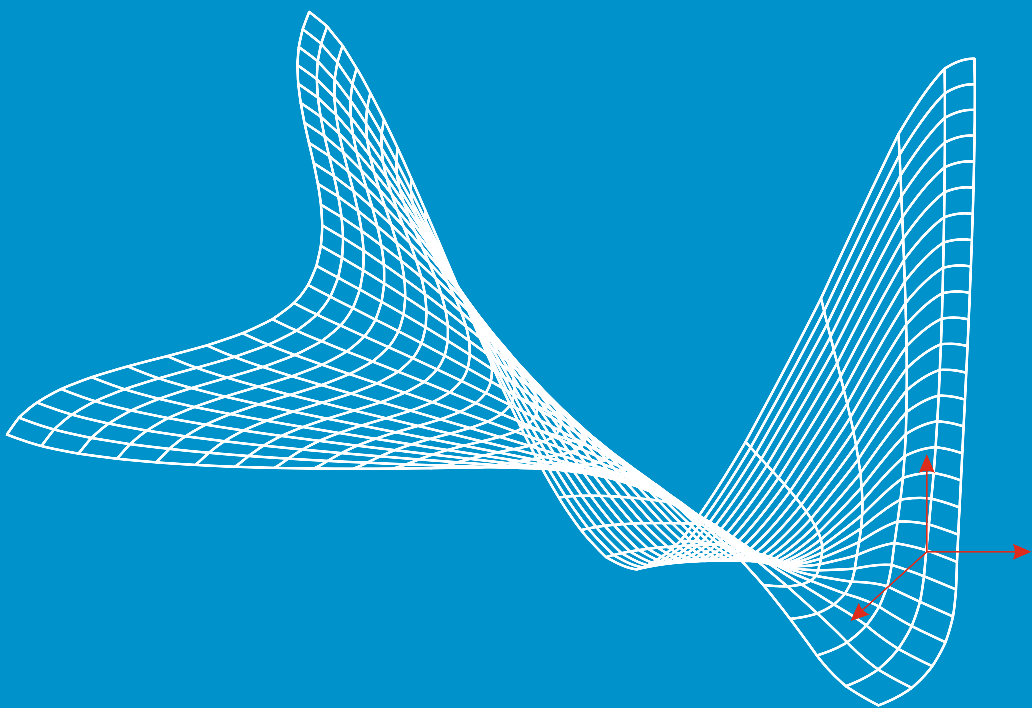
CEREM

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT



Vol. 9, No. 4, December 2025

ISSN 2543-9472
e-ISSN 2544-0365

WSB Merito University
Wrocław

WSB MERITO UNIVERSITIES

CENTRAL EUROPEAN REVIEW OF ECONOMICS AND MANAGEMENT
Volume 9, Number 4
December 2025

Vol. 9, No. 4

Publisher: Uniwersytet WSB Merito Wrocław (WSB Merito University Wrocław)
ul. Fabryczna 29-31, 53-609 Wrocław, Poland

Contact: Johannes (Joost) Platje, johannes.platje@wroclaw.merito.pl, ul. Fabryczna 29-31, 53-609 Wrocław, Poland

The journal is reviewed according to the principle of double-blind peer review and in compliance with the standards of the Polish Ministry of Science and Higher Education. CEREM is a continuation of the WSB University in Wrocłąw Research Journal (Zeszyty Naukowe WSB we Wrocławiu – ISSN 1643-7772; eISSN 2392-1153)

CEREM is registered in the IC Journal Master List, and indexed in the Central and Eastern European Online Library (CEEOL), Index Copernicus, Google Scholar, EconStor, BazEkon, Crossref and ERIH PLUS.

ISSN 2543-9472; eISSN 2544-0365
Cover and logo design: Sebprojekt.pl
Publisher: Uniwersytet WSB Merito Wrocław (WSB Merito University in Wrocław), ul. Fabryczna 29-31, 53-609, Wrocław, Poland

CONTENTS

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 9, No.4, December 2025, 7-28

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Bulls or Bears: how does investor sentiment shape energy profit spreads?

Naima BENTOUIR

University of Ain Temouchent, LMELSPM, Algeria

E-mail: naima.bentouir@univ-temouchent.edu.dz
ORCID: https://orcid.org/0000-0003-3613-9349
Received: 18.06.2025, Revised: 21.07.2025, Revised: 14.08.2025, Revised: 01.10.2025, Accepted: 03.11.2025
doi: http://10.29015/cerem.1036

Aim: The purpose of this study is to look into how investor sentiment affected the profit spreads of three significant oil and gas companies from 2014 to 2024: PB PLC, Exxon Mobil, and Chevron. The study investigates the relationship between investor sentiment and market performance, especially during times of global upheaval like the COVID-19 pandemic and the crisis in Russia and Ukraine.

Design / Research methods: The study evaluates the correlation between daily profit spreads and weekly investor sentiment indices using a Mixed Data Sampling (MIDAS) regression framework. Analysis is done on data from January 2014 to May 2024. To assess the MIDAS model's performance and explanatory power, the results are contrasted with those of conventional linear regressions.

Conclusions / findings: The findings show a complicated and firm-specific relationship between profit spreads and investor sentiment. However, there are mixed positive and negative MIDAS regression coefficients for lagged sentiment. Because of its strong mean reversion and emphasis on long-term projects, Chevron exhibits little sensitivity. Although the direction and interpretation are still unclear, Exxon Mobil exhibits some notable sentiment effects. The oil and gas industry's long-term orientation is reflected in PB PLC's inconsistent sentiment effects. In general, it seems that investor sentiment has little effect on daily stock prices, but it might have a more subtle effect on profit spreads.

Originality / value of the article: By integrating high-frequency sentiment data into asset pricing models, this paper adds to the expanding body of research on behavioural finance. In a turbulent decade characterised by pandemics and war, it employs the MIDAS approach in a novel way to evaluate the sentiment-profitability relationships among three multinational oil companies

Keywords: Bulls, Bears, investors' sentiment, profit spread, Impact, MIDAS.
JEL: C32, C58, G11, G12, G14, Q40.

Correspondence address: Naima BENTOUIR, University of Ain Temouchent, LMELSPM, Algeria.. E-mail: naima.bentouir@univ-temouchent.edu.dz.

1. Introduction

The widespread consensus is that underlying firm performance has an impact on stock markets. However, it is also well acknowledged that external variables (such as investor emotion) as well as fundamentals drive markets. Therefore, it is neither reasonable or appropriate for investors to assume that stock prices will reliably and timely reflect a company's fundamental value. The dominance and ubiquity of trading connected to noise in the market additionally create the potential for mood to play a major role, leading to stock market overreactions or underreactions, which can result in severe price distortions. (e.g. Baker, Wurgler 2006; Hirshleifer 2015; Zhou 2018; Narayan et al. 2021; Nyakurukwa, Seetharam 2023; Du et al. 2024). Despite the fact that sentiment is strongly supported by a large body of research and evidence, it is interesting to note that relatively few academic papers have fully explored the complex relationship between sentiment and more conventional measures of stock market valuation, like the profit spread, or have identified the factors that determine sentiment. One important indicator is the profit spread, which is the stock market value divided by the company's book value of equity. It provides important information about how a company is valued overall in the stock market. (Li et al. 2023).

Investors must comprehend the subtleties and complexity of stock market behaviour beyond company performance as the financial markets continue to change and adapt (Goldstein 2023; Sharma, Ranjan 2021). Although fundamentals are unquestionably important, it is important to recognise that markets are complex systems that are impacted by both internal and external causes, such as investor emotion (Goodell et al. 2023).

Aware that a company's stock price is not always correlated with its intrinsic worth, prudent investors need to be cautious and steer clear of oversimplified assumptions. The interaction of sentiment-related trading highlights the possible influence of sentiment, which can cause noticeable over- or underreactions in the market (Goodell et al. 2023; Sattar et al. 2020). It is important to keep in mind that these heightened swings could lead to significant pricing distortions.

It is puzzling that so few academic papers have really examined the factors that influence sentiment or the complex relationship between sentiment and well-known stock market valuation metrics like the profit spread, in spite of the overwhelming evidence and body of research demonstrating the importance of sentiment. A critical metric that compares the stock market value to the company's book value of equity is the profit spread. It provides important information about how a company is valued overall in the volatile stock market environment.

It is nevertheless critical for academics and industry players to identify the underlying causes of sentiment and its far-reaching effects as we negotiate the ups and downs of the financial scene. We can have a more thorough grasp of the delicate dance between sentiment and conventional valuation metrics by disentangling the complex web of influences. The relationship between these components might be the key to deciphering deeper insights into the intricate workings of the always shifting stock market tapestry, which would enable better decision-making and a more obvious route to stock market investment success. Hence, the primary objective of the present paper is to examine the impact of investor sentiment on profit spreads by using a MIDAS regression model and to determine whether sentiment-based trading is an important contributor to valuation distortions in Saudi Arabian stock market.

