DATA ANALYTICAL APPROACH FOR THE IDENTIFICATION OF EMERGING FOOD SAFETY RISKS

ZSUZSA FARKAS, ÁKOS JÓZWIAK

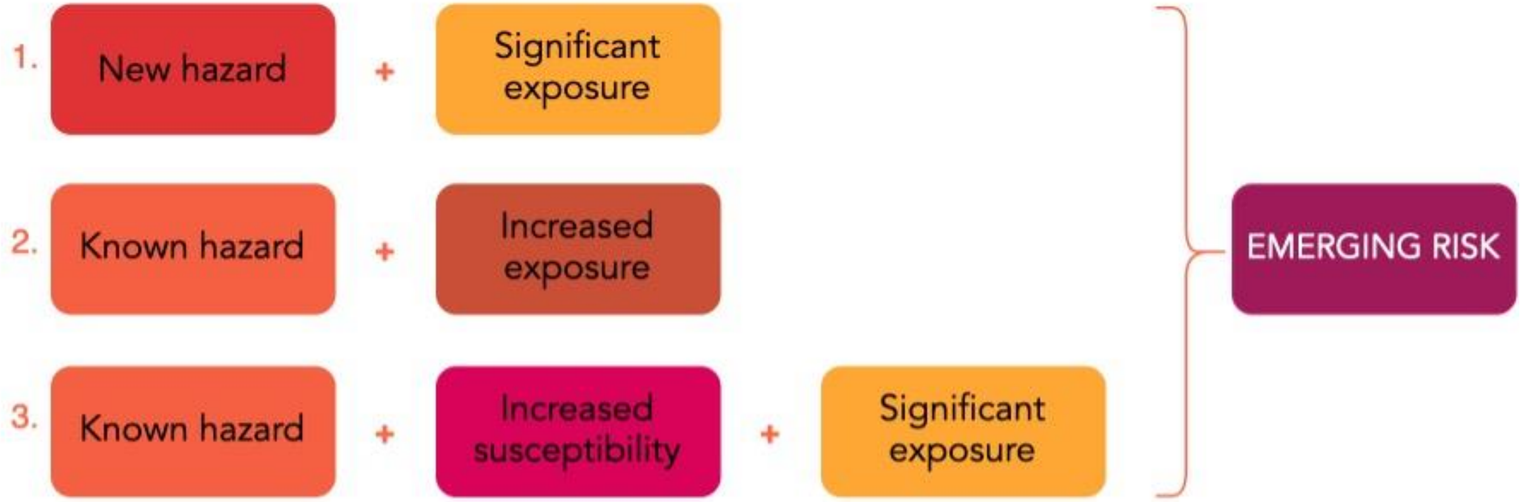
DIGITAL FOOD INSTITUTE
UNIVERSITY OF VETERINARY MEDICINE, BUDAPEST, HUNGARY

MODERN INSPECTIONS WITH NEW TECHNOLOGY SYMPOSIUM
26-27 FEBRUARY, 2024. THE HAGUE, THE NETHERLANDS

INTRODUCTION

EMERGING RISK DEFINITION

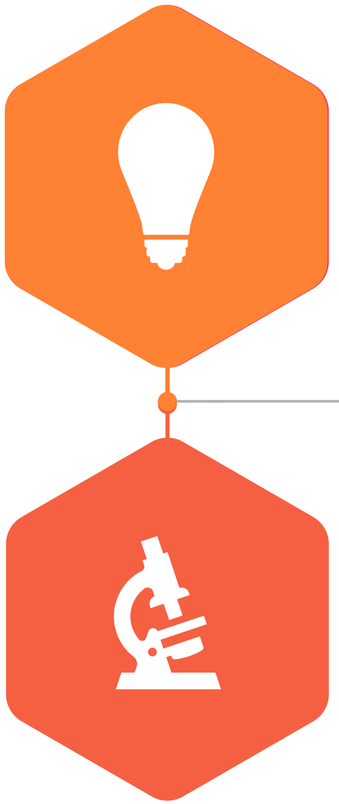
- ‘An emerging risk to human, animal and/or plant health is understood as a risk resulting from a newly identified hazard to which a significant exposure may occur or from an unexpected new or increased significant exposure and/or susceptibility to a known hazard’ (EFSA)



- **RISK BASED APPROACH**: Timely identification of food systems risks needs a profound knowledge on the prevalence and severity of risks
- Continuous evaluation of risks and continuous knowledge genaration → complex process

