EARLY WARNING TOOL DEVELOPED FOR THE SWEDISH TOXICOLOGICAL COUNCIL

Suzanne Bruks, Patrik Andersson

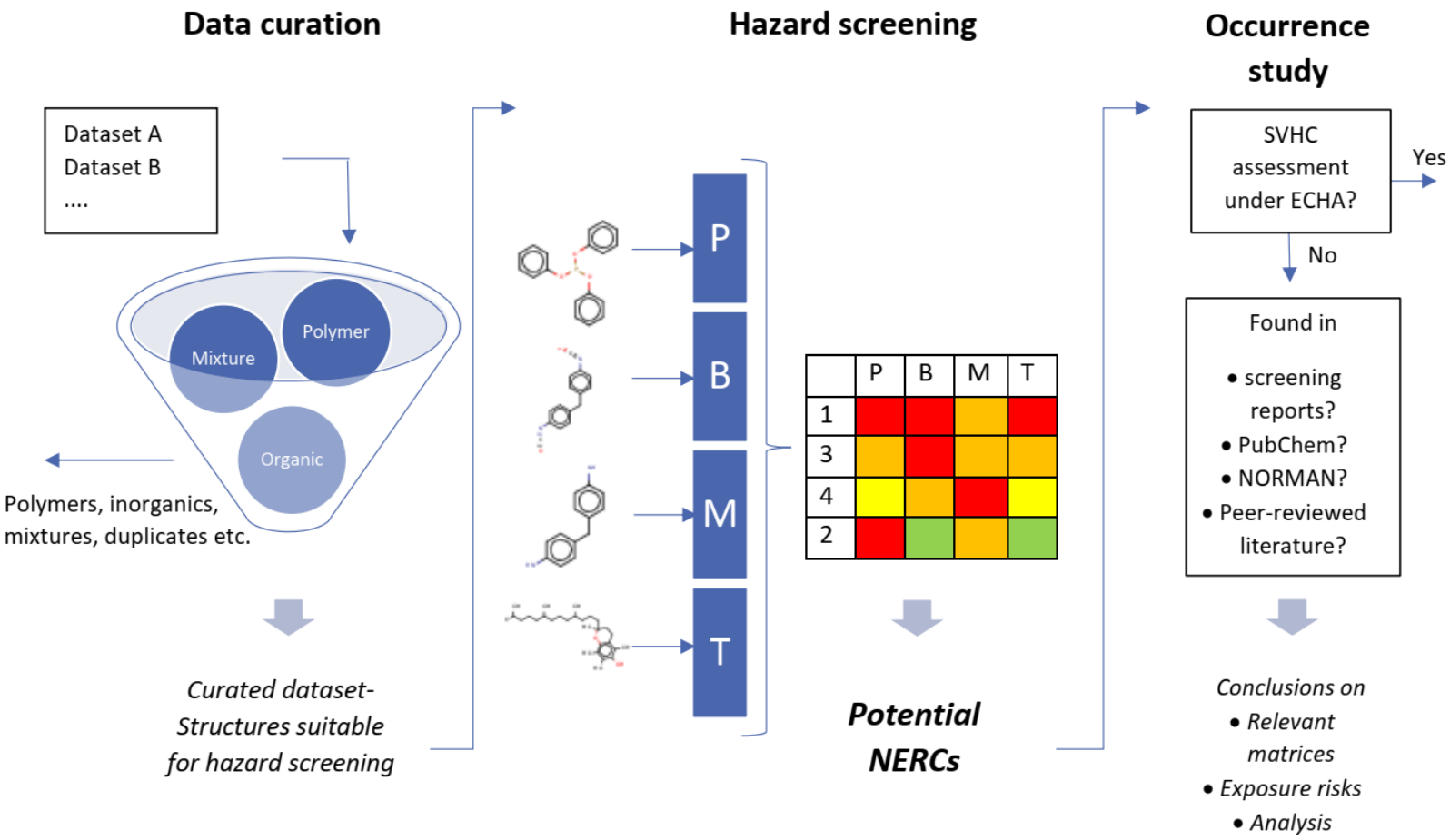
Chemistry Department

UMEÅ UNIVERSITET

AIMS OF THE PROJECT

- Initiate development of a systematic tool for identification of potential New or Emerging Risk Chemicals (NERCs)
- Development of data curation strategy
- Generating lists of NERCs
- (Comparison with literature data – Karin Wiberg, Vera Franke SLU)