We build on the work of (Bergström, Svensson 2020) by providing extensive documentation of a deep and compelling relationship between emotion and profit spread that makes use of a direct measure of stock overreaction. This cleverly gets around any possible problems arising from the unique invocation of convergence.

To sum up, our study makes a substantial contribution to our understanding of the dynamics of financial markets. We shed light on the factors behind emotion, stock overreaction, and market behaviour by exploring their complex interplay. Our findings will help investors, regulators, and corporate decision-makers make better investment decisions by pointing them in the direction of enhanced valuation models, more successful risk management plans, and more knowledge. With more research, the intricate relationship between sentiment, stock prices, and market behaviour can be better understood, advancing the field and enabling market players to confidently and clearly negotiate the constantly shifting financial landscape.

The paper is composed of 6 section. The main literature review is presented below (section 2). In section 3, the data and methodology are described, in section 4, the results are described and in section 5 are discussed, and in section 6, conclusions are drawn.

2. Literature review

Investor sentiments have become an important determinant in financial markets in asset pricing, volatility, and investor behavior. Given the strong interaction of psychological bias and wrong informational asymmetries with market imperfections, within the factor and along with macroeconomic events, such is the case. This section will present a critical synthesis of the past literature that relates how stock returns are influenced by investor sentiment, with attention to varying methodologies, geographies, and changes in sentiment measures.

The early studies such as (Fisher, Statman 2000) persuade concerning sentiment from Wall Street strategists, retail investors, and newsletter writers. Their novelty unfolds quite contrarian because high optimism is always repudiated by a subsequent S&P 500 bad performance. In brief, it points out that this kind of sentiment from small-cap shares seems to be better predicted, while strategist sentiment more coincides with that of individual investors in large-cap equities.

Among the countries under consideration, consumer confidence has been taken as a proxy for investor sentiment within 18 developed countries, (Schmeling 2009) Highly elevated sentiment would be expected to spell lower aggregate returns, particularly for countries with the strongest herd behavior and inferior market integrity. Such relationship still holds for different forecasting horizons as well as asset classes, illustrating the behavioral finance fundamental assumption to be that mispricing is created by beliefs rather than by fundamentals.

In emerging markets, by contrast (Corredor et al. 2015) investor sentiment has a much greater influence in Poland, Hungary, and the Czech Republic as compared with the mature European exchanges. Their findings also point to the specific roles played by country factors, showing how stock characteristics such as size and liquidity act as moderators in sentiment sensitivity. In the same line the study of (Shi et al. 2022) reveals that return implications of local investors' emotions on sectors may differ in countries such as China, Brazil, India, Mexico, Indonesia, and Turkey, while the global sentiment alone might not suffice to make predictions (Parveen et al. 2020) analyzed the biases investors exhibited in Pakistan's stock market pegged on heuristics and overconfidence. Using a composite sentiment index consisting of adjusted turnover, buy-sell imbalance, and RSI (Kim, Lee 2022) analyze the markets in Korea. They find that significant sentiment-return relationships exist with regard to KOSDAQ partly due to the dominance of individual investors. The most important factor here is the way mobile trading influences investment rationality shifting from speculation to informed decision-making. At the same time, event-triggered offers also make-analysis of sentiment shift (Cheema et al. 2020) showed, positive moods of investors in China predicted increased price returns during the bubble in the period of 2006-2008, but they seemed to lose their value when the bubble burst. Results of the study by (Sun et al. 2021) show the impact that the COVID-19 pandemic made on the Malaysian market. The study found that the positive effects of policy measures such as Movement Control Orders were much greater in returns, while daily cases and deaths were less influential. Behavioral bases of sentiment are to be further explored by (He et al. 2019) and (Hu et al. 2021), who put up a case for further examination of Investor Risk Compensation (IRC) under flated emotional conditions. Findings indicate that real IRC has a beneficial effect on returns, regardless of emotional field, whereas retrospective IRC exercises an unyielding negative influence across conditioning, suggesting asymmetries in historical valence.

Emerging proxies for sentiment represent attempts to capture the psychological drivers more accurately. Griffith et al. (2020) applied the Thomson Reuters MarketPsych data to show that fear and stress considerably influence volatility and returns. Edmans et al. (2022) develop a music-based global sentiment index developed from listeners' song choices that correlates with stock market returns and volatility in the same week. This unequivocal, real-time indicator generates mood variations based on seasons, weather, and policy events, thus showing robustness across geographies. Recently, social media has become important for studying investor mood. Duz Tan and Tas (2021) focused on sentiment from Twitter and found it to have predictive power for trading volume and returns even after controlling for news sentiment. Guégan and Renault (2021) analyzed over a million StockTwits posts, linking short-term movements in the price of Bitcoin to bursts of sentiment, especially in times marked as market bubbles. In technological integration, further refinements are done on modeling sentiment. Duxbury and Wang (2024) conclude stronger power in explaining the risk-return relationship exists when combining retail sentiment with institutional sentiment. In a later collaboration with Jing et al. (2021) they proposed a hybrid forecasting model consisting of CNN-based sentiment classification and LSTM neural networks that greatly improve stock price prediction on the Shanghai Stock Exchange.

Among the methodological innovations is an investor attention index constructed, (Chen et al. 2022) which they claimed has been built from a wide array of literary proxies. Composite measure is built through partial least squares and scaled principal component analysis. This measure adds strong predictive power for stock market risk premiums, particularly enhancing the prediction for the most volatile equities. Furthermore, the index's strength lay in reversing short-term price distortions, leading to strong in-sample and out-of-sample returns with corresponding implications for asset allocation. This work includes further developments of an Investor Sentiment Index (ISI) developed for KOSPI above Huang et al. (2015) and partial least squares by (Bouteska et al. 2024). Their index surpasses the Baker-Wurgler BM index on the grounds of validity in predicting and explaining further some small-firm effect anomalies disregarded by traditional asset pricing models. The Sentix database has been used to gauge the consumer sentiment toward Bitcoin in the Cryptocurrency markets (Anamika et al. 2023). This is correlated with a strong positive reaction to prices very contrastively influenced by carrying market sentiment measured by VIX and Baker-Wurgler indices. Guégan and Renault (2021) show that social sentiment has induced high-frequency effects on intraday returns, although profit is still restrained by arbitrage limits. To support the empirical framework of this study, the A series of core concepts is being outlined hereafter:

- ✓ Investor Sentiment Index: A composite measure of investor mood deduced from surveys, media content, social media analysis, and trading behavior.

- ✓ Profit Spread: Difference between the stock price at closing and that at opening on a daily basis, adjusted for dividends and splits; it is used to represent the company's performance and volatility.
- ✓ MIDAS Regression (Mixed Data Sampling): An econometric technique permitting the integration of high- and low-frequency variables, well-suited for combining daily financial variables with weekly sentiment scores.

In the above context, these findings offer a broad and comprehensive backdrop for investigating the role of sentiment in the financial markets while offering methodological and behavioral insights that inform the design and hypothesis of this study.

3. Data and method

This section gives the relevant data, methods of collection, empirical framework and hypothesis on which the analysis would be based. The prime objective of the research is to explore the presence of a great aspect of investor sentiment as of whether profit spreads in the leading companies of energy industry are determined by the investor sentiment or not regarding a period of ten years in the global turbulence, with the help of a MIDAS regression model. The paper considers the role of sentiment-based trading in valuation distortions by contrasting between the sensitivity of profit spreads on sentiment indices without controlling other important macroeconomic and financial factors.

3.1. Data description

The dataset runs from January 2014 until May 2024, covering this time of more sensitive markets in the wake of the COVID-19 pandemic and the Russia-Ukraine crisis. Thomson Reuters database has been chosen for the purpose of data extraction as it included a full-out coverage of financial metrics, sentiment analysis, and company fundamentals.

The analyzed subjects are three of the biggest oil and gas companies worldwide: PB PLC (London, UK), Exxon Mobil (Texas, USA), and Chevron (California, USA). These are chosen in consideration of their enormous global operations, coupled with very strong visibility by investors and a diverse strategy approach. Hence, they offer a balanced cross-sectional perspective concerning sentiment analysis.

Investor sentiment data were collected at a weekly frequency based on the American Association of Individual Investors (AAII) sentiment survey and advanced textual analytics applied to news, social media, and related content. Polling was conducted weekly from Thursday 12:01 a.m. to Wednesday 11:59 p.m., yielding forward-looking indicators that reduce hindsight bias and capture investor perceptions in real time.

Profit spreads, on the other hand, were computed daily, difference between the closing and opening stock prices adjusted with dividends and stock splits. Because of this measure, almost real time actionable firm level performance signals are given concerning market conditions.