HORIZON SCANNING UNIVERSE

DEFINITIONS IN THE HORIZON SCANNING UNIVERSE



EARLY WARNING

SHORT TERM

Rapid alert systems

Immediate action required

Ongoing outbreaks/incidents somewhere else

RASFF
Rapid Alert System for Food and Feed

INFOSAN

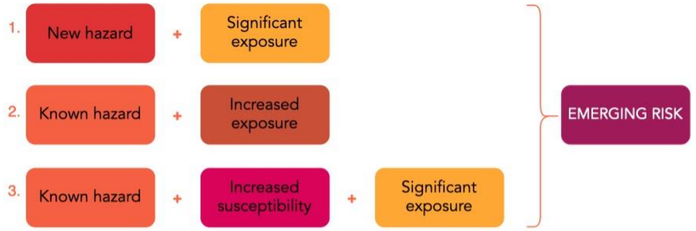
EMERGING RISKS

MEDIUM TERM

Screening systems

Increases preparedness

Initiates risk assessment



FORESIGHT

LONG TERM

Driver and scenario analysis

Affects strategic actions

POPULATION GROWTH

GLOBAL TRADE

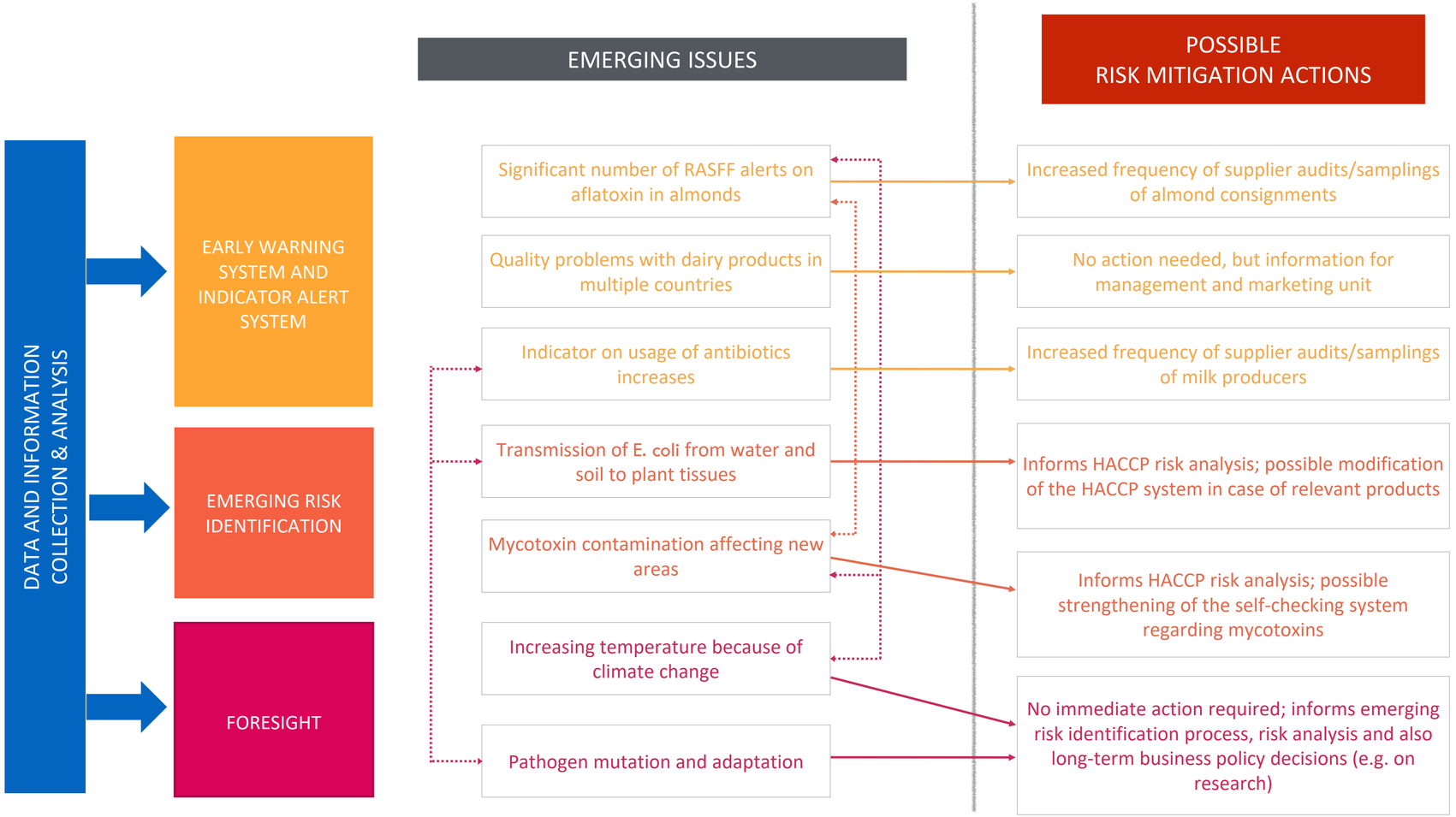
DEPLETION OF NATURAL RESOURCES

CLIMATE CHANGE

TECHNOLOGICAL DEVELOPMENT

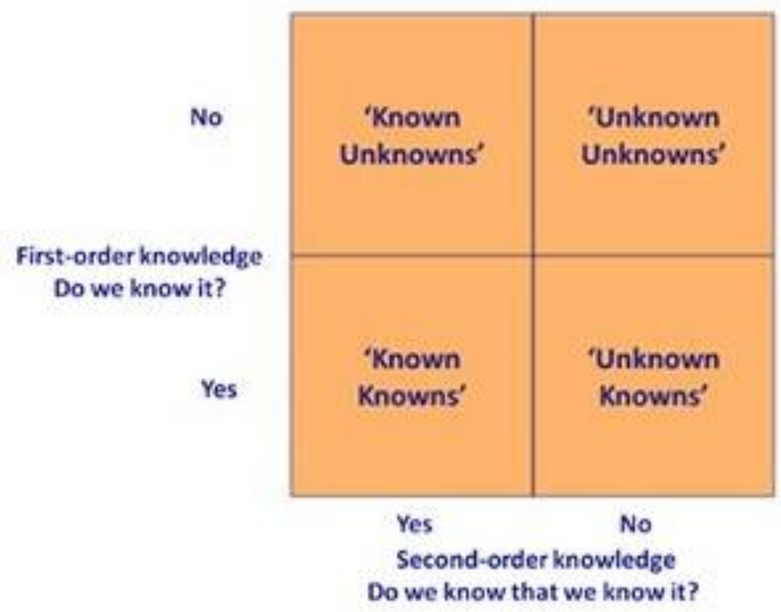
EMERGING ISSUES: EXAMPLE

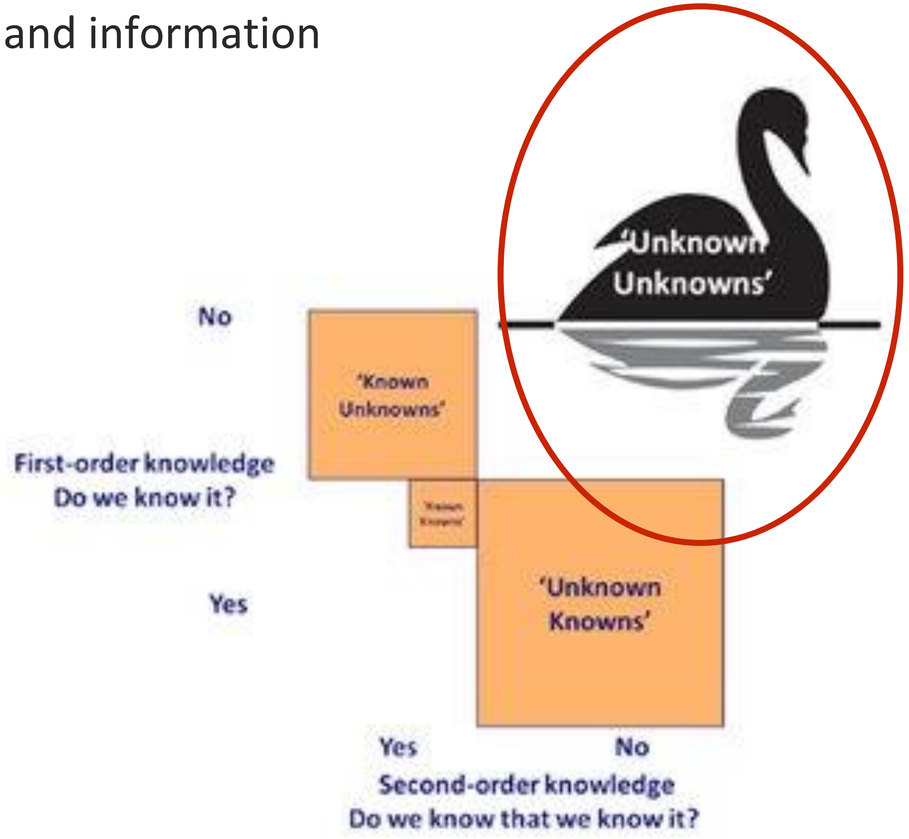
PRACTICAL EXAMPLES



CHALLENGES

- Finding the “next food systems issue”
- Systematic analysis of short, medium and long timescale data and information
- Challenges: data/information gaps, different timescales etc.



