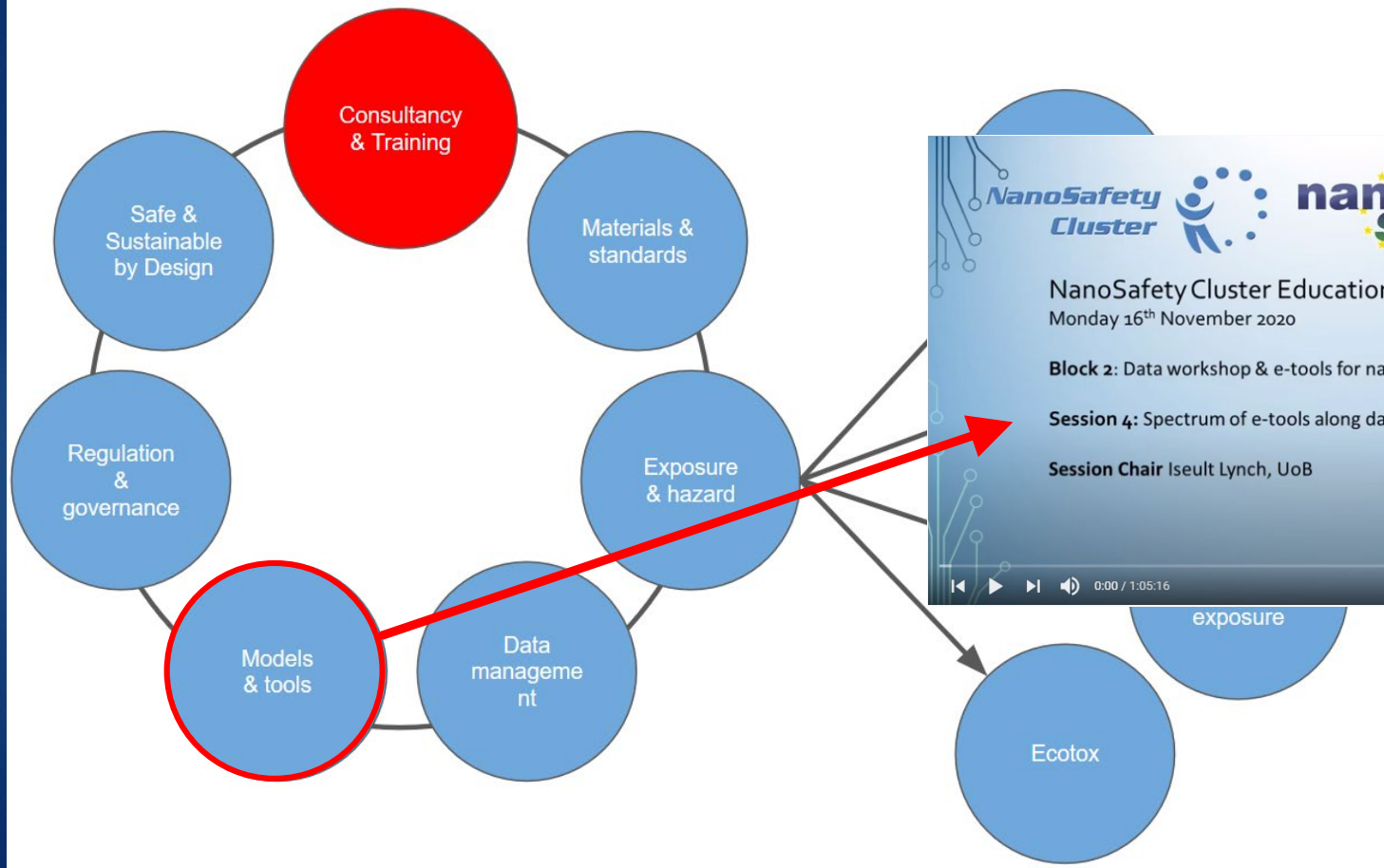
NanoCommons  
Nano-Knowledge Community




# ...aligned with the EU NSC WGs



NanoCommons  
Nano-Knowledge Community



**NanoSafety Cluster**  **nanoSAFE' 20**

**NanoSafety Cluster Education Day**  
Monday 16<sup>th</sup> November 2020

**Block 2:** Data workshop & e-tools for nano & beyond

**Session 4:** Spectrum of e-tools along data life cycle

**Session Chair** Iseult Lynch, UoB

0:00 / 1:05:16



- Start
- Entdecken
- Abos
- Mediathek
- Verlauf
- Später ansehen
- Videos, die ich mag

# NanoSafety Cluster



NanoSafety Cluster



NanoSafety Cluster  
28 Abonnenten

ABONNIEREN

- ÜBERSICHT
- VIDEOS
- PLAYLISTS
- KANÄLE
- KANALINFO

- ABOS
- UNIVERSAL MUSI...
  - Kanäle finden
- MEHR VON YOUTUBE
- YouTube Premium
  - Filme
  - Gaming
  - Live
  - Sport
  - Einstellungen
  - Meldeverlauf
  - Hilfe
  - Feedback senden