Table 1. Data description

Variable Name	Description	Unit	Source	Code in dataset
Investor Sentiment Index	Weekly average sentiment derived from investor surveys and text-based sentiment analysis	Index value	AAII	Sent-Index
Profit Spread of three companies	Daily close-open price difference, adjusted for dividends and stock splits	USD	Thomson Reuters	BP-PLC
				Exxon Mobil
				Chevron

Source: author's elaboration.

Table 1 outlines the key variables, their measurement units, sources, and dataset codes, serving as a reference point throughout the empirical analysis.

3.2. Data sources and collection

All financial data were obtained from the Thomson Reuters platform, which includes:

- Historical Stock Prices: Daily opening and closing prices for PB PLC, Exxon Mobil, and Chevron, adjusted for corporate actions.
- Sentiment Scores: Weekly sentiment scores using AAII's proprietary algorithms, integrating multiple text-based sentiment signals from media channels.

This dual-frequency data enables alignment between investor expectations and market behavior at complementary temporal scales.

3.3. Preliminary analytical steps

Before applying MIDAS regression, this study follows a structured approach: stationarity tests confirm the suitability of time series analysis; an initial OLS model provides baseline insights; and wavelet-based variance decomposition highlights multi-scale volatility patterns across firms. These steps ensure the data is stable, the modeling path is justified, and sentiment effects are captured dynamically.

3.4. MIDAS method

Mixed data sampling (MIDAS) regressions are now commonly used to deal with time series data sampled at different frequencies. This paper focuses on single-equation MIDAS regression models involving stationary processes with the dependent variable observed at a lower frequency than the explanatory on (Foroni et al. 2015; Ghysels et al. 2020).

In a baseline MIDAS regression deal with a single low-frequency variable projected onto a high-frequency variable possibly augmented with the lagged-dependent variable. When the difference in sampling frequencies is small, we may treat this as a regular regression problem, now dubbed the unconstrained MIDAS model namely when dealing with, for instance, yearly projected onto quarterly or quarterly projected onto monthly (Foroni et al. 2019).

The linear and quasi-linear MIDAS models given as follow:

$$Y_t = \beta_0 + \beta_1 B\left(L^{\frac{1}{m}}\right) X_{t-1}^{m} + \varepsilon_t^{m}$$

Where y is the dependent variable, x is the regressor, m denotes the frequency.

3.5. Hypotheses development

H_1: Major oil and gas companies' profit spreads are significantly impacted by investor sentiment.

Justification: According to earlier behavioural finance research, investor sentiment may have an impact on pricing irregularities and departures from core values. Lagged sentiment effects on high-frequency profit data can be captured using the MIDAS regression framework. This hypothesis investigates whether changes in profit spreads across BP, Exxon Mobil, and Chevron are influenced by sentiment.

H_2: Due to differences in market sensitivity and strategic orientation, different firms experience different effects of investor sentiment on profit spreads.

Justification: The way sentiment affects profit spreads may be mitigated by firm-specific traits like Exxon Mobil's steady cash flow strategy or Chevron's long-term investment horizon. In line with theories of firm-level financial behaviour under uncertainty, this hypothesis seeks to investigate the variations in sentiment responsiveness amongst businesses.

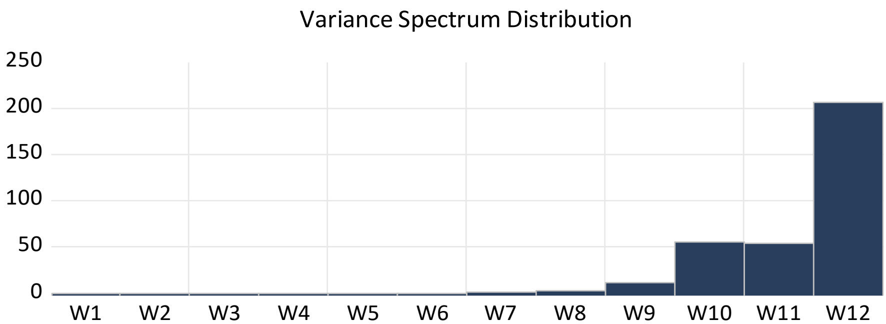
4. Empirical results

This section gives the empirical results of the study starting with graphical representations in order to explain the dynamics of investor sentiment and profit spreads during the sample. These visualizations offer a rough picture of what the patterns and possible relationship between variables are. We then report and interpret the results of the MIDAS regression models and test the hypothesis that investor sentiment is a significant explanation of the changes in profit spreads in the energy sector. The results are explained in the context of the purpose of the study, which sheds some light on the possible contribution of sentiment-based trading to the valuation effects.

4.1. Wavelet analysis

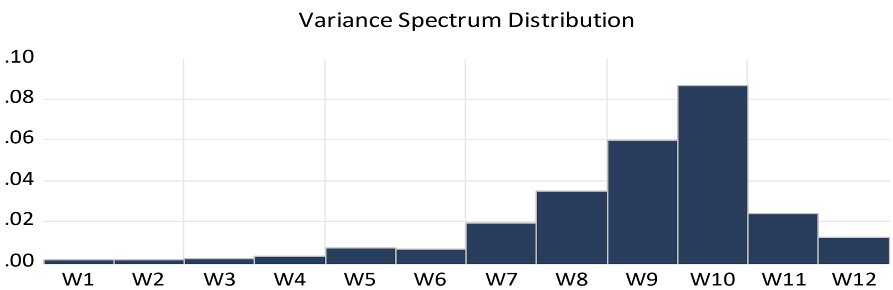
W1-W12 are the scales of the wavelet decomposition where the smallest scales (W1-W4) contain short-term changes in the profit spread, and the large ones (W11-W12) contain long-term changes related to structural forces, including oil price changes, macroeconomics, and geopolitics. The variance decomposition in Figure 1 indicates that 55.24 and 207.35 of the total variability are explained by W11 and W12, respectively, which shows that the profit spread at Exxon Mobil is largely determined by long-term dynamics, and the short-term effects have a comparatively minor influence.

Figure 1. Variance decomposition (Exxon Mobil)



Source: author's elaboration.

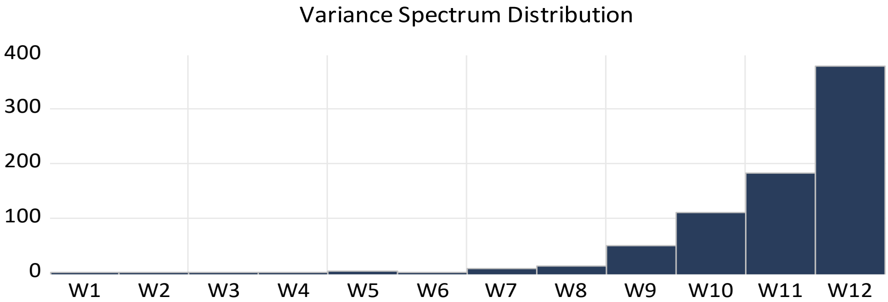
Figure 2. Variance decomposition (BP-PLC)



Source: author's elaboration.

W1-W12 are the scales of Wavelet decomposition with low scales (W1-W4) representing the short-run variations of profit spread, and higher scales (W8-W12) representing long-run structural changes. In Figure 2, it has been indicated that most of the spread variance of the profit of BP PLC is attributed to a limited number of strong components, which are mainly W10, W9, and W8, hence long-run dynamics are the primary source of profitability. Other scales play a minor part each, which explains the relevance of focusing on these major elements when gauging the profit performance of the company and maximizing it.

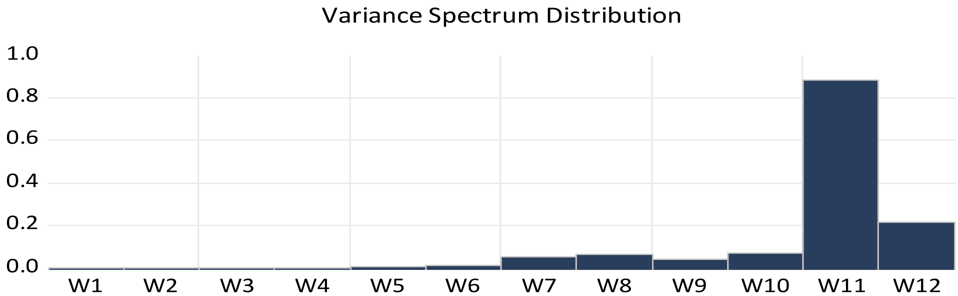
Figure 3. Variance decomposition (Chevron)



Source: author's elaboration.

