SBORNÍK

prací účastníků vědeckého semináře
doktorského studia
Fakulta informatiky a statistiky
Vysoké školy ekonomické

Abstrakty

Vědecký seminář se uskutečnil dne 28. ledna 2026
pod záštitou děkana FIS
prof. Ing. Jakuba Fischera, Ph.D.

Sestavení sborníku
prof. Ing. Petr Doucek, CSc.
proděkan pro tvůrčí činnost a zahraniční vztahy

Nakladatelství Oeconomica – Praha 2026

ISBN 978-80-245-2582-2

OBSAH

Předmluva

Jak se stalo dobrou tradicí, tak na konci ledna proběhla tradiční akce Fakulty informatiky a statistiky „Den doktorandů“. Seminář proběhl 28. ledna 2026 pod gescí děkana Fakulty informatiky a statistiky prof. Ing. Jakuba Fischera, Ph.D.

Fakulta garantuje celkem tři studijní programy doktorského studia. Celkový počet doktorandů, kteří prezentovali výsledky své vědecké práce byl ovšem poměrně skromný – celkem dvanáct doktorandů. Do tohoto počtu přispěl program „Aplikovaná informatika“ (AI) šesti účastníky, program „Statistika“ (ST) dvěma účastníky a program „Ekonometrie a operační výzkum“ (EOV) čtyřmi účastníky.

Hodnoticí komise posoudila vystoupení doktorandů a na jejich základě určila tři nejlepší vystoupení. V letošním roce tak získali prestižní „Cenu děkana FIS“, s níž je spojena i určitá finanční odměna, následující studentky a studenti:

Studijní program – Aplikovaná informatika

1. místo: **Ing. Alžbeta Gburíková**, Analyzing results of marketing campaigns by procedures of mechanizing hypothesis formation – case study NAPO data (školitel prof. RNDr. Jan Rauch, CSc.)
2. místo: **Ing. Kateřina Haniková**, Towards Graph-Based and Logic-Based Methods for Explaining Misleading Claims (školitel prof. Ing. Vojtěch Svátek, Dr.)
3. místo: **Ing. Veronika Kostrouchova**, Ontological Foundations of the Index of Resilience (školitel prof. Ing. Václav Řepa, CSc.)

Studijní programy – Ekonometrie a operační výzkum a Statistika

1. místo: **Ing. Petr Krautwurm**, The Essence of Product Relationships: Mapping CES to AIDS in the Two-

Product Case (školitel prof. RNDr. Ing. Michal Černý, Ph.D.)

2. místo: **Ing. Karolína Bakuncová**, Time-series clustering application in fertility (školitel prof. RNDr. Luboš Marek, CSc.)

3. místo: **Ebrahim Rezaei**, A Generalized Inefficiency Model for Interpreting Tax Collection Efficiency (školitel prof. Ing. Josef Jablonský, CSc.)

Oceněným studentům doktorského studia upřímně blahopřeji a pevně věřím, že získané zkušenosti uplatní při své další práci, ať už vědecké nebo v praxi. Uznání také patří všem vědeckým a pedagogickým pracovníkům FIS – školitelům doktorandů, kteří se „Dne doktorandů" zúčastnili a svým vedením a radami byli nápomocni při zpracování příspěvků.

Nedílnou součástí „Dne doktorandů" je i práce hodnoticích komisí, jejichž členové pečlivě sledují jednotlivá vystoupení a potom vybírají nejlepší práce k ocenění. Za práci v hodnoticí komisi děkuji prof. Ing. Vojtěchu Svátkovi, Dr. (Katedra informačního a znalostního inženýrství), Ing. Janu Kučerovi, Ph.D. (Katedra informačních technologií), doc. Ing. Mgr. Zdenku Smutnému, Ph.D. (Katedra systémové analýzy), prof. RNDr. Ing. Michalu Černému, Ph.D. a prof. Ing. Josefu Jablonskému, CSc. (Katedra ekonometrie), prof. RNDr. Luboši Markovi, CSc. (Katedra statistiky a pravděpodobnosti). Komise se zhostily své práce na výbornou.