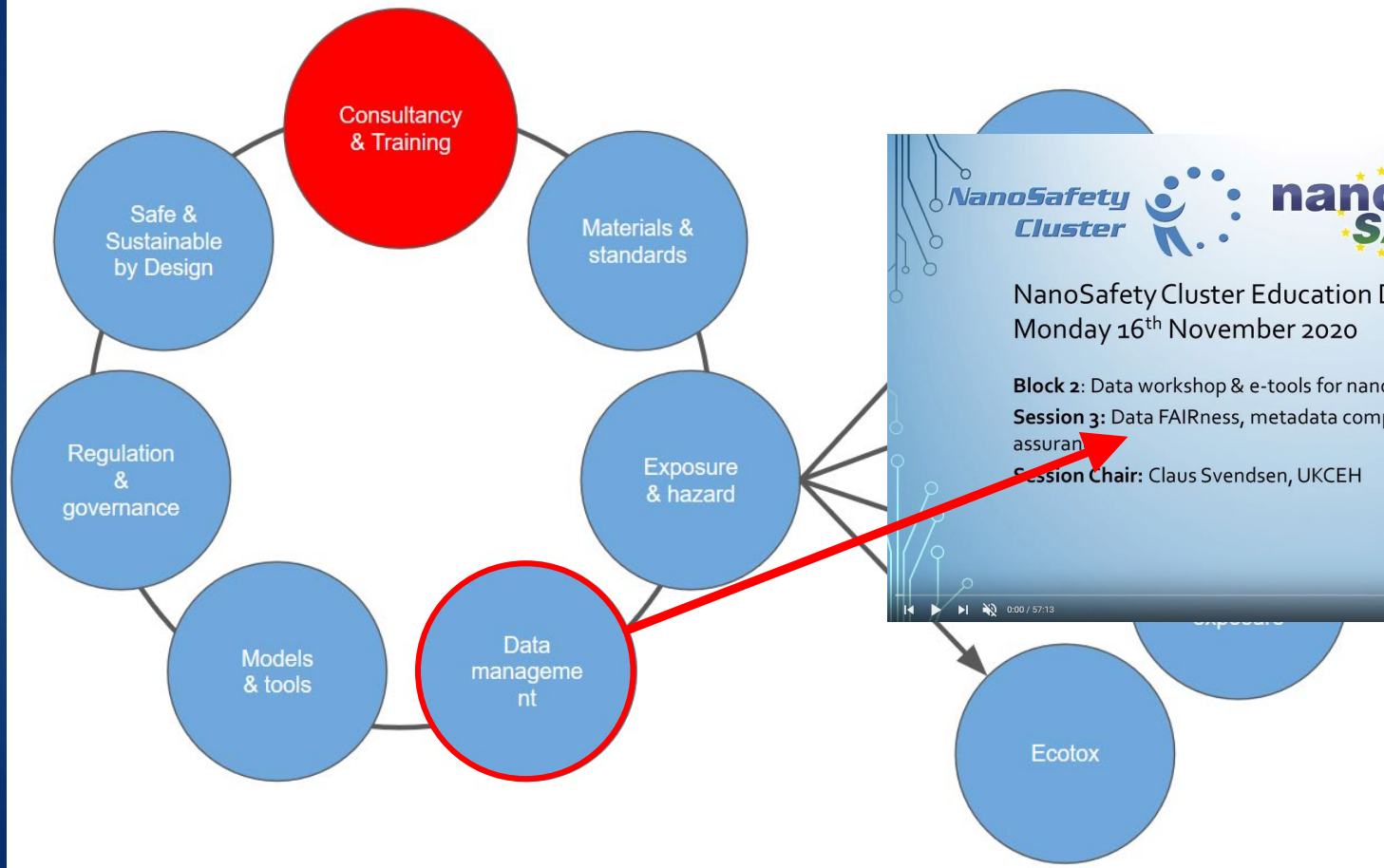
## Uploads


Thumbnail	Video Title	Duration	Views
	NSC Training Day: (Rm B 3) Stakeholder engagement...	2:13:34	17 Aufrufe • vor 1 Monat
	NanoSafety Cluster General Assembly, 16th November...	56:08	13 Aufrufe • vor 1 Monat
	NSC Education Day: Session 6 - Elements of Sbd for...	1:08:38	39 Aufrufe • vor 1 Monat
	NSC Education Day: Session 5 - Features of nanoRGFs in...	1:07:01	15 Aufrufe • vor 1 Monat
	nanoSAFE20 - Introduction to the NanoSafety Cluster...	6:12	11 Aufrufe • vor 1 Monat
	NSC Education Day: Session 1 - New developments in...	57:15	15 Aufrufe • vor 1 Monat

...aligned with the EU NSC WGs



NanoCommons  
Nano-Knowledge Community



**NanoSafety Cluster**  **nanoSAFE'20**

NanoSafety Cluster Education Day  
Monday 16<sup>th</sup> November 2020

**Block 2:** Data workshop & e-tools for nano & beyond  
**Session 3:** Data FAIRness, metadata completeness, scientific quality assurance  
**Session Chair:** Claus Svendsen, UKCEH

0:00 / 57:13



# Data FAIRness

- To **remove barriers** for nanosafety **regulatory and industry** processes
- To develop an **integrated knowledgebase** to facilitate development and application of regulatory tools such as **grouping & read-across**
- To create an **interconnected community** *via* a **FAIR data single market**
- To enable **full exploitation of EU-funded research data** & promotion of **data-driven innovation** leading to **positive socioeconomic impact**



# The NanoCommons Knowledge Base



NanoCommons  
Nano-Knowledge Community



BioXM™ Knowledge Portal

Username:

Password:

Log in



NanoCommons

Nano-Knowledge Community

You don't have an account yet? Please [register here](#).

[Forgot password?](#)

Contact Biomax if you need help [support@biomax.com](mailto:support@biomax.com)

Copyright © 2020 by [Biomax Informatics AG](#). All rights reserved.

# Nanosafety Training School: From Basic Science To Risk Governance

Event Date: 21st June 2021 - 25th June 2021  
Online



## Nanosafety Training School: From Basic Science To Risk Governance

Interprofessional Education Training School

### Organisers



SIGN UP FOR OUR NEWSLETTER

SEND



### TOPICS:

- Hazard to Human Health & Environment
- Fate & Exposure Assessment
- Nanomedicine: from the lab to the market
- Modelling
- Grouping & Read Across Approaches
- Risk Governance



NanoCommons  
Nano-Knowledge Community



# Useful links:

## User Guidance Handbook @ [www.nanocommons.eu](http://www.nanocommons.eu)

- ▶ <https://www.nanocommons.eu/e-infrastructure/user-guidance-handbook/>

## Training events and materials @ NanoCommons Infrastructure

- ▶ <https://infrastructure.nanocommons.eu/events/>

- ↳ NSC Education Day Session 1

- ↳ SmartNanoTox Final Conference

## NanoCommons @ ELIXIR TeSS

- ▶ [https://tess.elixir-europe.org/content\\_providers/nanocommons#events](https://tess.elixir-europe.org/content_providers/nanocommons#events)

## NanoCommons community @ Zenodo

- ▶ <https://zenodo.org/communities/nanocommons>

## NanoCommons Channel @ YouTube

- ▶ <https://www.youtube.com/channel/UCuawpRvXNpglwyeltfTctw>

## mailing list of WG-A Education, Training, Communication

- ▶ [www.nanosafetycluster.eu](http://www.nanosafetycluster.eu)



**NanoCommons**  
Nano-Knowledge Community



Thank you for joining our Webinar!