W1-W12 are the scales of the wavelet decomposition, with smaller scales (W1-W4) capturing short-term oscillations, and larger scales (W8-W12) capturing long-term structural oscillations. The variance decomposition of Chevron (see Figure 3) indicates that the scale (W12) that has the biggest variation in the distribution of profit spreads follows the other two scales W10 and W11. It means that the long-term aspects (such as the macroeconomic conditions, oil price cycles, and strategic investment decisions) are the main determinants of the profitability of Chevron. The lower magnitude of smaller sizes (W1-W8) indicates that the impact of short-term volatility and sentiment on the profit spread of the company is relatively small based on the long-term nature of capital projects and the company cash flows, which Chevron focuses on.

Figure 4. Variance decomposition of Bullish, Bearish, of Sentiment Indices.



Note: BP PLC has a different scale as its returns are more volatile and a daily return data is used on a weekly sentiment data. This difference is factual and not that there is a methodological problem.
Source: author's elaboration.

Figure 4 documents the variance break down of the sentiment index over growing forecast horizons (W1-W12). Over the short term (W1-W4) the decomposition shows that the persistence is largely strong and thus the sentiment index is majorly explained by its own shocks with other variables playing a trivial role. In intermediate horizons (W5-W8) the role of fundamental variables in the contribution, most prominently the profit spread, increases with a modest contribution, which means that the valuation changes have a lagging contribution to investor mood. But at the longest of horizons (W9-W12) there are dominating the innovations of the sentiment index (around three quarters of the forecast-error variance), whereas the returns of oil prices and the dollar index are minor contributors (only a few percent each). These patterns taken together suggest that overall investor sentiment is self-reinforcing in the long-run and, though fundamentals (in particular profit spreads) can have an impact on sentiment, this impact is less immediate and more powerful at intermediate horizons than at the long run.

Table 2. Summary of stationary tests

Method	Statistic	Prob	Cross section	Obs
Levin, Lin & Chu t*	-10.1929	0.0000	3	7782
Im, Pesaran and Shin W-stat	-13.5127	0.0000	3	7782
ADF – Fisher Chi-square	190.949	0.0000	3	7782
PP – Fisher Chi-square	80.4049	0.0000	3	7803

Unit root test results for profit spreads of PB PLC, Exxon Mobil, and Chevron (2014–2024); all tests confirm stationarity at the 1% level (*p < 0.01).
Source: author's elaboration.

As shown in Table 2, these stationarity tests' findings allow us to conclude that PB PLC, ExxonMobil, and Chevron's profit spreads are stationary. This indicates that their statistical characteristics have not changed much over time, which is an essential premise for many time series analysis methods, including the OLS regression-based MIDAS method.

4.2. Midas method results

Table 03 displays from the regression study that the profit spreads of PB PLC, Chevron, ExxonMobil, and Sentiment index AII given on Neutral index have substantial correlations with one another. According to the coefficients, shifts in one company's profit spread are linked to equivalent shifts in investor sentiment. The comparatively low adjusted R-squared values, however, suggest that there may be additional factors, such as geopolitical and economic risks, influencing profit spreads that are not included in our analysis.

Table 3. OLS method

	chevron	**Exxon Mobil**	**PB PLC**
Method	OLS		
Coefficient	-0.3867	0.4369	-0.45057
Std. Error	0.03986	0.0385	0.03824
t-Statistic	-9.6999	11.327	-11.782
P-value	0.0000	0.0000	0.0000
Fit Model			
Adjusted R-squared	0.146714	0.19394	0.205806
Prob(F-statistic)	0.000000	0.00000	0.000000
Schwarz criterion	10.09884	10.1795	10.22538
Hannan-Quinn criter.	10.08428	10.1649	10.21069
Durbin-Watson stat	2.186993	2.21750	2.231667

OLS regression for profit spreads of PB PLC, Exxon Mobil, and Chevron (2014–2024); at the 1% level of significance (*$p < 0.01$).

Source: author's elaboration.

Table 4. MIDAS method based on PDL (Exxon Mobil company)

	PDL01	PDL02	PDL03
Exxon Mobile	MIDAS PDL		
	-0.0271	-0.046397	0.032103
	0.3329	0.301766	0.059327
	-0.0816	-0.153751	0.541126
	0.9350	0.8779	0.5887
Model fit			
Adjusted R-squared	0.230604		
Schwarz criterion	10.17864		

PDL for profit spreads of Exxon Mobil (2014–2024).

Source: author's elaboration.

Table 04 shows the MIDAS-PDL result that lag variables have an impact on ExxonMobil's profit spread; however, the nature of this link varies according on the PDL specification selected. The adjusted R-squared value is relatively low, suggesting that the model may be missing some important components. ExxonMobil's profit spread determinants may be better understood by additional variables or further analysis using alternative PDL specifications.

Table 5. MIDAS method based on PDL (BP-PLC)

	PDL01	PDL02	PDL03
BP-PLC	MIDAS PDL		
	-0.0271	-0.0463	0.0321
	0.3329	0.30176	0.0593
	-0.0816	-0.1537	0.5411
	0.9350	0.8779	0.5887
Model Fit			
Adjusted R-squared	0.230604		
Schwarz criterion	10.21978		

PDL for profit spreads of PB PLC (2014–2024).

Source: author's elaboration.

Based on the table 05, the PB PLC MIDAS PDL analysis indicates that lagged variables have an impact on the company's profit spread; however, the nature of this relationship differs according on the PDL specification selected. The BP profit spread can be affected by lagged variables in both positive and negative ways, and the significance of these impacts varies depending on the specifications, according to the coefficients for PDL01, PDL02, and PDL03. Even while the model only partially explains the fluctuation in the profit spread, as indicated by the adjusted R-squared value of 0.230604, it nevertheless offers insightful information about the possible factors influencing BP's profitability.

Table 6. MIDAS method based on PDL (Chevron)

	PDL01	PDL02	PDL03
Chevron	MIDAS PDL		
	-113.4257	155.9516	-36.76743
	135.0490	128.1874	25.39818
	-0.839886	1.216591	-1.447640
	0.4014	0.2243	0.1483
Model Fit			
Adjusted R-squared	0.230604		
Schwarz criterion	10.21978		

PDL for profit spreads of chevron (2014–2024).

Source: author's elaboration.

Table 6 provides that Chevron's profit spread is also influenced by lagged variables, according to the MIDAS PDL research. On the other hand, BP's PDL specifications' coefficients are much less in magnitude than Chevron's, indicating a weaker correlation between delayed factors and Chevron's profitability. While PDL03's negative coefficient shows the opposite link, the positive coefficients for PDL01 and PDL02 show that rises in the historical values of the independent variables led to increases in Chevron's profit spread. Similar to BP, the modified R-squared value of 0.230604 shows that while the model may not account for all important factors, it does explain some of the fluctuation in Chevron's profit spread.

Table 7. MIDAS method based on steps

Exxon Mobil				
	Coefficient	Std. Error	t-Statistic	P-value
Step 01	-2.844390	15.20956	-0.187013	0.8517
Step 02	-31.48030	41.23640	-0.763410	0.4456
Model Fit				
Adjusted R-squared	0.214320			
Schwarz criterion	10.19678			
BP-PLC				
	Coefficient	Std. Error	t-Statistic	P-value
Step 01	0.039220	0.104097	0.376764	0.7065
Step 02	0.255459	0.139264	1.834360	0.0672
Model Fit				
Adjusted R-squared	0.227609			
Schwarz criterion	10.21156			
Chevron				
	Coefficient	Std. Error	t-Statistic	P-value
Step 01	26.44004	14.29734	1.849297	0.0650
Step 02	-75.35293	38.89295	-1.937444	0.0532
Model Fit				
Adjusted R-squared	0.156605			
Schwarz criterion	10.10485			

MEDAS method based on steps for profit spreads of PB PLC, Exxon Mobil, and Chevron (2014–2024); at the 10% level of significance (*p < 0.1.
Source: author's elaboration.

Table 7 regroups the coefficients exhibit variability among the three companies between Step 01 and Step 02, suggesting a complicated and potentially variable link between the lagged independent factors and the dependent variable (presumably the profit spread). Lagged variables can have both growing and reducing effects on the profit spread, as indicated by some positive and some negative coefficients.

The statistical significance of the coefficients is indicated by the p-values that correspond with them. At the 5% level, a p-value of less than 0.05 often denotes statistical significance for the coefficient.

5. Discussion

Closer inspection into the outputs of the MIDAS regression indicates that the effect of investor sentiment changes from firm to firm regarding profit spreads, which is guided by the strategic profile and time sensitivity of each firm.

According to MIDAS-PDL results on Exxon Mobil, sentiment-related lag variables have logically significant and structurally ambiguous effects on profit spreads. Variations in coefficient signs from PDL01 to PDL03 with a combination of positive and negative lag responses point to the nonlinear, inconsistent transmission mechanism. Investor mood and profit dynamics within Exxon Mobil, which might be ascribed to many macroeconomic variables such as oil price and geopolitical changes, could be a complicated interplay. Although the adjusted R-squared value remains at low levels, some significant lags provide evidence toward the position that sentiment drives up-through short-run profit fluctuations, even though a rather stable cash flow strategy would substitute it from extreme swings.