OVERVIEW OF PROJECT

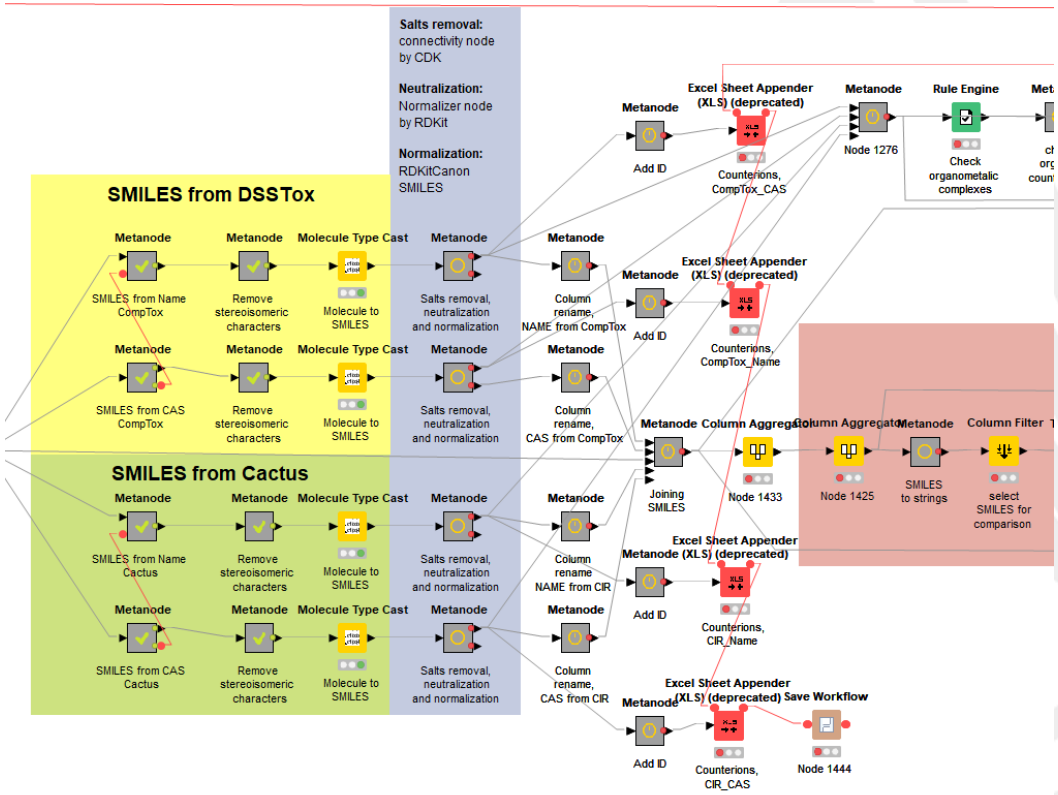


DATA CURATION

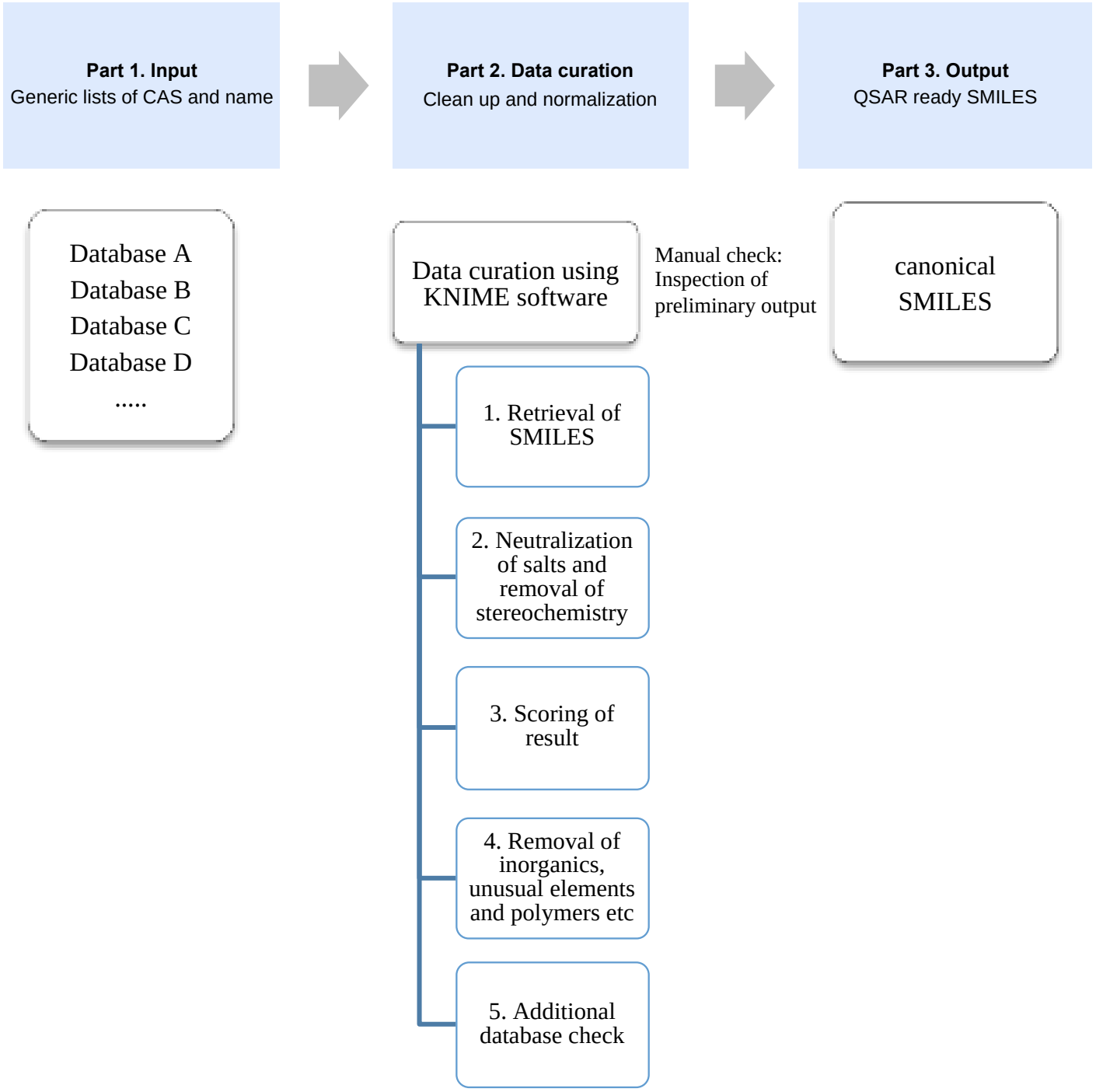
- **What is this?**
  - Data wash and filtrering
  - Standardization of molecular structures
- **Why is it needed?**
  - Ensure data quality
  - Prepare data for in silico modelling (QSAR)
- **Pros and conc**
  - Models suited for small organic molecules
  - Errorneuos data in inventories
  - Dependent on correct, curated data