## **D5.7 First tool for MIE prediction integrated into NanoCommons KnowledgeBase**

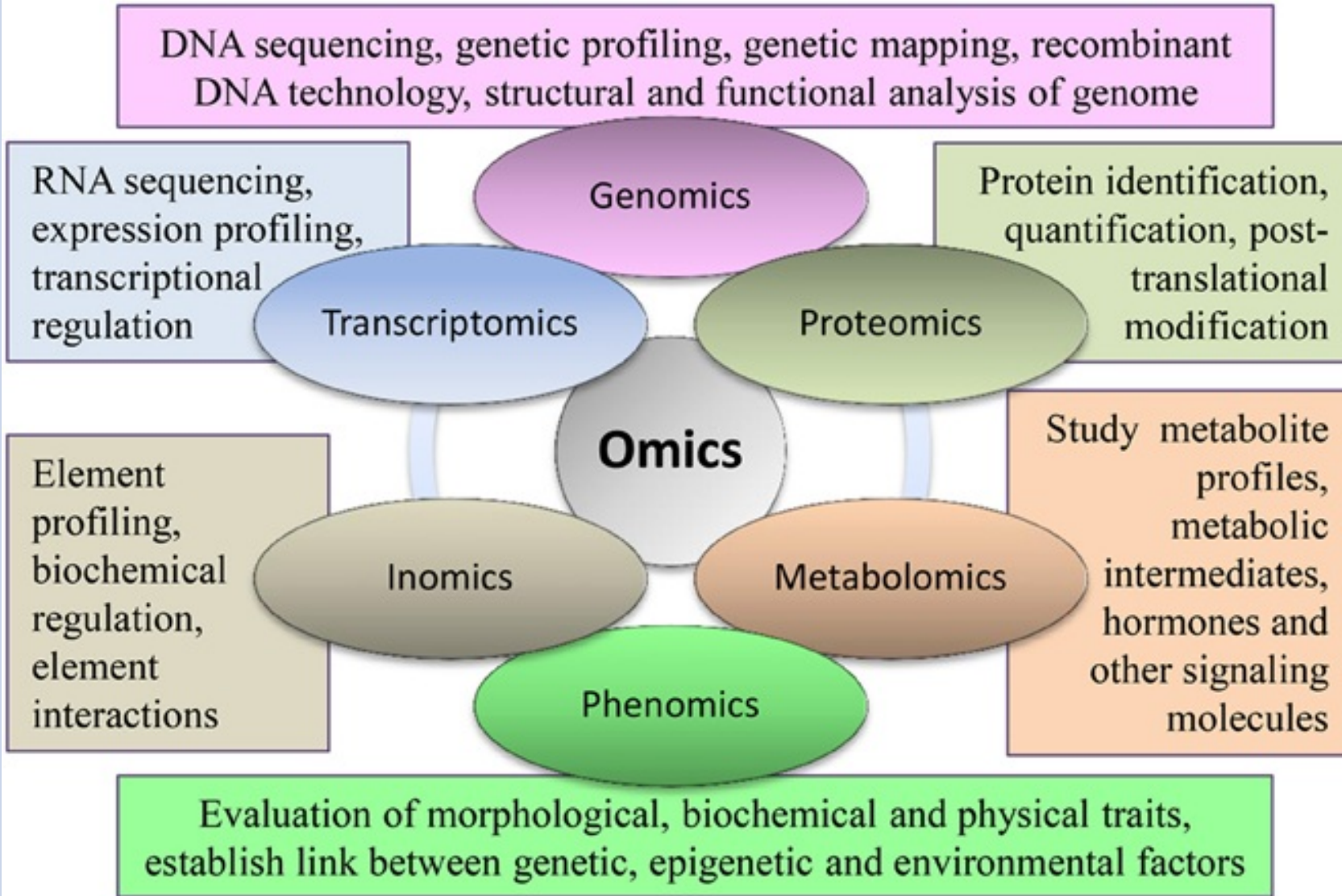
**---Abhijit Dasgupta**





# Some Terminologies

- **Key Events (KE):** Important molecular/chemical interaction initiating another important molecular incidents.
- **Molecular Initiating Events (MIE):** It is the first step in an adverse outcome pathway and can be considered as a chemical interaction between a chemical toxicant and a biological molecule. Key chemical characteristics can be identified and used to model the chemistry of these MIEs.
- **Adverse Outcome (AO):** An unintended and unwanted event or state occurring during or following medical care, that is so harmful to a patient's health that (adjustment of) treatment is required or that permanent damage results.
- **Adverse outcome pathways (AOP):** It is a simplified depiction of complex toxicological processes in a linear and modular format starting with a MIE and ending with an AO.
- **Omics:** It is high-content datasets with measurements of genes, proteins, and/or metabolites. It can provide a better understanding of potential molecular toxicity pathways.



# Omics

# KE, MIE, AO and AOP

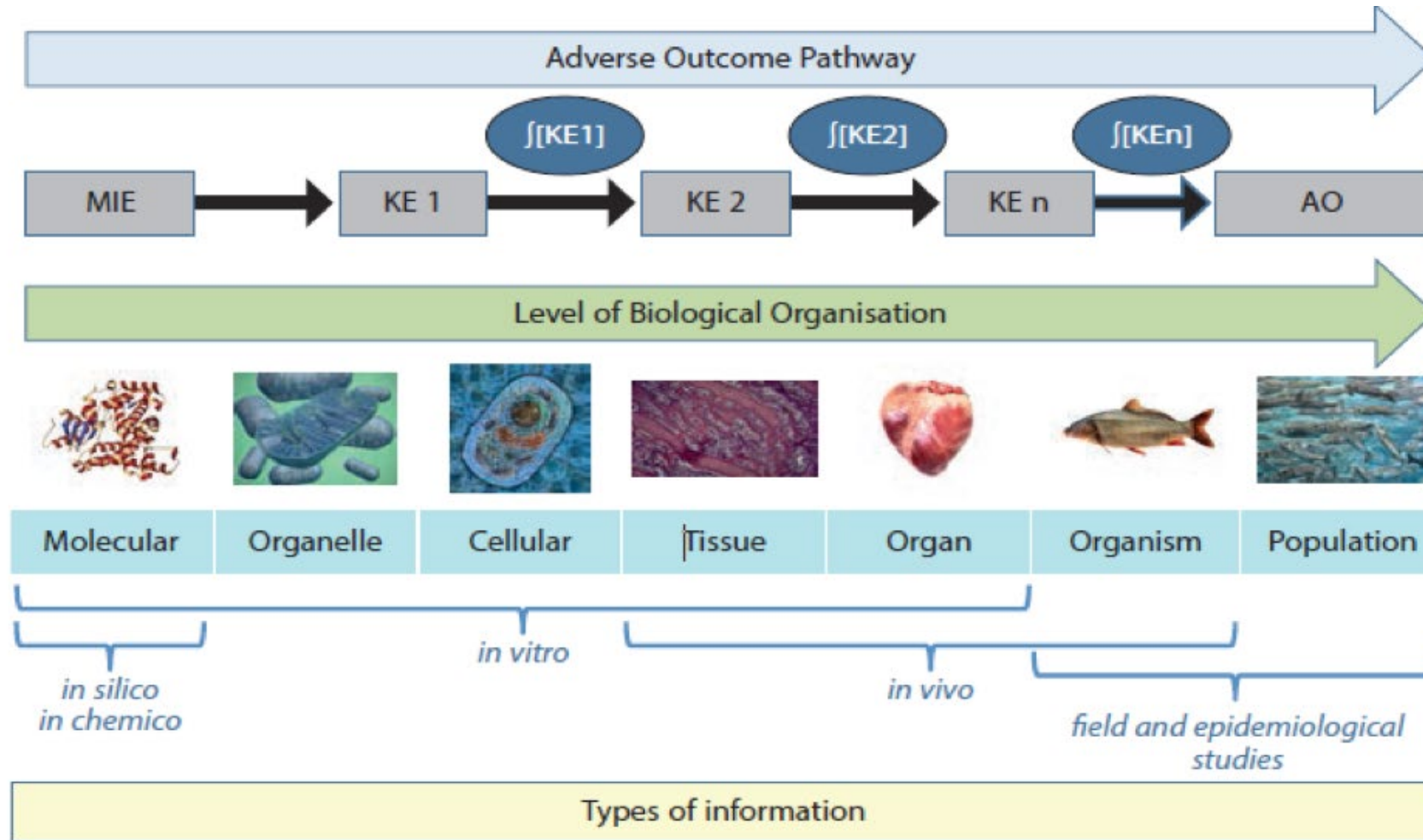
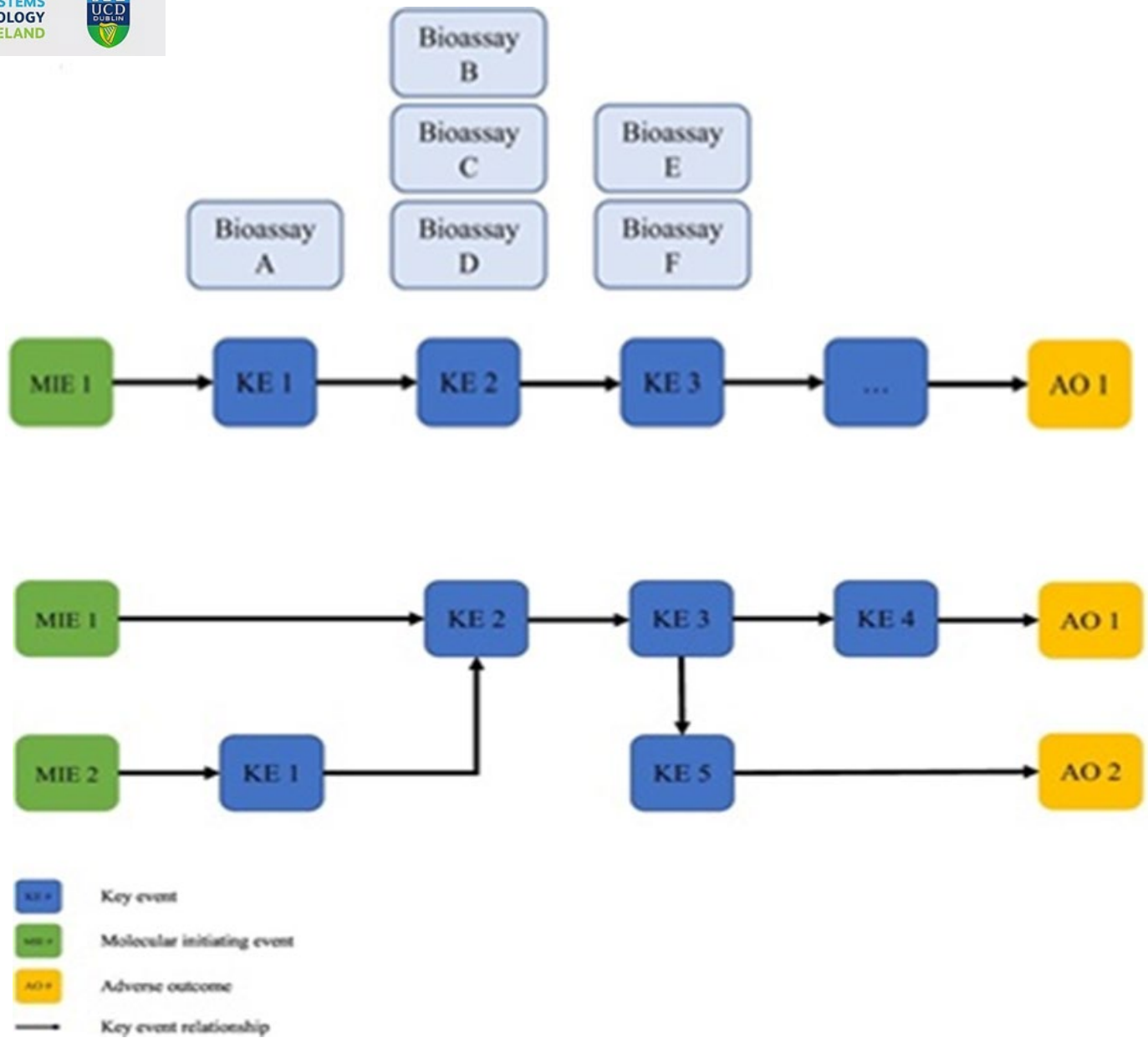