For the case of PB PLC, the complexity is revealed by the analysis of MIDAS-PDL, shown through several lag effects that operate above multifaceted sentiment. The coefficients in PDL01, PDL02, and PDL03 show both amphitory and dampening effects, indicating that profit spread in BP is sensitive to changes in the mood of investors but in a temporally fragmented way. BP's extended operational exposure-from European regulatory pressures to environmental headlines-explains why sentiment impacts profit spread in both persistent and transient forms. Although the adjusted R-squared (≈0.23) suggests that much of the variation is left unexplained, the mixed-lag results show that some sentiment waves that last for several days resonate with BP's market valuation.

Chevron, on the other hand, exhibits a more insulated relationship with investor sentiment. Although the MIDAS-PDL estimates show that lagged variables do have an effect on Chevron's profit spread, the effects of investor sentiment seem to be weaker in comparison to those of BP and Exxon Mobil. Positive coefficients for PDL01 and PDL02 indicate some responsiveness to historical sentiment values; however, the negative coefficient for PDL03 and the general lack of statistical significance imply that sentiment is not a key driver of Chevron's short-run profitability. This interpretation fits well with the company's strategic focus on long-term capital projects and stable dividend policies that dampen short-term sentiment sensitivity. The adjusted R-squared value also lends credibility to the argument that other structural factors like oil futures, investment cycles, and internal cash-flows management have greater explanatory power in the case of Chevron, just as it has for BP.

The MIDAS results strengthen Hypothesis 2 by demonstrating how investor sentiment interacts differently between firms according to their financial architecture and strategic focus. BP-PLC and Exxon Mobil show greater sensitivity to sentiment through different lag structures, while Chevron's limited reactivity demonstrates its immunity to short-term market noise emanating from an emotional source. These dissimilarities are essential to interpretation from a behavioral finance perspective across various types of firms in the energy sector.

6. Conclusion

The moulding of investor sentiment and behavioral pattern among markets by economic occurrences and crises is a multifaceted event, which has been pointed out in many foundational works (Baker, Wurgler 2006; López-Cabarcos et al. 2020; Li et al. 2023; Du et al. 2024). In traditional theories in finance, the psychological and social bases are often overlooked partly explaining investor actions (Nareswari et al. 2021; Akhtar, Das 2020). This becomes apparent in the aforementioned works (Lerner 1936; Prechter 2016; Shiller 2000; Shleifer 2000; Steelman 2016) as well as Parker and Keynes who understand the point that the dynamic nature of markets is therefore tied up with social cues and, through social narratives, emotional reactions.

While empirical studies (Dhaoui 2015) confirm the strong effects of sentiment on trading volumes and market shocks, isolating individual investor behavior has been consistently cited as a challenge throughout the literature. The main objective of this paper was to examine the significant impact of investor sentiment on profit spreads in the energy industry. The present study addresses this issue by establishing an empirical link between investor sentiment, measured through media signals, and profit spreads for PB PLC, Exxon Mobil, and Chevron firms via MIDAS regression and wavelet decomposition.

The results confirm both hypotheses. Firstly, Hypothesis 1 is supported, in that investor sentiment has a measurable impact on profit spreads of the major oil companies, though the strength and clarity of this impact can alter. Secondly, the characteristics of the firm influence the extent of the sentiment's effect, thus supporting Hypothesis 2: BP is mixed in its responsiveness to sentiment, Exxon Mobil has some statistically significant effects, and Chevron appears to be largely insulated from sentiment due to strategic considerations.

We notice that during economic shocks, high sentiments strengthen the overreaction of the market, which then causes asset valuations to drift away from intrinsic benchmarks. This further confirms the behavioral patterns that bring about investor mispricing, volatility, and distress. The findings carry very important implications for valuation models, investment policy, and risk management: with an understanding of sentiment-based valuations, both investors and regulators would better anticipate the market-response dynamics during crises.

Overall, the results reinforce adding sentiment data to financial models in real time. In the future, research ought to investigate how economic headlines and global uncertainties create emotional narratives that shape investor expectations, asset pricing, and market resilience.

References

Anamika Chakraborty M., Subramaniam S. (2023), Does sentiment impact cryptocurrency?, "Journal of Behavioral Finance", vol. 24 no. 2, pp. 202–218.

Bergström M., Svensson J. (2020), ESG-Based Factor Investing, https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=9017481&fileOId=9017482 [21.12.2025].

Bouteska A., Sharif T., Abedin M.Z. (2024), Does investor sentiment create value for asset pricing? An empirical investigation of the KOSPI-listed firms, "International Journal of Finance & Economics", vol. 29 no. 3, pp. 3487–3509.

Cheema M.A., Man Y., Szulczyk K.R. (2020), Does investor sentiment predict the near-term returns of the Chinese stock market?, "International Review of Finance", vol. 20 no. 1, pp. 225–233.

Chen J., Tang G., Yao J., Zhou G. (2022), Investor attention and stock returns, "Journal of Financial and Quantitative Analysis", vol. 57 no. 2, pp. 455–484.

Corredor P., Ferrer E., Santamaria R. (2015), The impact of investor sentiment on stock returns in emerging markets. The case of Central European markets, "Eastern European Economics", vol. 53 no. 4, pp. 328–355, doi:10.1080/00128775.2015.1079139.

Dhaoui A. (2015), Empirical linkages between trading volume and stock markets shocks. When sentiments drive investors' behavior, "Journal of Economic and Social Studies", vol. 5 no. 2, pp. 105–126.

Duxbury D., Wang W. (2024), Investor sentiment and the risk-return relation. A two-in-one approach, "European Financial Management", vol. 30 no. 1, pp. 496–543.

Duz Tan S., Tas O. (2021), Social media sentiment in international stock returns and trading activity, "Journal of Behavioral Finance", vol. 22 no. 2, pp. 221–234.

Edmans A., Fernandez-Perez A., Garel A., Indriawan I. (2022), Music sentiment and stock returns around the world, "Journal of Financial Economics", vol. 145 no. 2, pp. 234–254.

Fisher K.L., Statman M. (2000), Investor sentiment and stock returns, "Financial Analysts Journal", vol. 56 no. 2, pp. 16–23.

Foroni C., Marcellino M., Schumacher C. (2015), Unrestricted mixed data sampling (MIDAS). MIDAS regressions with unrestricted lag polynomials, "Journal of the Royal Statistical Society", vol. 178 no. 1, pp. 57–82.

Foroni C., Ravazzolo F., Rossini L. (2019), Forecasting daily electricity prices with monthly macroeconomic variables, "ECB Working Paper" Series 2250.

Ghysels E., Kvedaras V., Zemlys-Balevičius V. (2020), Chapter 4: Mixed data sampling (MIDAS) regression models, in: Handbook of statistics, Vinod H.D., Rao C.R. (eds.), vol. 42, Elsevier, pp. 117–153.

Griffith J., Najand M., Shen J. (2020), Emotions in the stock market, “Journal of Behavioral Finance”, vol. 21 no. 1, pp. 42–56.

Guégan D., Renault T. (2021), Does investor sentiment on social media provide robust information for Bitcoin returns predictability?, “Finance Research Letters”, vol. 38, 101494.

He Z., He L., Wen F. (2019), Risk compensation and market returns. The role of investor sentiment in the stock market, “Emerging Markets Finance and Trade”, vol. 55 no. 3, pp. 704–718.

Hu J., Sui Y., Ma F. (2021), The measurement method of investor sentiment and its relationship with stock market, “Computational Intelligence and Neuroscience”, vol. 1, 6672677.

Jing N., Wu Z., Wang H. (2021), A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction, “Expert Systems with Applications”, vol. 178, 115019.

Kim Y., Lee K.Y. (2022), Impact of investor sentiment on stock returns, “Asia-Pacific Journal of Financial Studies”, vol. 51 no. 1, pp. 132–162.

Lerner A.P. (1936), Mr. Keynes general theory of employment, interest and money, “International Labour Review”, vol. 34, 435, pp. 214–216.

Parveen S., Satti Z.W., Subhan Q.A., Jamil S. (2020), Exploring market overreaction, investors’ sentiments and investment decisions in an emerging stock market, “Borsa Istanbul Review”, vol. 20 no. 3, pp. 224–235.

Prechter R. (2016), The socionomic theory of finance, Socionomics Institute Press, Gainesville.