KNIME ANALYTICS

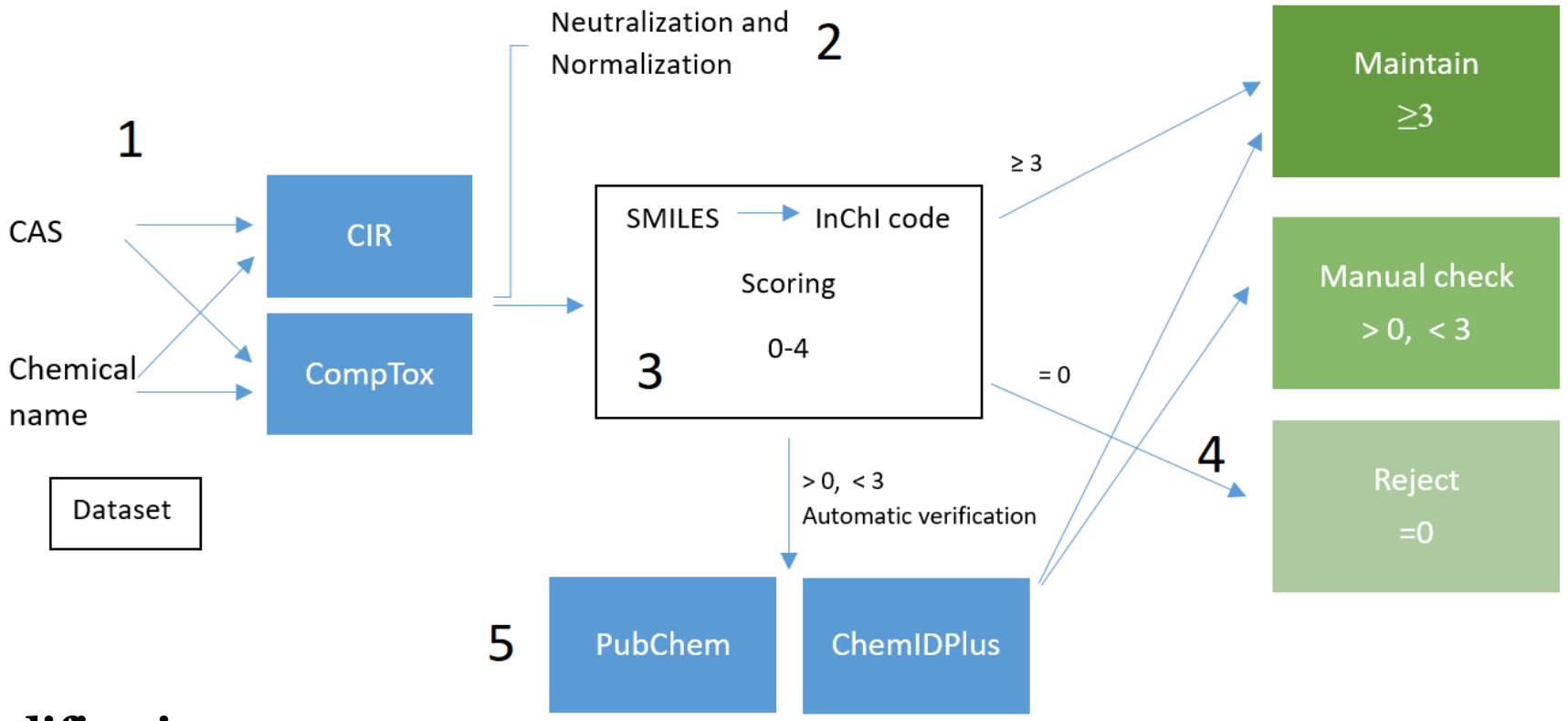
- Open source platform
- No expertise in programing needed
- Tailor made, flexible
- Simple, transparent
- Suitable for data mining
- Gadaleta et al 2018



UMEÅ UNIVERSITET


Part 1. Input
Generic lists of CAS and name
Part 2. Data curation
Clean up and normalization
Part 3. Output
QSAR ready SMILES
Database A
Database B
Database C
Database D
.....
Data curation using KNIME software
Manual check: Inspection of preliminary output
canonical SMILES
1. Retrieval of SMILES
2. Neutralization of salts and removal of stereochemistry
3. Scoring of result
4. Removal of inorganics, unusual elements and polymers etc
5. Additional database check