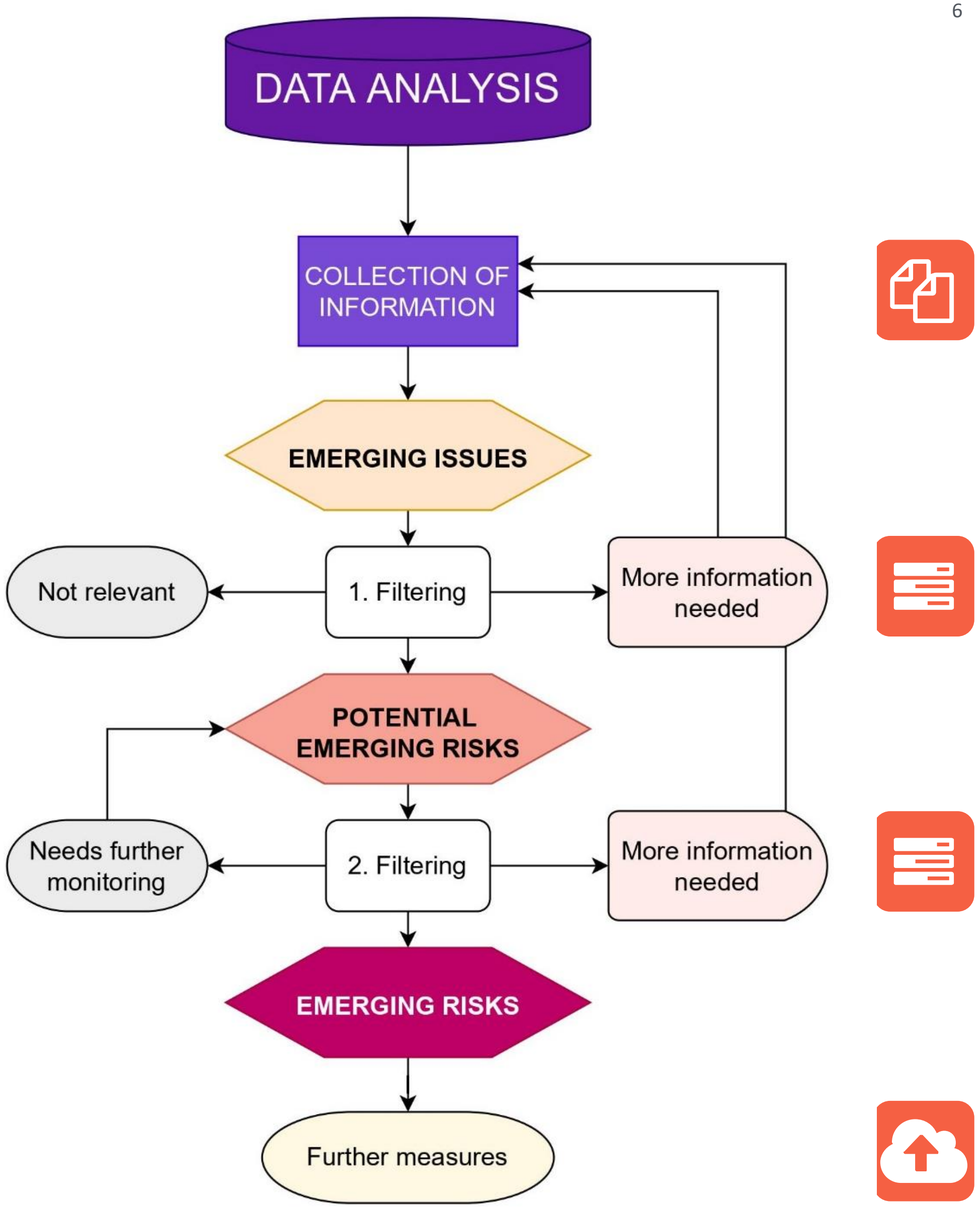


Source: De Spiegeleire, Stephan & Bekkers, Frank. (2012). Who says generals can't dance: Strategic agility and defence capability options.