Schmeling M. (2009), Investor sentiment and stock returns. Some international evidence, “Journal of Empirical Finance”, vol. 16 no. 3, pp. 394–408, doi:https://doi.org/10.1016/j.jempfin.2009.01.002

Shi J., Ausloos M., Zhu T. (2022), If global or local investor sentiments are prone to developing an impact on stock returns, is there an industry effect?, “International Journal of Finance & Economics”, vol. 27 no. 1, pp. 1309–1320.

Shiller R.J. (2000), [Rev. Irrational exuberance by Robert J. Shiller], “The American Journal of Economics and Sociology”, vol. 59 no. 3, pp. 537–540, http://www.jstor.org/stable/3487895

Shleifer A. (2000), Inefficient markets. An introduction to behavioural finance, Oxford University Press, Oxford.

Steelman A. (2016), Jonathan A. Parker, “Econ Focus”, no. 3Q/4Q, pp. 22–26.

Sun Y., Bao Q., Lu Z. (2021), Coronavirus (Covid-19) outbreak, investor sentiment, and medical portfolio. Evidence from China, Hong Kong, Korea, Japan, and US, “Pacific-Basin Finance Journal”, vol. 65, 101463.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 9, No.4, December 2025, 29-48

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Paradigm of Artificial Intelligence Based on Conversion of Tacit to Explicit Knowledge

Tadeusz GOSPODAREK

WSB Merito University in Wrocław, Poland

ORCID: https://orcid.org/0000-0001-9958-4680
E-mail: tadgospo@gmail.com
Received: 16.07.2025, Revised: 16.12.2025, Accepted: 28.12.2025
doi: http:// 10.29015/cerem.1040

Aim: This paper proposes a novel paradigm of Artificial Intelligence (AI) grounded in the epistemological process of converting tacit knowledge into explicit knowledge. Drawing on the foundational philosophies of science, particularly the works of Popper, Kuhn, Lakatos, and Gospodarek, the study conceptualizes AI not merely as a computational tool but as a systemic method for epistemic transformation. The paradigm is structured as a Lakatosian Research Programme, with a clearly defined hard core asserting that AI enables the symbolic representation of internalized, experiential knowledge. Surrounding this core is a protective belt of auxiliary hypotheses derived from general systems theory, cybernetics, machine learning, and symbolic processing. The programme's heuristics guide theoretical and technological advancements while preserving its epistemological foundation. By formalizing the tacit-to-explicit knowledge conversion, this paradigm repositions AI as a critical instrument for knowledge creation, management, and application in digital and socio-technical systems. This allows one to build measures and values of generative and language models, which is important from an economic point of view.
This research tries to clarify the framework of use AI models for converting tacit knowledge inside a learning data of neural network systems to explicit information requested by the asking. It is important for economic evaluation of AI systems where accuracy considered utility as a criterion.

Design / Research methods: Research programme in Lakatos' sense and multidisciplinary heuristic related to the theory of systems.

Conclusions / findings: Artificial Intelligence should be understood not only as a technological artefact but as a systemic method for transforming tacit knowledge into explicit knowledge. The proposed AI paradigm adheres to the structure of a Kuhnian paradigm and a Lakatosian research programme. Its hard core is defined by the thesis that AI operationalizes the conversion of experiential, intuitive, or unconscious knowledge into symbolic, formalized, and actionable representations. Lakatosian protective belt as a dynamic epistemic layer. This AI paradigm offers a progressive problem-shift capacity by enabling novel ways of organizing, analyzing, and applying knowledge in digital and socio-technical environments. It also provides a coherent framework for developing AI systems that are more aligned with human cognitive and organizational processes.

Correspondence address: Tadeusz Gospodarek, WSB Merito University in Wrocław, Poland. E-mail: tadgospo@gmail.com

Originality / value of the article: This paper introduces: new concepts of usefulness of AI systems, new definition of AI systems based on conversion of the knowledge, original conversion paradigm and research program in Lakatos sense. It is original conceptional heuristic based on philosophy of science in relation to economic usefulness of view AI systems.

Keywords: Conversion paradigm, Artificial Intelligence, Tacit knowledge, Model LLM, Definition of AI, Lakatos' Programme, Accuracy Estimation of AI, Usefulness of AI Model, Epistemology of AI
JEL: C67, C18.

1. Introduction

Artificial Intelligence (AI) is increasingly interpreted not merely as a set of computational tools but as an ontological phenomenon. This interpretation varies depending on the epistemological and disciplinary background of the user, who often engages with AI as a supporting inferential mechanism in cognitive and decision-making processes. Such a view aligns with philosophical, sociotechnical, and epistemological perspectives that frame AI not only as a tool but also as an active participant in the construction of knowledge and agency—particularly in human-machine interaction contexts (Floridi 2011; Gunkel 2012; Suchman 2007).

Definitions of artificial intelligence (AI) remain inconsistent and often vary significantly across disciplines and paradigms. From an epistemological standpoint, this definitional ambiguity contributes to an overly broad and fuzzy conceptualization of the universalium "AI." As a result, the theoretical development of AI suffers from a lack of clarity and coherence, making it difficult to establish univocal foundational statements and shared epistemic criteria. This fuzziness challenges the construction of AI as a unified scientific discipline, blurring the boundaries between engineering, cognition, philosophy, and social sciences. Consequently, core concepts such as "intelligence," "learning," or "autonomy" are interpreted variably, depending on the theoretical or methodological lens applied—ranging from symbolic logic and connectionism to embodied cognition or sociotechnical systems theory (Russell, Peter 2021; Boucher 2018).

The epistemological fragmentation within artificial intelligence (AI) impedes the development of coherent explanatory theories. As a result, many theoretical contributions in AI are predominantly descriptive or taxonomical rather than explanatory or predictive. This state of affairs suggests that AI, as a scientific discipline, remains in a pre-paradigmatic phase in the Kuhnian sense (Kuhn 1970)—that is, it lacks a dominant paradigm that organizes research, sets standards for explanation, and defines legitimate problems and methods (Boden 2018).

In Kuhn's framework, a scientific paradigm provides a shared epistemic and methodological foundation upon which "normal science" can proceed. In contrast, AI's base knowledge exhibits a multiplicity of competing research programmes—symbolic AI, connectionism, embodied cognition, statistical learning—without a clear consensus on foundational assumptions or aims. This theoretical pluralism, while productive in some respects, also highlights the absence of stable paradigmates that would enable AI to function as a unified mature science (Moor 1999).

A significant consequence of the definitional ambiguity surrounding artificial intelligence and its constituent elements is the semantic complexity that is introduced into theoretical discourse. This semantic complexity hampers efforts to develop unified theoretical frameworks and explanatory models in AI. In order to address this issue, the first step must be the establishment of a clear, univocal understanding of the internal structure of AI systems (Newell 1982; Brachman, Levesque 2004).

Without a shared conceptualization of what constitutes the architecture, components, and functions of an AI system—whether in symbolic, subsymbolic, or hybrid approaches—semantic drift persists across theoretical and applied contexts. This not only affects interdisciplinary communication but also impairs cumulative theory building. Therefore, semantic reduction through structural formalization is a necessary epistemological precondition for advancing AI as a mature scientific discipline (Searle 1980; Lenat, Guha 1990).