Illustration of an AOP linking molecular and cellular perturbations to impacts at the tissue, organ, organism and population levels, and the types of assays that can be used to explore the perturbations at the different levels. From OECD, 2017.





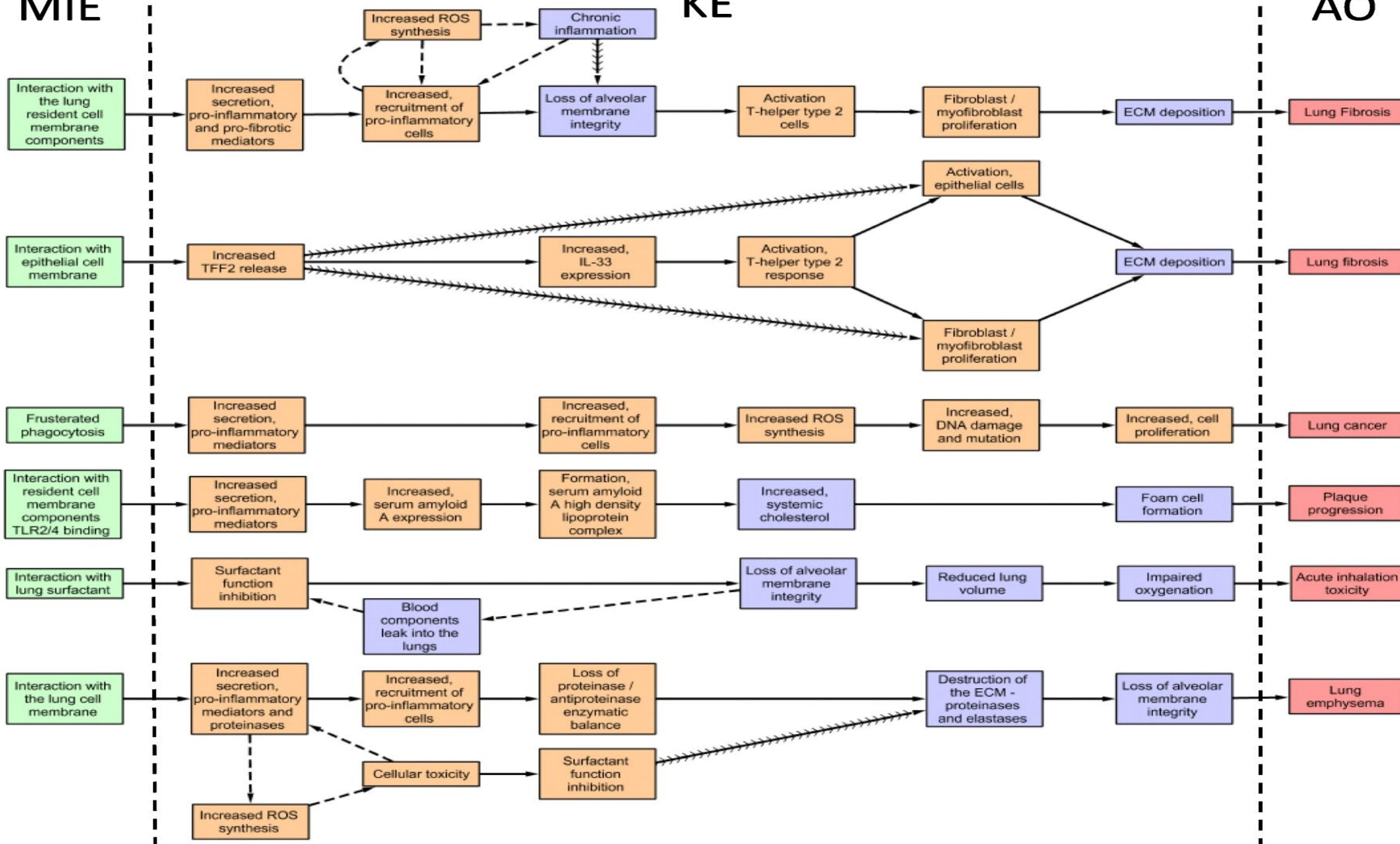
Generalized AOP showing the relationship between MIE, KEs and AOs and the KERs that connect them. Bioassays targeting the MIE and KEs in an AOP developed as part of an IATA (Halappanavar et al. NanolImpact, 2019).