Na závěr bych chtěl vyjádřit zvláštní poděkování studijní referentce doktorského studia paní Ing. Tereze Krajíčkové, DiS, díky níž byl seminář skvěle organizačně zajištěn, dále paní Ing. Martině Jandové a paní Daně Zachariášové za technickou podporu akce a Mgr. Lee Nedomové za práci při editaci a sestavení tohoto sborníku abstraktů.

prof. Ing. Petr Doucek, CSc.

proděkan pro tvůrčí činnost a zahraniční vztahy

STUDIJNÍ PROGRAM APLIKOVANÁ INFORMATIKA

-

AI Act and People Management: Regulatory and Ethical Challenges

Philippe Dubost
qdubp51@vse.cz
Ph.D. student of Applied informatics
Supervisor: doc. Ing. Mgr. Zdeněk Smutný, Ph.D.
(zdenek.smutny@vse.cz)

Organizations increasingly deploy Artificial Intelligence (AI) for People Management use-cases such as employee recruiting, workforce planning, employee performance evaluation, employee training, retention (turnover prediction), and overall workforce analytics, raising urgent ethical and regulatory questions. This systematic review examines how the European Union (EU) AI Act is interpreted and applied in people-management contexts and how it compares with other AI regulations globally on ethical and enforcement aspects. Two questions guide the review:

1. What gaps remain for AI-powered people management in the EU AI Act?
2. Is the EU AI Act stricter or looser than others on ethics/enforcement?

Following PRISMA-2020 (Preferred Reporting Items for Systematic reviews and Meta-Analyses), we searched Scopus for peer-reviewed articles from the past five years. Several hundred records were identified, and 40 papers met the inclusion criteria after screening and full-text review. To enrich the comparison, we introduce a quantitative text analysis (inspired by sentiment analysis techniques) to position several AI regulations on a "regulatory stringency" scale.

Preliminary results show limited coverage of the existing literature when it comes to people management use cases and AI use, as well as the lack

of operational guidance in the EU AI Act to ensure an ethical use of AI in an AI-augmented employer/employee relationship. We assessed that the AI Act is arguably the most stringent legislation to date, compared to other existing AI regulations in different geographies. Many reviewed papers mention the likelihood of the Brussels' effect, meaning that the decisions of the EU may strongly impact upcoming regulations in other geographies. Overall, the review outlines that corporations would benefit from a practical set of guidelines (checklist, procedures, or audit process) to enable ethical use of AI in people management.

Keywords: Recruiting, Workforce planning, Employee performance, Employee training, Employee retention, Comparative AI regulation, Regulatory stringency

JEL Classification: J53, K31, M54, O32

Modern Data Architectures: Evaluation Framework for Selecting Suitable Data Platforms

Felix Espinoza

xespf01@vse.cz

Ph.D. student of Applied informatics

Supervisor: doc. Ing. Miloš Maryška, Ph.D.
(milos.maryska@vse.cz)

This article deals with helping organizations select the most suitable data platform in the context of growing data complexity, where data has become a strategic company asset. It mentions that data complexity and increased volume are closely connected to recent phenomena, namely the Internet of Things, Big Data, and Artificial Intelligence, which bring data heterogeneity. Companies are obligated to react to these trends; therefore, a paradigm shift from traditional data handling practices to sophisticated architectural frameworks can be observed. The article aims to design a universal process through a set of criteria and their evaluation, which will help companies choose the most suitable variant. The variants considered are Data Warehouse (DWH), Data Lake (DL), and Data Lakehouse. The evaluation process and the outputs respect the Design Science Research (DSR) methodology, which is particularly suited for developing and validating practical artifacts in information systems. For a more lucid interpretation, the criteria are divided into three groups: Technical Criteria (TC), Organizational Criteria (OC), and Economic Criteria (EC). This division reflects the fact that companies develop their activities under different conditions, and the selection is a multidimensional problem involving not only technical capabilities but also strategic alignment, organizational readiness, cost considerations, and governance maturity. And since every company expects to satisfy different needs, it is possible to assign different weight to every criterion on the list. The criteria are also evaluated in function of its fulfillment within considered variant. So, for example, the traditional variant (DWH) is assigned only 50 points (maximum 100) in criteria related to type of processed data, since only structured data sources are supported.