SYSTEMATIC APPROACH

Process:

1. Collect and collate information
2. Analyze and filter
3. Share information



SYSTEMATIC APPROACH

COLLECTION AND COLLATION OF DATA & INFORMATION

- Screening various data & information sources
  - media and scientific literature
  - data from food safety authorities
  - patent databases
- Collecting expert knowledge

ANALYSIS AND FILTERING

- Characteristics assessed:
  - novelty, significance, susceptibility
- Prioritization: evaluation based on pre-defined criteria
  - soundness, imminence, scale, severity
  - risk management situation

SHARING

- Risk management/preparedness:
  - new procedures, modifying HACCP plans, etc.
- Communication with various target audience:
  - consumers, business, authorities
- Research

TOOLS & METHODS

- Automated data retrieval, Text mining, Data mining, Network analysis, AI&ML, Visualization
- Multidisciplinary team with high level expertise

DATA/INFORMATION SOURCES

DATA/INFORMATION SOURCES FOR EMERGING ISSUE IDENTIFICATION

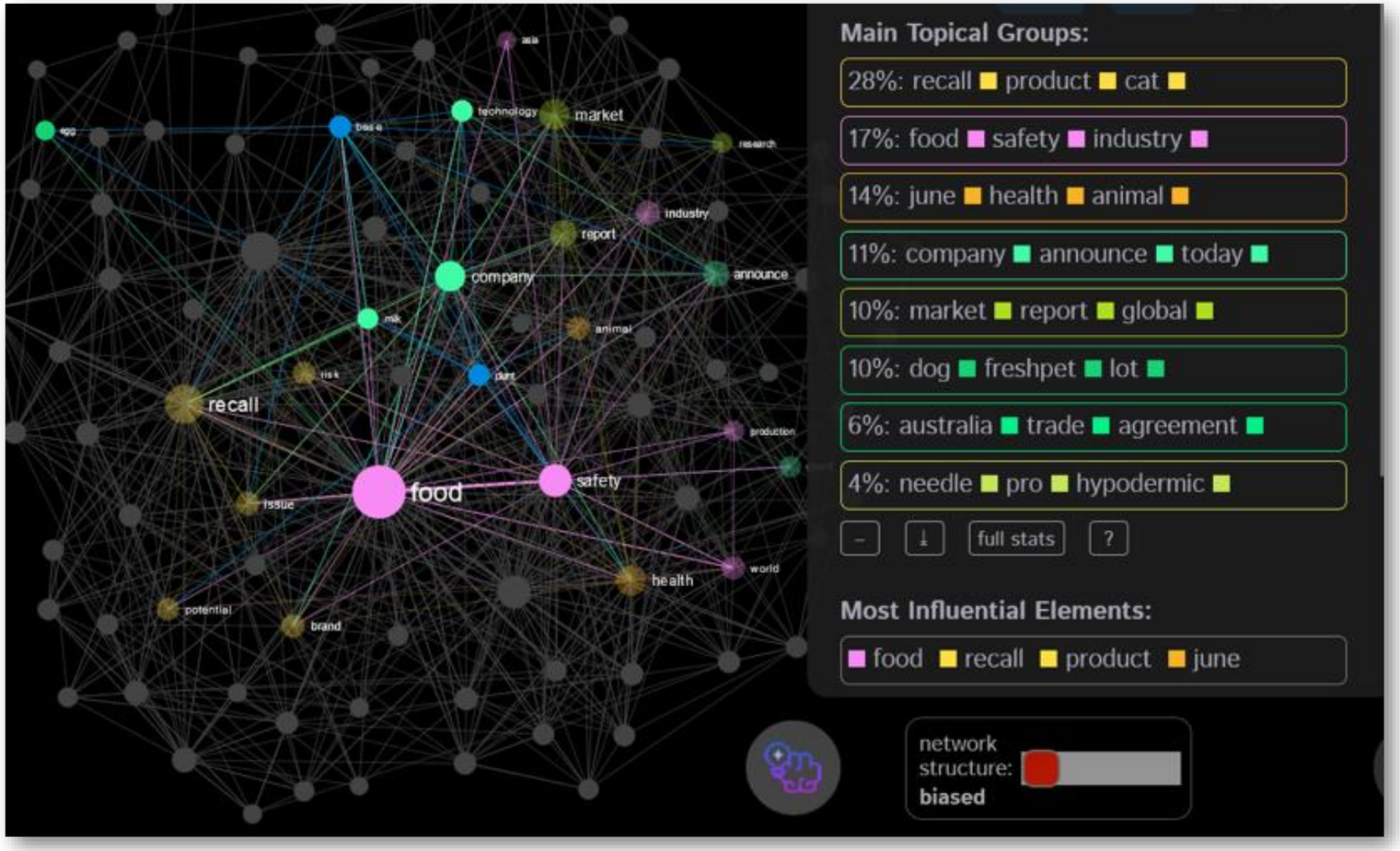
DATA ANALYTICAL METHODS FOR EMERGING RISK IDENTIFICATION

TOPIC DETECTION

DATA ANALYSIS

IDENTIFYING TRENDING TOPICS IN NEWS
BASED ON TEXT MINING AND NETWORK ANALYSIS

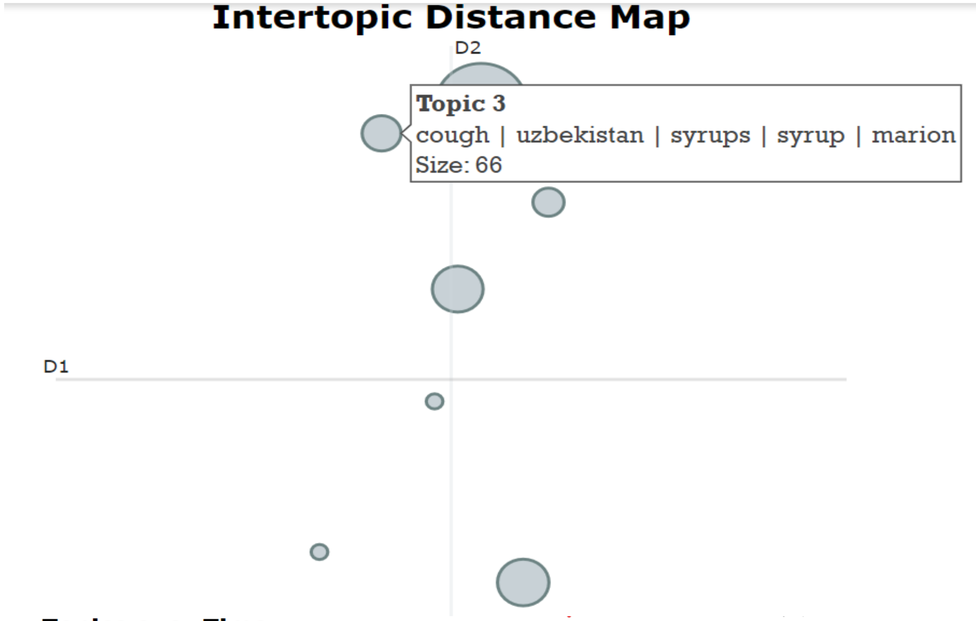
- EMM – European Media Monitor – food safety news from one week is retrieved from RSS feed
- The co-occurrence network of words is arranged into topics
- Network and topics are analysed further manually by experts, who select news which contain information on possible emerging issues/risk

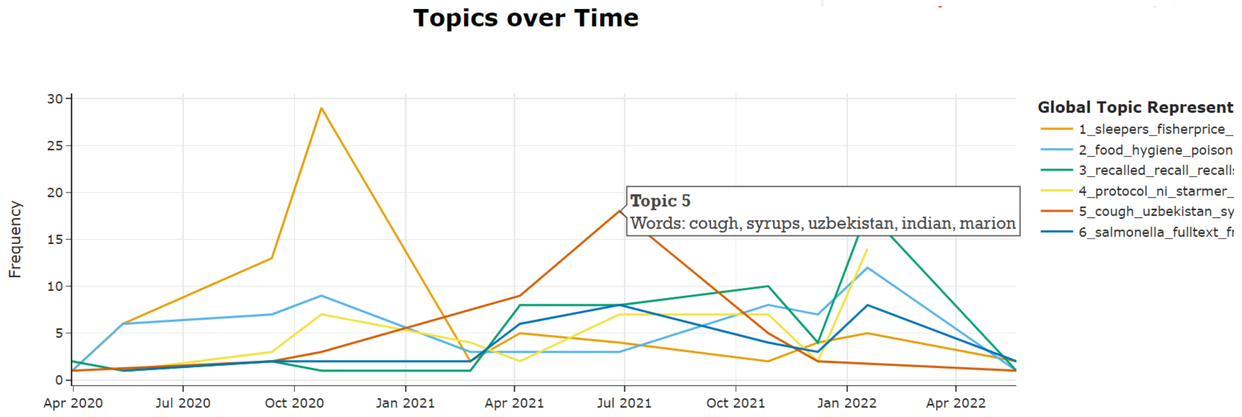


DATA ANALYSIS

IDENTIFYING TRENDING TOPICS IN NEWS
DYNAMIC TOPIC MODELLING

- **BERTopic** framework: most advanced default algorithms
- Dynamic topic modelling is possible
- Challenges:
  - Best algorithm parameters – we don't know beforehand – iterative process for fine-tuning as different datasets require different parameters to get the most meaningful results
  - Best timescales
  - Best input data in terms of text content