AOPs can form interlinked networks based on overlapping MIEs, KEs and AOs that better capture the complex biology of disease processes (Halappanavar et al. NanolImpact, 2019).

MIE

KE

AO



# Motivation

- **Toxicity testing and regulation of advanced materials at the nanoscale *i.e.*, nano safety, is challenged by the growing number of nanomaterials.**
- **The existing animal-reliant toxicity testing tools are onerous in terms of time and resources.**
- **There is a need for faster, cheaper, sensitive and effective animal alternatives that are supported by mechanistic evidence.**
- **There is an urgency for developing alternative testing strategies.**
- **The Adverse Outcome Pathway (AOP) provides pragmatic insights to promote the development of alternative testing strategies.**
- **The application of omics in AOP-based risk assessment is a long-term goal.**



# MIE: Challenges

- **To understand an MIE completely, a lot of information is required:**
  - ❖ Information about chemicals that are associated with the MIE.
  - ❖ Structural features or properties of the chemical that causes its association.
  - ❖ The types of interactions that occur between the chemical and biomolecule or biosystem.
  - ❖ The nature or structure of the entity with which the molecule interacts.
- **Obtaining all aforementioned information is very difficult.**
- **The most well studied chemicals lack detailed reports of molecular interactions.**
- **Nano materials (NMs) is less known.**
- **Risk assessment requires information on the exposure conditions (e.g., route, dose, duration and frequency) needed to cause an AO.**
- **Predicting actual MIEs without time-resolved data establishing the MIE is challenging.**

- Ede et al. Nanomaterials (Basel), 2020
- Allen et al. Chem Res Toxicol, 2014
- Vinken et al. Curr Opin Toxicol, 2019
- Christmann et al. Nucleic Acids Res, 2013

## MIE: Probable Solution

- **Partial information from different sources must be brought together when evaluating MIEs.**
- **Quantitative AOPs (qAOPs), which use quantitative data, may help to predict risk of an AO under specific exposure conditions.**
- **Omics-based data can be used to feed virtually all information blocks in AOPs.**
- **Availability of data for the initial response period (1-3 hours after NM exposition), can help most likely to predict actual MIEs.**

# Brief Description of the MIE Tool

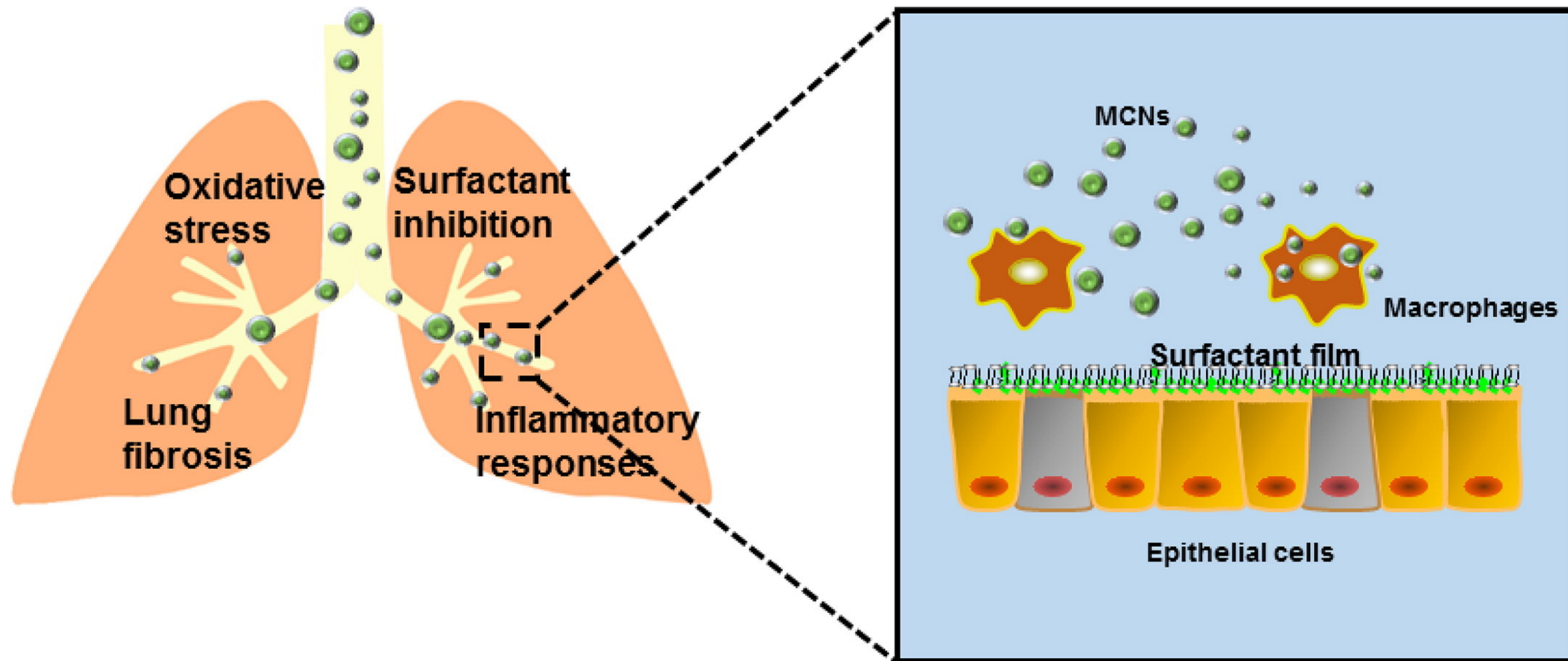
## NanoCommons MIE gene set database (NanoCommons GS-MIE DB):

It captures-

- I. Gene signatures (GS) of MIEs by integrating knowledge from KEGG, REACTOME, GO, WikiPathways public databases.
- II. Custom gene sets from published data.
- III. To date, manual collection of 132 gene sets representing three different types of MIE actions:
  - MIE1. Disruption of lung surfactant functionality.**
  - MIE2. Lysosomal destabilization.**
  - MIE3. Oxidation of cell membrane.**