This approach makes it possible to determine the extent to which requirements are met. The final score is then calculated using the formula Score_architecture = Σ (weight_i × score_i), where the weight_i = weight of the i-th criterion or question and the score_i = score of the i-th criterion or question

The functionality of the proposed methodology was verified through a practical case study within the financial sector, a field currently undergoing digital transformation. Evaluation criteria were assessed in collaboration with staff members across multiple departments. The DLH solution was identified as the most appropriate choice, a conclusion supported by the professional judgment of several other BI specialists.

Keywords: modern data architecture; data warehouse; data lakehouse; enterprise architecture; data platform selection

JEL Classification: M15, O32

Analyzing results of marketing campaigns by procedures of mechanizing hypothesis formation – case study NAPO data

Alžbeta Gburíková

alzbeta.gburikova@vse.cz

Ph.D. student of Applied informatics

Supervisor: prof. RNDr. Jan Rauch, CSc., (jan.rauch@vse.cz)

The article addresses the use of advanced analytical methods for evaluating data from public opinion surveys, which serve as a basis for marketing and communication campaigns. Traditionally, these data are analyzed only with basic statistical tools, despite their potential for deeper interpretation using data science techniques.

The paper introduces the principles of Mechanizing Hypothesis Formation (MHF), implemented through the GUHA method, which enables applications such as association rule mining, subgroup discovery, exception mining, and contrast set mining.

Based on a case study of the NAPO campaign, the paper demonstrates the use of these procedures for analyzing data from respondent panels. The initial form was data collection by a research agency, which was carried out in August 2024 by the ResSolution research agency on a representative sample of 1027 respondents from the Czech national panel.

The results show that applying MHF provides a more detailed insight into relationships between variables and allows the identification of specific respondent segments, which can lead to more effective targeting of marketing activities.

Keywords: marketing campaigns; data science; GUHA method; association rules, pairs of association rules.

JEL Classification: M15, M31.

Towards Graph-Based and Logic-Based Methods for Explaining Misleading Claims

Kateřina Haniková
katerina.hanikova@vse.cz
Ph.D. student of Applied Informatics
Supervisor: Prof. Ing. Vojtěch Svátek, Dr. (svatek@vse.cz)

This article introduces methods designed to support the explanation of misleading claims that have been fact-checked. Fact-checking is a type of journalism used to debunk claims that are spread in the public space. The aim is to find relevant evidence and facts to support or refute a given claim. Fact-check reports tend to be lengthy, and people unfamiliar with the given topic may struggle to understand the core issue in the claim. This can discourage people from going through the report. The proposed methods aim to provide tools that make the explanation more engaging, thereby drawing more attention to it. For explaining misleading claims, we primarily use semantic web technologies, enabling the explanation to be both human-readable and machine-readable. This article presents three different methods: explanations based on entity graphs, logic, and knowledge graphs. Entity graphs explore fact-checked claims at the level of entities and relationships between them, capturing them as simply as possible and highlighting exactly what is wrong with the claim. This approach relates to logic-based explanations that are at the same level of detail. By capturing facts and misleading claims with first-order logic, we can identify hidden logical fallacies. The last approach aims to enhance current knowledge graphs that focus on collecting fact-checked reports and related information. This third method involves extracting arguments and associated facts as complete sentences and linking them to a claim, making facts and arguments reusable.