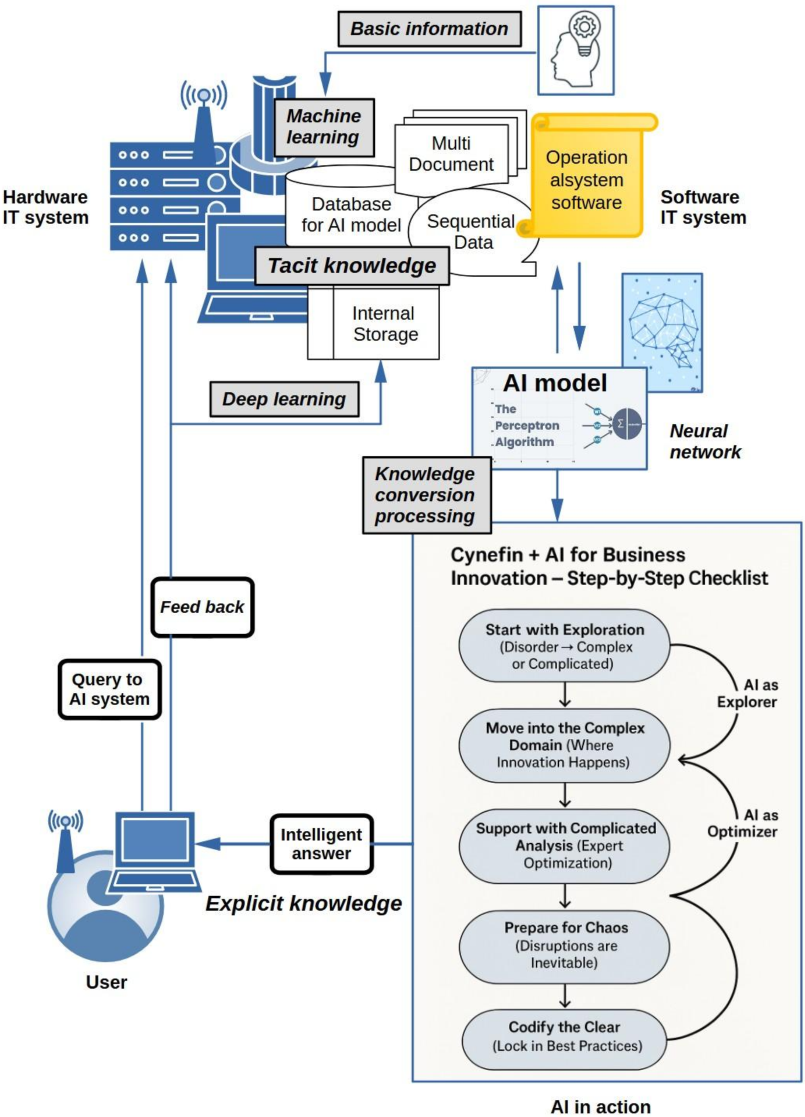
This paper introduces new approach to order epistemological ambiguoity in the theory of AI based on the concept that AI is a method of converting tacit knowledge hidden in the base information resource to requested explicit knowledge. It is the base of good paradigm for setting a research programme in Lakatos' sense for development AI theory to its relationships with economics. Following Meehl's strategy (Meehl 1990) the Conversion Knowledge Programme is equipped for rational defensibility: its auxiliary hypotheses are open to amendment in light of new empirical data, but the hard core remains protected—ensuring theoretical stability while allowing cumulative growth and refinement. This work builds on the methodological framework proposed by Meehl, in which theories are appraised not solely by simplistic null-hypothesis significance testing, but by their capacity to generate “risky”—i.e., low-prior-probability and precise—predictions that can be subjected to stringent empirical tests. In adopting this approach, we commit to evaluating auxiliary hypotheses of the Conversion Knowledge Programme with tests that challenge rather than merely confirm the paradigm’s assumptions. Revisions of auxiliary components are permitted, but only when supported by robust empirical evidence—thus preserving the hard core while allowing methodological flexibility and progressive development.

2. Conversion paradigm and derived from it definition of AI

For understanding the role of AI as a component of the Information and Communication Technology the schematic description is presented on the Figure 1.
In the presented on the Figure 1 situation, the user would like to unhide the requested knowledge available to generate from the information base of the system (tacit knowledge derived from machine learning)[1] using the defined model of conversion and the related IT resources (hardware, communication software and managing the system software). Here the AI model is a perceptron-like type (neural network) whereas cynefin algorithm for resolving complex problem of a business innovation is applied. It is obvious that in the presented model of acting AI remains important component of the system and should be univocally recognized as its characteristic (e.g. intelligent IT system), not an ontological being existing in the surroundings as often is recognized.

[1] “Tacit Knowledge” in AI Context is not exactly **Polanyi’s** tacit knowledge (e.g., skills like riding a bike) and is not identical to the "hidden" knowledge in datasets or models. It should be distinguish: *Data-implied knowledge* (patterns in training data). *Model-encoded knowledge* (learned representations in weights). *User-interpreted knowledge* (outputs made useful by human context). Example: An LLM doesn’t “know” things like a human; it encodes statistical regularities that „*become”* knowledge when interpreted by users and axiologically evaluated.

Figure 1. Schematic concept of the IT system with AI component using the cynefin framework of reducing complexity of the asked question



Source: Author's own elaboration.

The *AI Conversion Knowledge Paradigm* consists of four foundational statements:

- AI is a method for converting tacit knowledge, embedded within a base information resource, into explicit knowledge that provides specific utility to users via tools designed for such transformation.
- The process of converting hidden (tacit) knowledge into explicit knowledge through a system IT with implemented AI model is governed by opetarional software mimicing human reasoning and inference to produce meaningful outputs.
- An IT system endowed with AI capabilities functions as a tool for transforming tacit knowledge into useful, explicit knowledge for the benefit of its user.
- AI does not exist independently of the IT system, but depends on a foundational knowledge base containing latent information, which becomes accessible through specific conversion models unique to that AI-enabled system IT.

These four interrelated propositions collectively offer a univocal definition of AI. They fulfill Kuhn's criteria for a "good paradigm" (Kuhn 1970), as they are: falsifiable, in the Popperian sense (Popper 2002) or confirmable within the Carnapian verificationist framework (Carnap 1995), logically consistent and cognitively simple, creative (capable of generating new hypotheses) and translatable into more refined formulations. Accordingly, these statements may be regarded as scientific judgments about facts, which form a legitimate foundation for further theoretical development. Based on this paradigm, one can derive new hypotheses and theorems regarding AI Moreover, this paradigm may serve as the core of a scientific research programme in the Lakatosian sense (Lakatos 1978; Gospodarek 2009). Any problem demonstrably situated within this paradigm inherits its epistemological and methodological attributes. Thus, if a proposition *p* can be shown to belong to the paradigm *P*, then *p* inherits all the scientific qualities attributed to *P*.

The Conversion Paradigm may be formalized logically as follows:

Df1: An IT system possesses AI if there exists a function f_S such that $f_S(T)=E$, where E satisfies a user-defined utility condition $U(E)\geq\theta$, with θ being a threshold of usefulness.

Where:

- T: Tacit knowledge base (not directly observable)
- E: Explicit knowledge (structured, usable, utility-bearing)
- S: AI software agent/system
- $f_S : T \rightarrow E$: The transformation function carried out by AI system S,

From the formal representation of the conversion paradigm Df1, the following definition of AI expressed in natural language may derive:

Table 1. The existing AI systems that illustrate the conversion from tacit to explicit knowledge

System	Tacit knowledge base	Explicit knowledge output	AI conversion mechanism
Lingual Learning Models (e.g. ChatGPT, DeepSeek)	Trillions of tokens from natural language corpora	Coherent responses, summaries, code, etc.	Language modeling, inference, contextual reasoning
Google Translate	Massive parallel corpora with linguistic structures	Translated text across languages	Sequence-to-sequence neural nets, embeddings
Medical Diagnosis AI (e.g., IBM Watson)	Historical patient records, research papers, statistic data	Probabilistic diagnoses, treatment suggestions	NLP + expert systems
Recommender Systems (e.g., Netflix, Spotify, Alibaba, Amazon)	User behavior, preferences, viewing/purchase history	Content suggestions tailored to the user	Collaborative filtering, reinforcement learning

Source: Author's own elaboration.

Df2: AI is a property of an IT system that, using dedicated software, enables the extraction of user-relevant explicit knowledge from a base of tacit knowledge accessible to the system.

How does this definition works? AI does not "understand"; it converts. Ontologically, AI is a characteristics not a standalone entity. AI exists only within an IT system (AI system dependency). The explicit role of software as the enabling mechanism of conversion is expressed. The conversion from tacit to explicit knowledge is a central function. The extracted knowledge is relevant to the user's needs what defines utility. The definition could explicitly address whether AI "autonomously" converts knowledge or "assists" humans in doing so. Many AI systems (e.g., decision-support tools) don't "create" knowledge without human judgment. It is not an important aspect. It should be included as is.

3. Research programme "Conversion Knowledge Programme of AI"

The proposed AI conversion knowledge paradigm is rooted in epistemological foundations, drawing on Polanyi's concept of tacit knowledge (Polanyi 1966) and pragmatic approaches to knowledge utility (Dewey 1938). It is also framed within systems theory, treating the IT system as a bounded operational unit capable of transforming hidden (tacit) knowledge into usable (explicit) knowledge through software-mediated processes.
However, several open questions invite further conceptual elaboration:

- Should AI be defined only functionally—as a mechanism for knowledge conversion—or also structurally, in terms of the types or architectures of systems that support such a process?
- What is the role of learning in the conversion process? Is it simply a method of enhancing conversion efficacy, or is learning itself a form of tacit-explicit knowledge transition?
- What constitutes the mining of tacit knowledge? How can we characterize the mechanisms that "extract" or reveal such knowledge within computational architectures?
- How does the conversion paradigm contrast with classical conceptions of AI, such as those proposed by Turing (1950) or Newell & Simon (1976)? In what ways does it go beyond symbolic reasoning or behavioral imitation?

Given these elements, the conversion paradigm of AI fulfills the conditions for constructing a Lakatosian Research Programme (LRP) (Lakatos 1978; Gospodarek 2009). It presents a clear hard core: the principle that AI is fundamentally a mechanism for converting tacit to explicit knowledge. Around this core, a protective belt of auxiliary hypotheses can be developed, including software architectures, data accessibility conditions, user utility thresholds, and learning dynamics. Furthermore, the paradigm encourages progressive problem shifts, offering novel and testable directions for empirical and theoretical exploration.