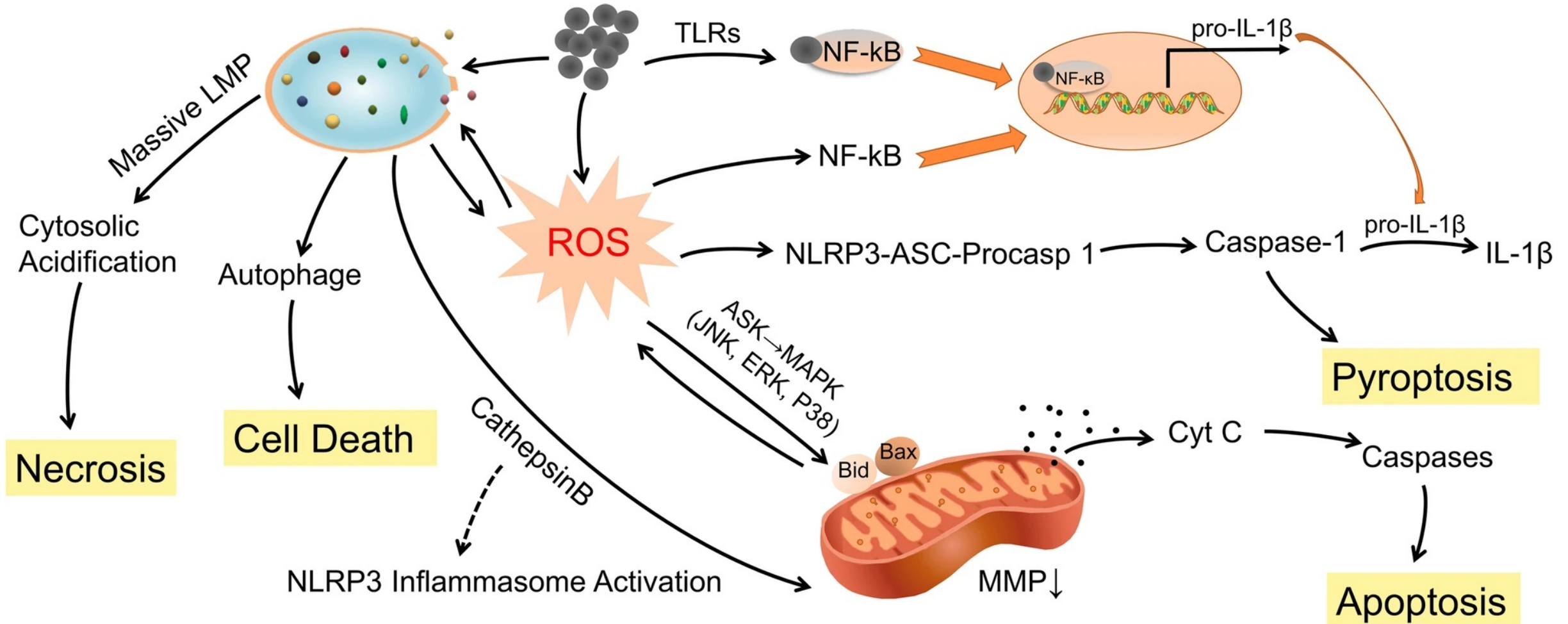
# MIE1: Disruption of lung surfactant functionality



The comprehensive toxicological profile of mesoporous carbon nanomaterials (MCNs) under the scenario of moderate environmental exposure.

Chen, Yunan, et al. "Mesoporous carbon nanomaterials induced pulmonary surfactant inhibition, cytotoxicity, inflammation and lung fibrosis." *Journal of Environmental Sciences* 62 (2017): 100-114.

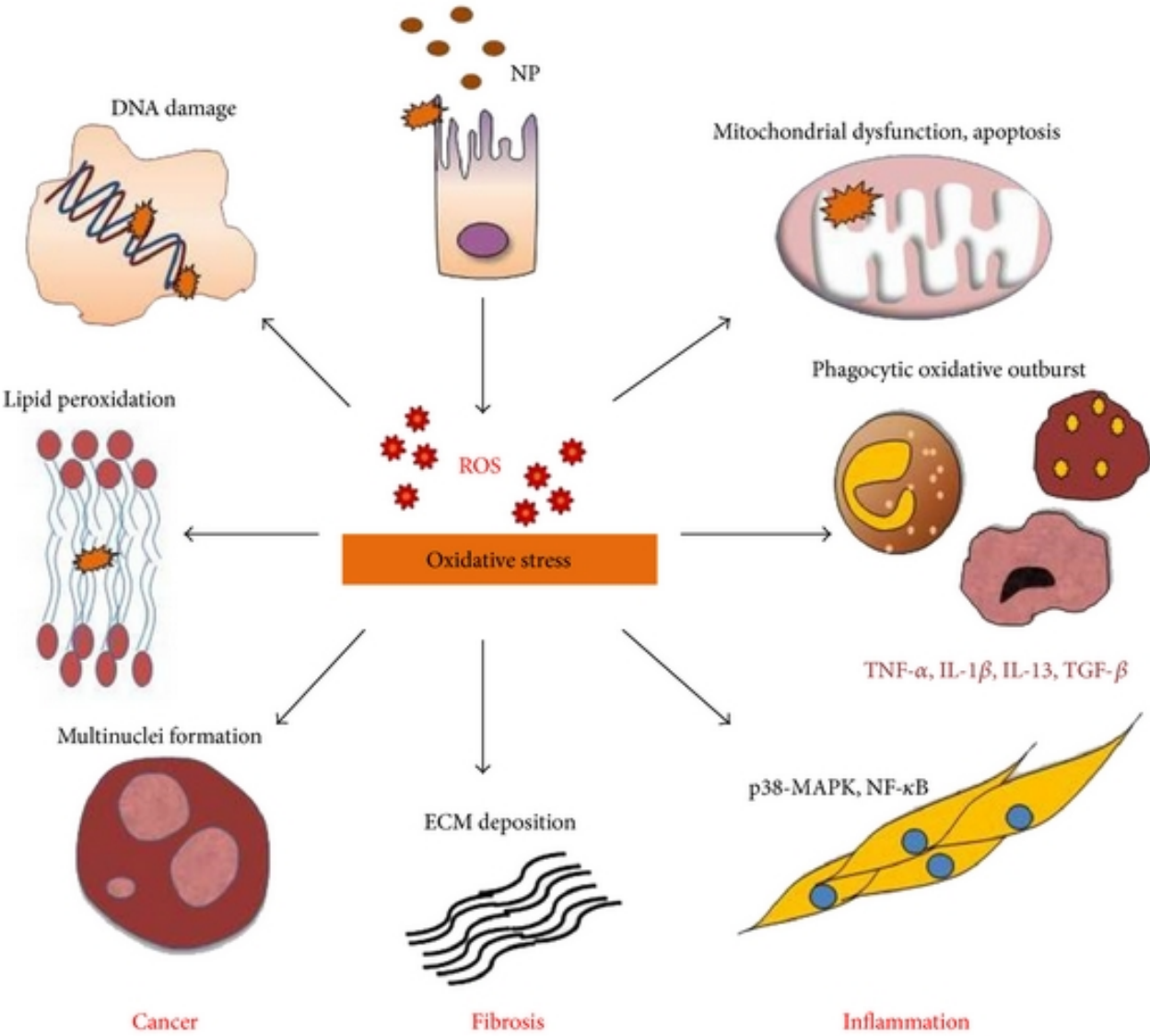
# MIE2: Lysosomal destabilization



The mechanisms by which carbon-based nanoparticles induce cytotoxicity of macrophages.

Yuan, Xia, et al. "Cellular toxicity and immunological effects of carbon-based nanomaterials." *Particle and fibre toxicology* 16.1 (2019): 1-27.

# MIE3: Oxidation of cell membrane



Prooxidant pathway for NP-induced toxicity: various NP exhibit oxidative stress dependent toxicity.

Manke, Amruta, Liying Wang, and Yon Rojanasakul. "Mechanisms of nanoparticle-induced oxidative stress and toxicity." *BioMed research international* 2013 (2013).