Keywords: misleading claims, entity graphs, knowledge graphs, logic-based explanation

JEL Classification: D83

Adopting Semantic Web technologies as a backbone of future-proof software development

Patrik Kompuš

qkomp00@vse.cz

Ph.D. student of Applied informatics

Supervisor: prof. Ing. Vojtěch Svátek, Dr., svatek@vse.cz

In this PhD project, new framework for software development in businesses is being developed and evaluated, based on qualitative research among different key role players during development phase, e.g. data domain experts, software developers and product owners. The separation of responsibilities for data model development and later usage is being used as a key idea, around which the framework is built. The components of Semantic Web technologies, namely ontologies, are evaluated as the means for such separation. Firstly, qualitative research is conducted with the goal to research and understand the different levels of Semantic Web technology acceptability across businesses dealing with software development. Afterwards, selection of latest state-of-the-art methodologies for data modeling across domains within the scope of Semantic Web technology is evaluated and selected to be used within the framework. As the next step, the selection of latest state-of-the-art tools for software developers with Semantic Web technologies native support will be evaluated and selected to be used within the framework. Following, the whole framework will be used to develop demonstrative software as part of the A-B testing, which will be needed for the framework evaluation in terms of time and resources cost optimization during the software development. Lastly, a qualitative research of the framework adoption will be performed and evaluated, which will deliver final results of the PhD project.

Keywords: semantics in services and processes, interoperability, business processes optimization, ontology modeling tools

JEL Classification: L15, L86

Ontological Foundations of the Index of Resilience

Veronika Kostrouchova

Kosv08@vse.cz

Ph.D. student of Applied informatics

Supervisor: prof. Ing. Václav Řepa, CSc. (repa@vse.cz)

Recent research on resilience has highlighted persistent conceptual ambiguity across disciplines, limiting both theoretical coherence and empirical applicability. Barcelos et al. address this challenge through an ontological analysis grounded in the Unified Foundational Ontology (UFO), defining resilience as a capability inherent in an object at risk, constituted by a specific configuration of capabilities and vulnerabilities. While this framework provides a domain-independent and conceptually rigorous definition of resilience, it deliberately treats resilience as binary (either present or absent) and leaves questions of gradation, internal composition, and empirical differentiation open for future research.

This article proposes a structured extension of the ResiliOnt framework by integrating the Index of Individual Resilience (IIR) developed by the Solvo Institute as an empirical specialization at the level of the individual and, by extension, society. The IIR operationalizes individual resilience as a multidimensional configuration of eight domains: values, trust in institutions, social cohesion, adaptability, mental health, skills, material security, and physical activity and health. These domains are measured through a standardized questionnaire and aggregated into a weighted index, allowing systematic comparison across individuals and contexts without embedding sociodemographic attributes directly into the resilience construct.

From an ontological perspective, the article demonstrates that each IIR dimension can be interpreted as a cluster of capabilities that modulate specific vulnerabilities when the individual is modeled as an Object at Risk. Physical health, mental health, material security, skills, adaptability, social cohesion, trust in institutions, and values function as context-dependent capabilities that either disable vulnerabilities or, when absent or weakened, give rise to them. In this way, we do not redefine

resilience but empirically specify the internal capability configurations that ResiliOnt defines abstractly.

By mapping the IIR onto the ResiliOnt ontology, the article shows how resilience can be simultaneously understood as an ontologically well-founded capability configuration and as an empirically measurable phenomenon. This integration enables the analysis of resilience across micro- and macro-levels, supports comparative research across populations and societies, and provides a pathway toward addressing gradation and robustness of resilience without compromising ontological clarity. The proposed extension thus bridges formal ontology and empirical social research, while supporting the systematic development of resilience models across levels and contexts, advancing both the theoretical and practical study of resilience.

Keywords: Resilience, ontology, UFO, OntoUML, capability, vulnerability, individual resilience, ResiliOnt, Index of Individual Resilience

JEL Classification: D81, I31, Z13

STUDIJNÍ PROGRAM EKONOMETRIE A OPERAČNÍ VÝZKUM

Modelling Race Dynamics in Formula One: A State-Space Approach for Lap Time Decomposition

Karolína Jindrová

karolina.jindrova@vse.cz

Ph.D. student of Econometrics and operations research

Supervisor: doc. Mgr. Vladimír Holý, Ph.D.