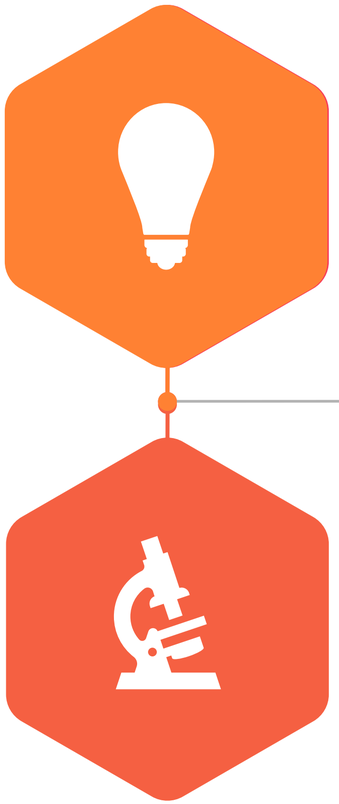


DATA ANALYTICAL METHODS FOR EMERGING RISK IDENTIFICATION

WEAK SIGNAL MINING

DATA MINING UNIVERSE

DEFINITIONS IN DATA MINING: WEAK SIGNALS, STRONG SIGNALS, TRENDS



WEAK SIGNALS

- Low visibility
- Low number of events
- Low interpretation
- **IN DATA MINING: LOW TERM AND DOCUMENT FREQUENCY BUT HIGH GROWTH-RATE**

STRONG SIGNALS

- High visibility
- High number of events
- High interpretation
- **IN DATA MINING: HIGH TERM AND DOCUMENT FREQUENCY AND HIGH GROWTH-RATE**

TRENDS/DRIVERS

- Phenomena that are already known to many people
- Already manifested

Early signal

Early signal

Early signal

EMERGING RISKS

MEDIUM TERM

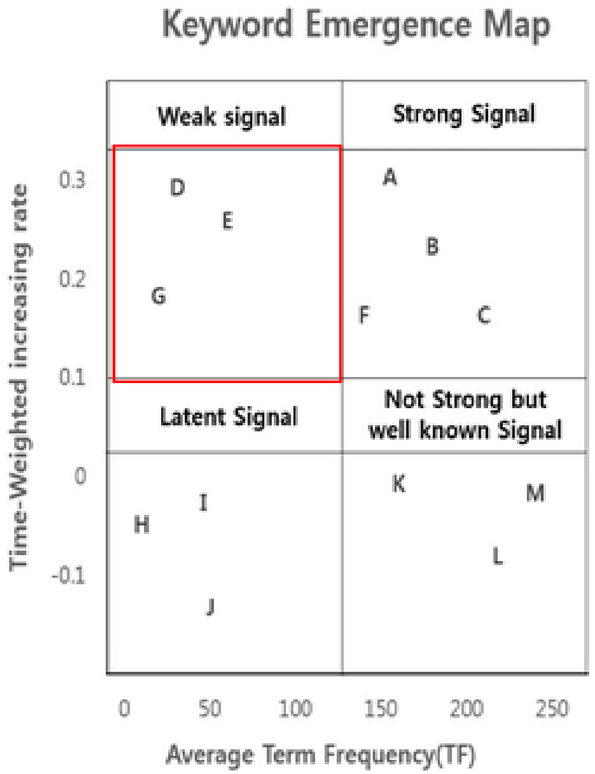
Screening systems

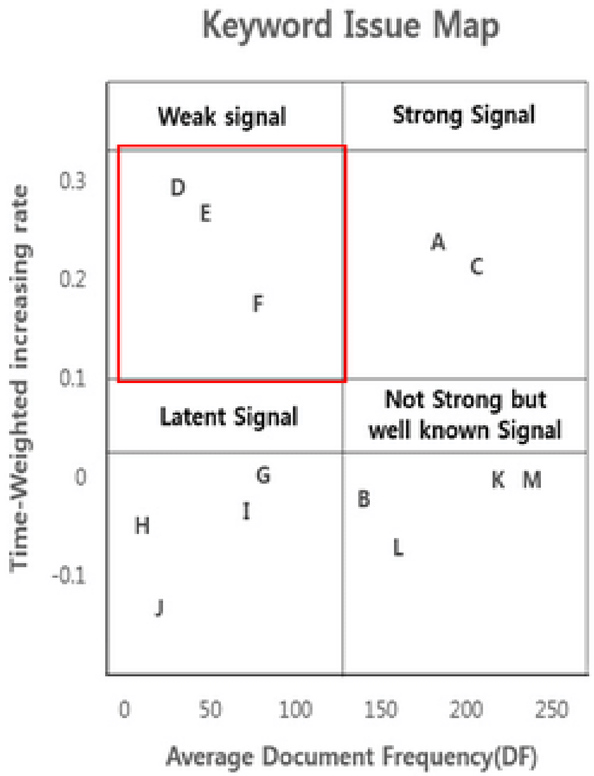
Increases preparedness

Initiates risk assessment

FUTURE RESEARCH DIRECTIONS

WEAK SIGNAL MINING





Future Sign

Category	Common terms
Weak Signal	D, E
Strong Signal	A, C
Not Strong but well known signal	K, L, M