## Brief Description of the MIE Tool (Continued...)

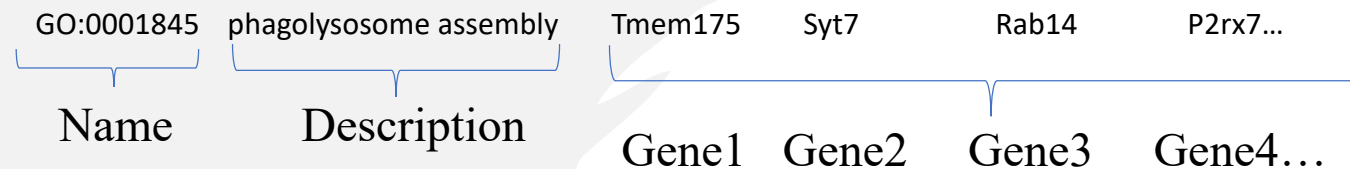
### A web interface & R modules:

- ❖ It integrates the MIE prediction tool into the NanoCommons Knowledgebase.
- ❖ It provides the R codes for access to the database and calculation of the overrepresentation of distinct biological processes for each MIE.
- ❖ The tool uses a list of differentially expressed genes/proteins as input from high-throughput experiments.
- ❖ It enables to calculate a prioritized list of MIEs with identified biological processes.
- ❖ The services are available for Transnational Access (TA).

# NanoCommons GS-MIE DB

- We create databases responsible for three MIEs by creating the MIE gene sets using information from the KEGG, REACTOME, GO, HP and WikiPathways, public databases as well as published data.
- The database was constructed based on the Gene Matrix Transposed (GMT) file format

## GMT File Format



# NanoCommons GS-MIE DB

MIE	Description
Disruption of lung surfactant functionality	<p>Three gene sets:</p> <ul style="list-style-type: none"> <li>• GO:0043129 Surfactant homeostasis</li> <li>• REAC:R-MMU-5683826 Surfactant metabolism</li> <li>• SBI:S001 Surfactant homeostasis custom gene set</li> </ul>
Lysosomal destabilization	<p>Number of gene sets: 54</p> <p>Examples of the gene sets:</p> <ul style="list-style-type: none"> <li>• GO:0097212 Lysosomal membrane organization</li> <li>• REAC:R-MMU-432720 Lysosome Vesicle Biogenesis</li> <li>• HP:0004356 Abnormality of lysosomal metabolism</li> <li>• KEGG:04142 Lysosome</li> </ul>
Oxidation of cell membrane	<p>Number of gene sets: 75</p> <p>Examples of the gene sets:</p> <ul style="list-style-type: none"> <li>• GO:0006979 Response to oxidative stress</li> <li>• GO:0071451 Cellular response to superoxide</li> <li>• REAC:R-MMU-1222556 ROS and RNS production in phagocytes</li> <li>• WP:WP1496 Oxidative Damage</li> </ul>

# The web interface

## MIE prediction

Threshold:

Enter your list of genes (one gene per line):

Il6  
Cxcl5  
Saa1  
Mt2  
Gm1960  
Cxcl1  
Timp1  
Pbp2  
Aldh1a3  
Hnxn

[\[example 1\]](#) [\[example 2\]](#)

MIE prediction tool performs functional enrichment analysis using custom gene sets:

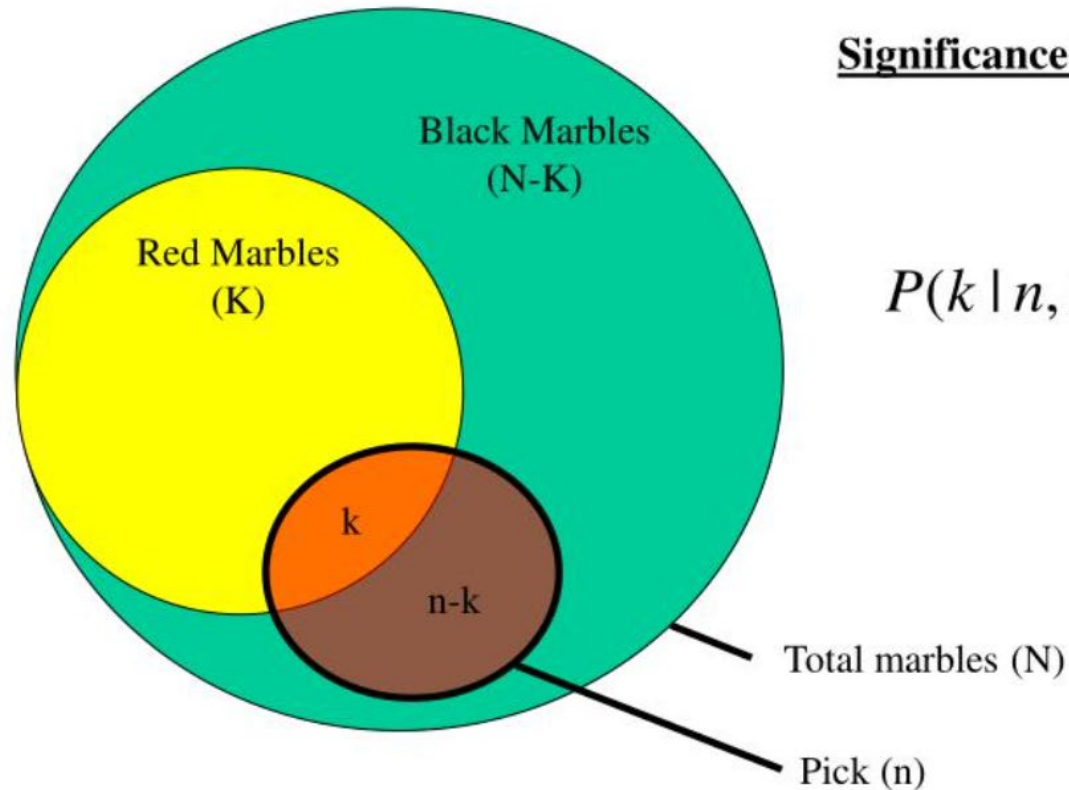
Show  entries Search:

	Event	term_id	name	p_value	genes
1	MIE2	KEGG:04142	Lysosome	0.00584125972250547	Litaf, Cd63, Hexa, Ctsk, Cltb, Ctsw, Ppt1, Manba, Gnptab, Slc11a1, Abca2, Ppt2, Gaa, Ap4e1, Ap3m2, Aga, Nagpa, Gga2, Atp6v0a1, Arsb, Ctsf
2	MIE3	GO:0006801	superoxide metabolic process	0.000222152617578984	Cxcl1, Noxo1, Sod2, Agt, Fbln5, Gch1, Tnf, Duox1
3	MIE3	GO:0090322	regulation of superoxide metabolic process	0.00299823855002102	Cxcl1, Agt, Fbln5, Gch1, Tnf
4	MIE3	GO:0006979	response to oxidative stress	0.00623995061061375	Il6, Zc3h12a, Sod2, Ptgs1, Dusp1, Slc7a11, Arg1, Fbln5, Areg, Sphk1, Lcn2, Plk3, Gch1, Tnf, Il18rap, Duox1, Epor, Fos, Arntl, Pdgfra, Hif1a, Bmp4, Fkbp1b, Akr1b3, Il18bp, Pla2r1, Mpv17, Axl, Tat, Ccl19, Ercc6, Glrx, Mmp2, Pxdn, Mmp14, Txnrd1, Plekha1, Rgs14, Ptk2b, Btk, Jak2, Mcl1, Chuk



# Over-Representation Analysis

If you draw  $n$  marbles at random, what is probability of  $k$  red ones?