This paper proposes a structural time-series methodology for modelling lap-level performance in Formula One races. The analysis is based on the premise that observed lap times reflect the joint influence of directly measurable race-specific factors and latent performance dynamics that evolve gradually over the course of a race. While elements such as fuel load, pit stops, and strategic interactions between drivers can be incorporated explicitly, key processes including tyre degradation and track evolution remain unobserved and must be inferred from the data.

To address this structure, the paper introduces the *Racing Pace and Conditional Events State-Space Model* (R-PACE-SSM), a state-space framework that decomposes lap times into a stochastic latent performance state and a set of observable regressors. The model is estimated via maximum likelihood using the Kalman filter and smoother, allowing for a flexible representation of time-varying performance without imposing rigid functional assumptions on underlying degradation processes. This formulation enables the data to govern model complexity at the driver level and accommodates heterogeneous race dynamics within a unified framework.

The empirical application uses lap-by-lap timing and positional data from a Formula One Grand Prix. Key contextual variables — including pit stops, blue flag disruptions, and on-track battles — are constructed through fully automated novel algorithms, enabling consistent event detection across drivers. The results confirm that fuel mass, strategic interactions, and blue flag events significantly influence lap times. The

proposed methodology adapts flexibly to individual race data, leading to distinct patterns of model simplification across drivers. Overall, the R-PACE-SSM proves robust across heterogeneous race conditions and provides a unified, interpretable decomposition of lap-level performance.

Keywords: State-Space models, Kalman filter, Kalman smoother, Structural time series, Formula One, Circuit motorsport, Performance modelling

JEL Classification: C12, C13, C51

Analysis and calibration of mean-reversal strategies on capital markets

David Jukl

David.j@seznam.cz

Doctoral of program Finance at VŠFS

Supervisor: Doc. RNDr. Jan Lánský Ph.D.
(zizelevak@gmail.com)

The aim of this paper is to analyze and evaluate the performance of the mean reversion trading strategy applied to various financial assets, including currency pairs, commodities, and stock indices. The strategy is credited with the ability to generate profits by identifying short-term deviations in price from its medium-term average. The thesis emphasizes the selection of the optimal length of the moving average (MA) depending on the volatility of the monitored instrument.

In the initial phase, classic approaches to mean reversion are tested on basic currency pairs, including additional filters such as ATR or trend MA. The work is then extended to test other currency pairs, highly volatile assets, and stock indices. The results show that for different levels of volatility, it may be advantageous to use different models to calculate the length of the MA, including a simple inverse model and a linear regression approach.

Based on extensive testing, the author proposes a generalized formula for setting a mean reversion strategy depending on the characteristics of individual markets. The work provides practical recommendations for quantitative traders and academics involved in financial time series modeling.

Keywords: Mean reversion; quantitative trading; moving average; volatility; algorithmic trading; financial time series.

JEL Classification: C58, G11, G17.

The Essence of Product Relationships: Mapping CES to AIDS in the Two-Product Case

Petr Krautwurm

Petr.krautwurm@vse.cz

Ph.D. student of Econometrics and operations research

Supervisor: prof. RNDr. Ing. Michal Černý, Ph.D. (cernym@vse.cz)

This paper clarifies the gap between preference-based and market-based definitions of substitutes and complements; it does this by using CES as a preference-based benchmark and testing whether the Almost Ideal Demand System can reproduce the same relationship content in demand behavior. A Monte Carlo agent-based comparison of local demand responses yields a stable translation between CES' relationship parameter and the AIDS' relationship parameter: a monotone, nonlinear relationship well approximated by a power-law–type mapping. The analysis further shows that, in the two-product AIDS case, the elasticity of substitution and the cross-price elasticity of demand can be made to align in a systematic way, which provides a workable bridge for extending relationship encoding beyond two goods, while also indicating limitations of elasticity of substitution as a diagnostic in AIDS.