FILTRATION OF ENTRIES



Modifications;
SciFinder
OpenMolecule
ChemSpider
Filter

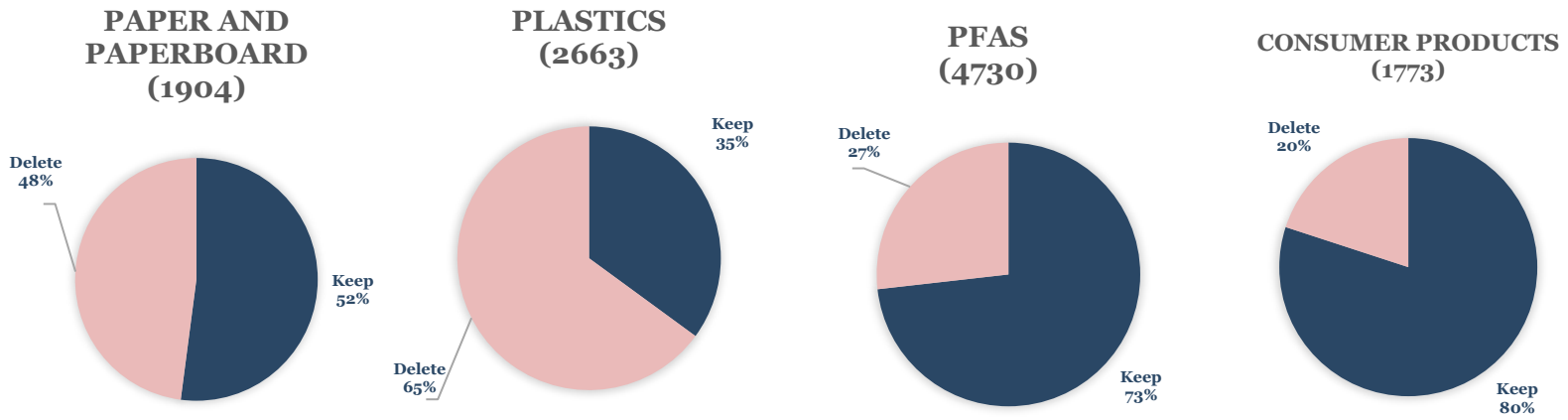
CASE STUDIES/DATA SETS

- Paper and paper board (1904 entries)
- Plastic additives (2663 entries)