Park et al., 2017

- Takes into account the timely dinamics
- Applicable for keyword detection

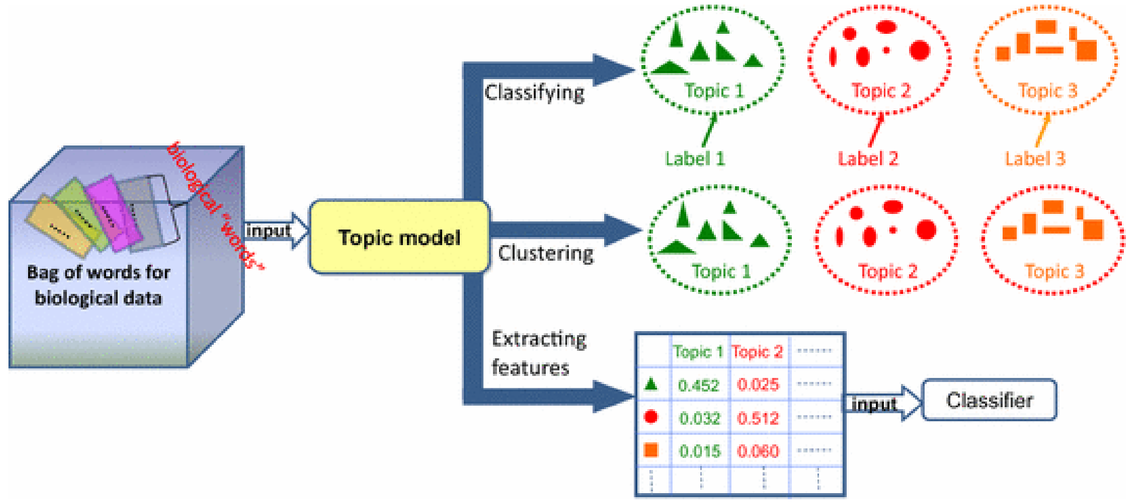
DATA ANALYTICAL METHODS FOR EMERGING RISK IDENTIFICATION

COMBINED METHODS

FUTURE RESEARCH DIRECTIONS

COMBINING METHODS: WEAK SIGNAL MINING AND TOPIC DETECTION

- Limitation of topic modelling: it has no time dimension, static picture (repeated weakly in our workaround)
- Limitation of weak signal mining: no topics and exact definitions for the scale of weak to strong signals
- Many research directions, e.g.:
  - Dynamic topic modelling (dynamic BERTopic)
  - Identifying keywords with weak signal mining to enhance topic modelling
  - Word embedding methods combined with topic modelling



Liu, L., *et al.* An overview of topic modeling and its current applications in bioinformatics. *SpringerPlus* **5**, 1608 (2016).

SUMMARY

KEY MESSAGES

- Data analytical methods are effective in helping ERI and early signal identifying processes, they help in identifying important issues from an extremely noisy environment.
- **But expert knowledge is still important: data validation, choosing the right analytical framework, interpretation of the results, etc.**
- **UTILIZATION: input for monitoring program planning, initiating risk assessments, communication activities**
- The process management system and the possible data analytical methods for emering risk identification (but not early signal detection) is published in Farkas et al. 2023. ***Emerging risk identification in the food chain – A systematic procedure and data analytical options.*** Innovative Food Science & Emerging Technologies. https://doi.org/10.1016/j.ifset.2023.103366

ELSEVIER

Innovative Food Science & Emerging Technologies

Available online 23 April 2023, 103366

In Press, Journal Pre-proof What's this?

Emerging risk identification in the food chain – A systematic procedure and data analytical options

Zsuzsa Farkas [a b], Erika Országh [a b], Tekla Engelhardt [a b], Andrea Zentai [a b], Miklós Süth [a b], Szilveszter Csorba [a b], Ákos Jóźwiak [a b]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.ifset.2023.103366

Get rights and content

THANK YOU FOR YOUR ATTENTION

CONTACT
DIGITAL FOOD INSTITUTE
UNIVERSITY OF VETERINARY MEDICINE BUDAPEST
farkas.zsuzsa@univet.hu
https://dfi.univet.hu/en/