Keywords: Product relationships, Constant Elasticity of Substitution, Almost Ideal Demand System, Monte Carlo Simulation

JEL Classification: C5, C63, D11

A Generalized Inefficiency Model for Interpreting Tax Collection Efficiency

Ebrahim Rezaei

Reze00@vse.cz

Ph.D. student of Econometrics and operations research

Supervisor: Josef Jablonsky (jablon@vse.cz)

This paper aims to present a unique interpretation of an efficient tax collection system. In light of the divide between 'optimal tax theory' and the practical design of a tax system, we have focused on the overlooked aspects of optimal taxation. By generalizing the simultaneous variations and dependencies of inputs and outputs, we have developed a generalized inefficiency model. Our results highlight the causal effects of policy tools such as endogenized tax expenditure and seigniorage, as well as collection costs on tax collection inefficiency. These factors are identified as the main origins of differences in the performance of clustered OECD economies from 2009 to 2021. Additionally, we conducted further modeling to confirm the validity of the initial results.

Keywords: Tax collection inefficiency, optimal tax theory, practical design of tax system, generalized model, panel causality test, OECD.

JEL Classification: H21, H26, H40, H83, C61

STUDIJNÍ PROGRAM
STATISTIKA

Time-series clustering application in fertility

Karolína Bakuncová

karolina.bakuncova@vse.cz

Ph.D. student of Statistics

Supervisor: prof. RNDr. Luboš Marek, CSc., (marek@vse.cz)

Time series clustering analysis is commonly used as a tool for finding hidden structure in the data. Currently, this method is most often applied in areas such as medicine or finance. The aim of this study is to verify the possibilities of using time series clustering analysis in a demographic context, specifically when examining structural similarities in the development of total fertility across regions of the Czech Republic. Attention is paid in particular to identifying common development patterns in regions grouped into the same cluster and, conversely, emphasizing the differences between clusters. The analysis shows that regions can be divided into three basic groups, which differ most in fertility development in the years 2007-2008 and 2022-2023. These key periods became the starting point for a more detailed examination of differences between clusters based on selected socioeconomic characteristics.

Keywords: Demography, Clustering, Time-series

JEL Classification: C10, J13

Gender and Wealth Inequality in the Baltic States

Alisha Marcinová

mara64@vse.cz

Ph.D. student of Statistics

Supervisor: doc. Ing. Diana Bílková, Dr. (diana.bilkova@vse.cz)

The presentation aims to explain the difference in net wealth distribution between households represented by men and women respondents, with a particular focus on the Baltic states (Estonia, Latvia, and Lithuania), due to having a slightly higher proportion of women in the population. Based on the latest available data from the Household Finance and Consumption Survey, the analysis is carried out at the microdata level. To study gender-related wealth disparities, the presentation employs decomposition methods, namely the Oaxaca-Blinder decomposition at the mean and the Melly decomposition across the 10th, 50th, and 90th percentiles. The analysis focuses on the role of income, the current price of the household main residence, and socio-demographic characteristics, including education, age, and household composition.

The results indicate that, on average, households represented by men exhibit higher levels of net wealth than those by women across the Baltic countries. Decomposition results suggested that a portion of the observed wealth gap was attributed to differences in observable characteristics. The Melly decomposition reveals that gender differences in wealth are not uniform across the distribution, with more pronounced gaps at the lower end and diminishing differences toward the upper tail.

Overall, the findings demonstrate that gender-based wealth inequalities in the Baltic states are shaped by both differences in economic resources and structural factors. By combining microdata analysis, decomposition techniques, and distributional perspectives, the presentation provides clear insights into how gender and household characteristics jointly influence net wealth outcomes.

Keywords: net wealth, gender, Baltic states, decomposition

JEL Classification: D31, J16