} Swedish Chemicals Agency

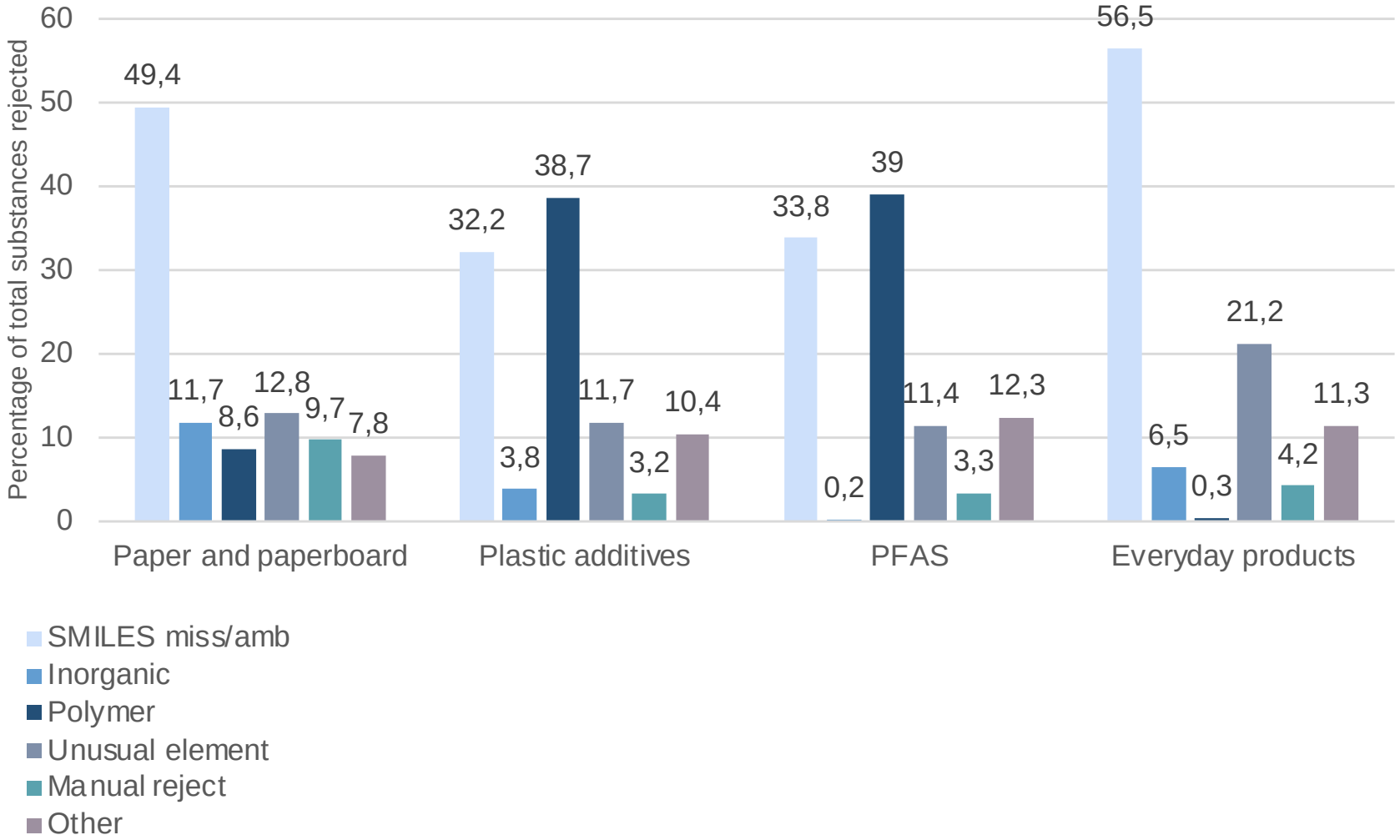
- OECD inventory of PFAS (4730 entries) (Chelcea et al 2020)
- Danish EPA (Miljøstyrelsens) inventory of chemicals in consumer products (appr 1800 entries)