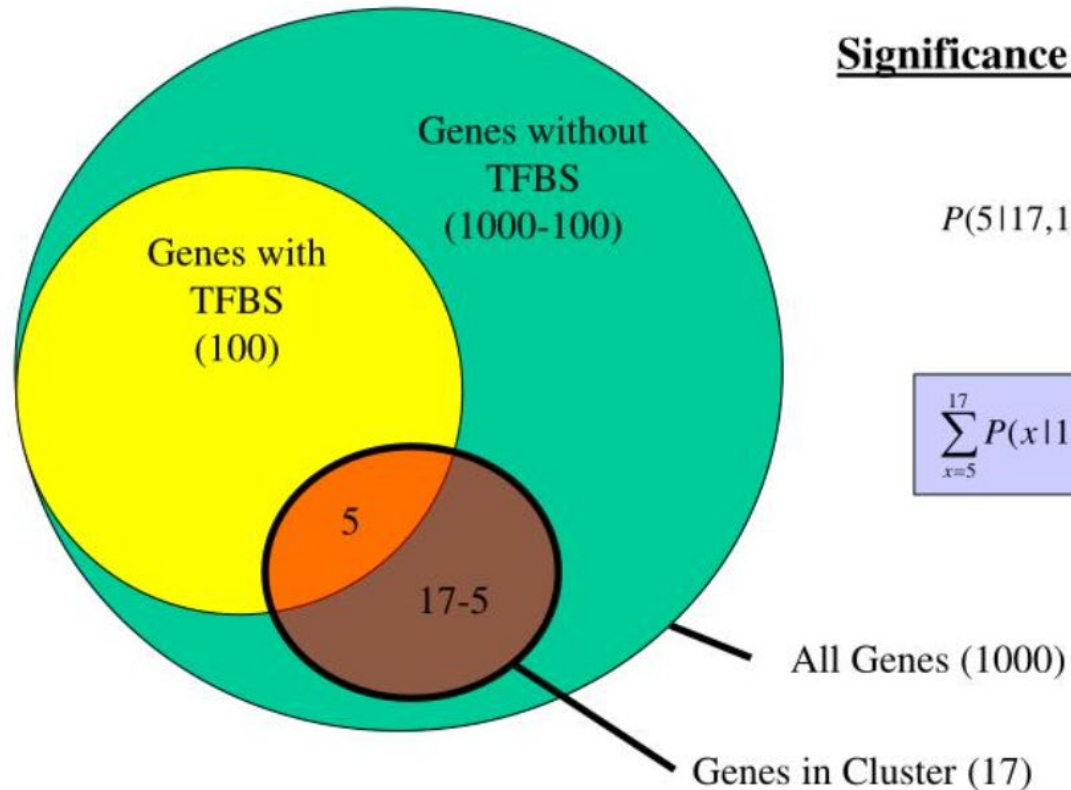
**Significance by Hypergeometric Distribution**

$$P(k | n, K, N) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

**Hypergeometric Distribution**

# Over-Representation Analysis

Assume 17 genes in cluster, 5 with binding site...



## Significance by Hypergeometric Distribution

$$P(5|17,100,1000) = \frac{\binom{100}{5} \binom{1000-100}{17-5}}{\binom{1000}{17}} = 0.017$$

$$\sum_{x=5}^{17} P(x|17,100,1000) = 0.02$$

Must choose threshold to define “differential expression”

# Hypergeometric Test: Summary

- It is a probability distribution that describes the probability of genes of interest ( $x$ ) being associated with a particular MIE, for all genes in the gene list ( $n$ ), from a population of all the genes ( $N$ ) in entire genome which contains  $k$  number of genes associated with the MIE.
- The calculation of  $x$  success follows the formula given below.

$$P(X = x) = h(x; N, n, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

- This test will result in an adjusted p value (after multiple test correction) for each category tested.

# R modules

- It can be utilized for integration into various automated pipelines, including the Jaqpot computational platform and the NanoCommons Knowledgebase.
- R library is a means to provide universal access to the tool and can be implemented into various automated pipelines. The `gost` function from `gprofiler2` library will calculate a prioritized list of MIEs with identified biological processes.

```
#####
```

```
#gprofiler2 can be installed from CRAN:
```

```
install.packages("gprofiler2")
```

```
#or via conda from the conda-forge channel:
```

```
conda install -c conda-forge r-gprofiler2
```

```
#running the code:
```

```
library(gprofiler2)
```

```
#Performing Over-Representation Analysis
```

```
gost(query, organism = 'gp__2ruD_Ave9_Hk4', domain_scope = "known", correction_method = "gSCS", user_threshold = p_value, evcodes = TRUE, ordered_query = TRUE)$result
```

```
#####
```

## Arguments of `gost` function:

- ❑ **query** - List of differentially expressed genes/proteins. Different types of identifiers can be used: gene symbols, Entrez Gene IDs, transcripts, microarray IDs, uniprot protein ID, *etc.*
- ❑ **p\_value** - User-defined p-value threshold provides a possibility to additionally filter results. The threshold defaults to  $p=0.05$ , meaning that all significant results are shown. If the threshold is set to less than 0.05, matches with p-values above threshold are not shown.
- ❑ **organism** = 'gp\_\_2ruD\_Ave9\_Hk4' - predefined token for MIE gene set database.
- ❑ **Correction\_method** - The tool uses `g:SCS` multiple testing correction algorithm by default. Alternatively, False Discovery Rate (FDR), Bonferroni correction (BC) and Benjamini-Hochberg FDR (False Discovery Rate) can be used.



# MIE Tool: Pipeline

To create the MIE gene set database: we used published data from *in vivo* and *in vitro* experimental studies of nanoparticles toxicity and different signaling and functional databases.

We used R and Python client libraries for g:Profiler tool to perform an overrepresentation analysis for MIE terms using the cumulative hypergeometric test.

The user needs to input a list of genes/proteins and press the “Run query” button.

The genes can be derived from omics experimental data, such as RNA-seq, microarrays, mass-spectrometry datasets, among others.

The tool performs incremental enrichment analysis.

This list of genes is interpreted as an ordered list where elements are in the order of decreasing importance.

This order of genes can represent some biological effects, for instance, fold changes of differential expression genes (DEGs), absolute expression values, among others.

The web interface accepts a list of genes (one gene per line) in different formats: gene symbols, gene IDs, microarray IDs, transcripts, uniprot protein ID, among others. Mixed types also can be entered.

# Conclusions

- The tool can fill the gap in understanding of bionano interactions and facilitate reconstruction of the sequential chain of KEs in a particular AOP.
- In this first version of the tool, we integrated gene expression predictors for three key types of MIEs: disruption of lung surfactant functionality, lysosomal destabilization, and oxidation of cell membrane.
- These MIEs were highlighted and described as the initial starting point of toxicological response to NMs for pulmonary AOPs in the SmartNanoTox project.
- The tool will be extended and updated during the NanoCommons project to address the needs of the nanosafety community, with this deliverable representing the first such tool to be integrated into the NanoCommons Knowledgebase.

# *Thank You*