RESULTS FROM DATA CURATION

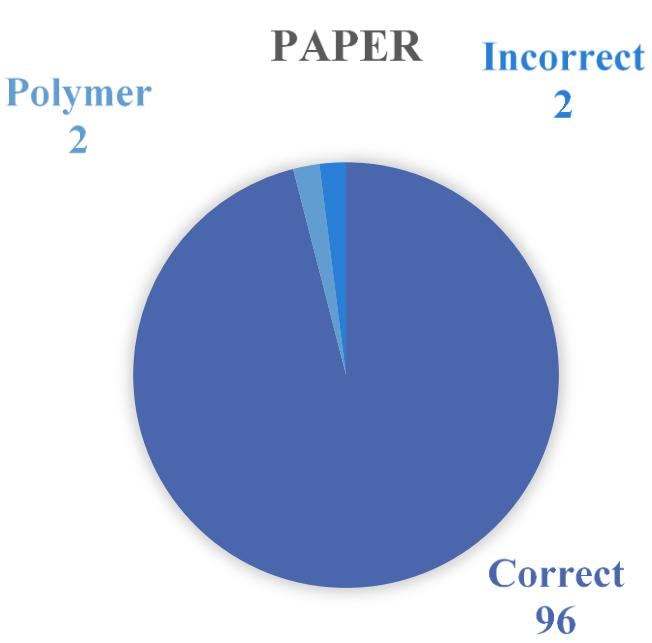


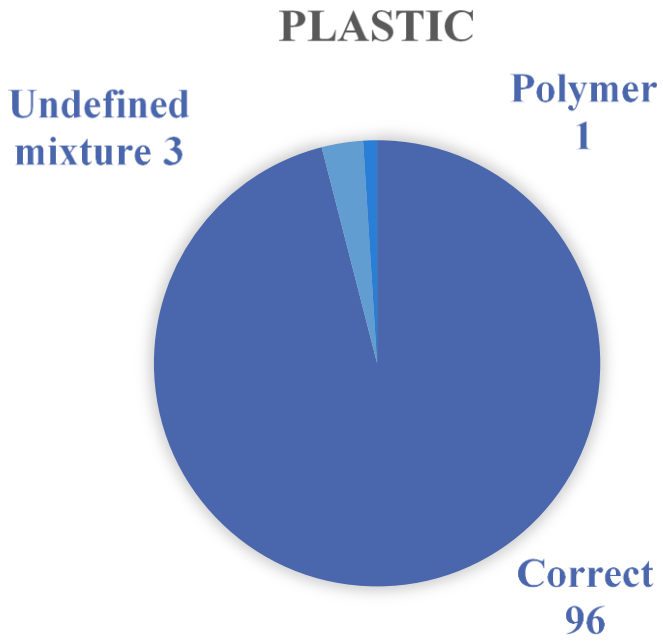
- Workflow 4-6h (PFAS 28h); manual check 4-8h
- Large share is rejected (20-65%)
- Variation between databases
- Plastic additives include many polymers
- Consumer products data base is in large based on target screening data

WHAT IS REJECTED?



HOW CORRECT IS CORRECT?





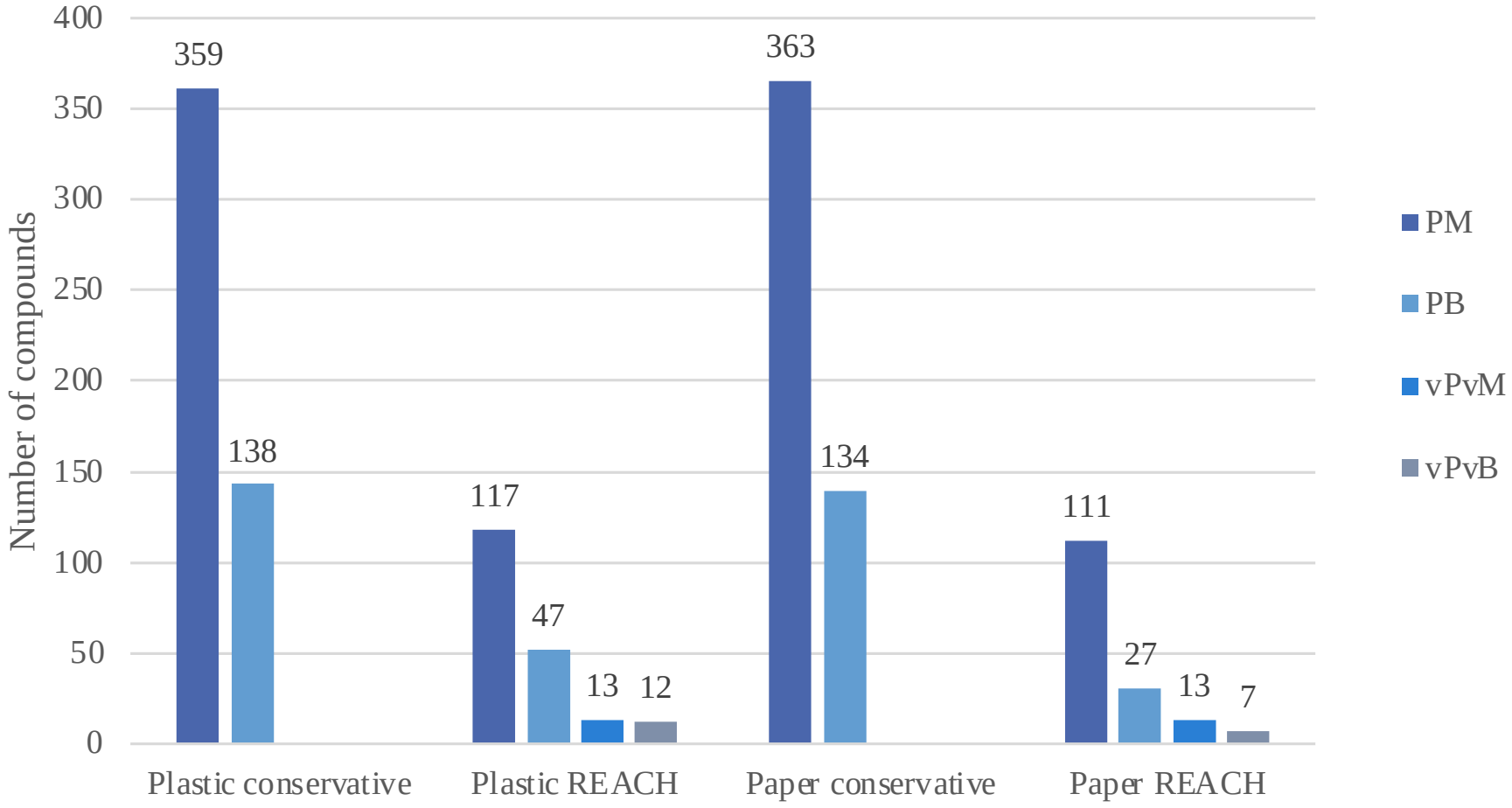
HAZARD SCORING - MODELS

- Bioaccumulation potential
  - BCF; three models – Episuite, Ceasar, Arnot-Gobas
- Persistence
  - BIOWIN; ULTIMATE, VEGA; nP; nP/P; P/vP; vP
- Mobility
  - LOGKOC; Episuite (MCI, Kow), Opera (Vega)
  - Water solubility (Epi, Vega)
- Toxic potential – to be analyzed
- How filter? Open for user.

POSITIVE CONTROLS

Abbreviation	Chemical name	CAS number	Hazard labelling
HBCDD	Cyclododecane, 1,2,5,6,9,10-hexabromo-	3194-55-6	PBT, POP, R, Ss
Triclosan	Phenol, 5-chloro-2-(2,4-dichlorophenoxy)-	3380-34-5	PBT, ED
DDT	Benzene, 1,1'-(2,2,2-trichloroethylidene)bis[4-chloro-	50-29-3	POP, C
TBBPA	Phenol, 4,4'-(1-methylethylidene)bis[2,6-dibromo-	79-94-7	PBT, ED
BPA	Phenol, 4,4'-(1-methylethylidene)bis-	80-05-7	ED, Ss, R
TPhP	Phosphoric acid, triphenyl ester	115-86-6	ED
Benzo[a]pyrene	Benzo[a]pyrene	50-32-8	PBT, POP, R, M, C, Ss
DEHP	1,2-Benzenedicarboxylic acid, 1,2-bis(2-ethylhexyl) ester	117-81-7	ED, R
TCEP	Ethanol, 2-chloro-, 1,1',1''-phosphate	115-96-8	R, C
PFOS	1-Octanesulfonic acid, 1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-heptadecafluoro-(	1763-23-1	POP, R, C

RANKING OF CHEMICALS



SUMMARY

- Efficient semiautomatic data curation workflow was developed using KNIME
- Identified need for further improvements, washing procedures, identification of unwanted structures
- A large share of entries can not be handled by todays hazard scoring models, e.g. polymers, oligomers, rare elements, inorganics, and PFAS
- Models in large missing for generating degradation products
- Next level would be to implement models in the KNIME workflow (VEGA)
- One important task in PARC, WP8 as a component of an Early Warning System
- Report available at kemi.se
- Danish database under study this autumn