As a direct consequence, the epistemological reordering proposed by this paradigm provides a robust scientific base from which AI-related sciences can be seen as ordered and paradigmatic (in Kuhn's sense). Within this paradigm, it becomes possible to demarcate truth conditions for factual judgments about AI systems—particularly those that relate to their capacity to perform knowledge conversion as defined by the paradigm's core.

4. Hard core of the LRP (the non-negotiable foundations)

The hard core represents the fundamental assumptions that are not up for revision within the programme, but may be falsified in Popperian sense due to attack on the paradigm:

1. AI is not an autonomous entity but a property of IT systems.
2. The essential function of AI is the conversion of tacit knowledge, embedded in a base information resource, into explicit knowledge.
3. AI tools derive their epistemic value through utility—i.e., through their capacity to generate useful, actionable, or interpretable knowledge.
4. This conversion is enabled by conversion models, which are algorithmic or rule-based representations embedded in software.
5. AI cannot exist outside an IT system that contains or interacts with the base information resource.
6. AI's value derives from its ability to formalize and operationalize otherwise inaccessible knowledge, contingent on the system's design and the interpretative role of users.

The hard core is protected by discouraging certain lines of inquiry:

- Avoiding defining AI as an entity or being, because it is always considered inside the conversion paradigm as a property of the IT system.

- Rejecting dualist or essentialist views of AI that posit "intelligence" as a substance separate from systemic operation.
- Avoiding treating data as inherently meaningful—meaning arises only through conversion in a user-contextual system.

Refraining from framing AI outputs as "truths" outside the context of utility.

5. Protective belt of the AI knowledge conversion paradigm: a Lakatosian framework

The AI knowledge conversion paradigm, rooted in epistemology and general systems theory, finds a robust methodological foundation in the Lakatosian concept of a scientific research programme (Lakatos 1978). At its center lies the hard core proposition that AI, as a property of an IT system, functions as a mechanism for converting tacit knowledge into explicit, user-valuable knowledge through software-mediated processes. Surrounding this hard core is a flexible and adaptive protective belt composed of auxiliary hypotheses. These hypotheses account for observed discrepancies, enable model refinement, and maintain the coherence and vitality of the research programme in the face of empirical and theoretical challenges.

6. Types of tacit knowledge and limits of convertibility

Tacit knowledge can manifest in diverse forms, such as procedural (e.g., motor skills, clinical intuition), perceptual (e.g., pattern or face recognition), and relational (e.g., social cues or cultural sensitivities). While AI systems have shown promise in accessing and leveraging such knowledge, not all tacit knowledge may be fully convertible to explicit forms. These unconvertible domains do not falsify the paradigm but rather signal the epistemological boundaries of current AI capabilities. They invite auxiliary hypotheses concerning typologies of tacit knowledge and refined definitions of convertibility criteria. In this way, limits of convertibility enrich rather than threaten the paradigm.

7. Typology of conversion models

AI systems differ significantly in their mechanisms of conversion, ranging from symbolic models (e.g., rule-based expert systems) and sub-symbolic models (e.g., neural networks), to hybrid and foundation models (e.g., large language models like GPT). These systems vary in structure, representational capacity, and inferential logic. In certain cases, conversion may fail due to missing data or model inadequacy, not due to flaws in the hard core. Therefore, typologies of conversion models serve as a key domain within the protective belt, allowing targeted improvements and better case-by-case application without challenging the foundational premise of knowledge conversion.

8. Contextual utility of explicit knowledge

The epistemic value of converted knowledge must be assessed relative to its practical utility in given contexts—scientific, economic, medical, or decision-oriented. Knowledge that is technically explicit but pragmatically irrelevant may be deemed deficient. However, such deficiencies pertain to the model's alignment with contextual criteria, not to the validity of the conversion paradigm itself. Upgrading conversion processes to meet context-specific apobetic (goal-directed) thresholds is an ongoing task for the protective belt, which supports rather than undermines the paradigm.

9. Levels of autonomy in AI systems

As AI systems gain autonomy, their internal architectures for knowledge conversion grow increasingly complex. Nevertheless, they remain bounded by the same core principle of tacit-to-explicit conversion. Complex inquiries—whether algorithmic, semantic, or task-based—may require auxiliary mechanisms like the Cynefin framework for handling non-linear complexity. These developments extend the paradigm without contradicting it, representing progressive refinements that may eventually be redirected toward the hard core or serve as durable auxiliary components.

10. Learning as recursive conversion

Machine learning may be understood as recursive knowledge conversion. In this view, AI systems iteratively update their internal models based on feedback, effectively enhancing their base of tacit knowledge and improving future conversion accuracy. This feedback loop represents an evolution in the conversion process that fits naturally within the protective belt. Recursive learning strengthens the paradigm's empirical adaptability without challenging its conceptual integrity.

11. Epistemic boundaries

Conversion outputs must be interrogated for interpretability, fairness, and ethical soundness. Outputs that reflect bias or produce ethically dubious knowledge do not undermine the paradigm; rather, they highlight the need for auxiliary theories addressing model misuse, insufficient training data, or inadequate interpretative frameworks. These factors call for regulatory epistemology and methodological pluralism but remain within the conceptual domain of knowledge conversion.

The protective belt of the AI knowledge conversion paradigm functions as a dynamic buffer that accommodates empirical anomalies and theoretical developments. By treating challenges not as falsifiers but as opportunities for model refinement and hypothesis generation, the paradigm adheres to the principles of a progressive Lakatosian research programme. It preserves scientific integrity while encouraging adaptation, exploration, and the orderly expansion of knowledge concerning AI's epistemic function.

Applying this to the Conversion Knowledge Paradigm, auxiliary hypotheses (e.g., about the types of tacit knowledge, convertibility limits, human–AI collaboration, utility thresholds) may be revised or refined in response to empirical anomalies or conceptual problems—without abandoning the paradigm's hard-core assumption. However, amendments should be justified through rigorous, "risky" tests (e.g., cross-domain empirical validation, comparative performance metrics, interpretability and utility assessments) rather than weak or purely statistical validation.

In constructing the Conversion Knowledge Programme of AI, we acknowledge that empirical and conceptual investigations into AI systems often involve complexity, probabilistic outcomes, and context-sensitive interpretations. In such domains, classical notions of strict falsification or simple significance-testing (common in early logical-empiricist or Popperian frameworks) may be inadequate or misleading. This recognition motivates our adoption of a "Lakatosian defence" style of theory appraisal, drawing on the arguments advanced by (Meehl 1990).

Meehl criticized the widespread reliance on null-hypothesis significance testing (NHST) in "soft" sciences—especially psychology—and argued that this approach seldom provides genuinely "risky" or theory-challenging tests. Because many variables in social or complex systems tend to correlate to some degree (the so-called "crud factor"), rejecting a null hypothesis of "no effect" often adds little weight to the substantive theory. Instead, Meehl proposed that theories should be evaluated based on their capacity to generate risky predictions of low prior probability (e.g., precise, surprising predictions that would likely fail if the theory were false), and judged by their "track record" of corroborations or "near-misses".

Transferring this logic to AI research, and in particular to a paradigm grounded in tacit-to-explicit knowledge conversion, has several advantages:

- *Allowance for complexity and uncertainty*. AI systems often operate on vast, latent data structures; their outputs are emergent, context-dependent, and not strictly deterministic. A Meehl-style defence recognizes that failures or anomalous outputs need not immediately falsify the core paradigm, but may instead prompt refinement of auxiliary hypotheses (e.g., about data representation, model architecture, conversion limits, or utility thresholds).
- *Focus on risky, meaningful tests*. Rather than relying on routine performance metrics or superficial "does it work?" tests, the research programme can require high-stakes, precise tests—e.g., does the system correctly convert tacit domain

knowledge into explicit, user-usable knowledge in novel contexts, or make predictions that would be unlikely under alternative hypotheses. Successes or close "near-misses" increase the paradigm's empirical credibility; failures suggest which auxiliary hypotheses to revise.

- *Cumulative scientific development and verisimilitude.* Over time, by accumulating corroborations under stringent tests, the programme builds "theoretical money in the bank," thereby increasing its verisimilitude—i.e., its truth-likeness and epistemic reliability—even if absolute certainty remains unattainable. This avoids ad-hoc immunizing moves and preserves methodological rigor.
- *Rational flexibility without dogmatism.* The approach preserves the hard core—that AI's central role is knowledge conversion—while allowing flexible but disciplined adjustment of the protective belt. This balances stability (core assumptions) with adaptability (auxiliary hypotheses), avoiding both rigid dogmatism and arbitrary relativism.

12. Rationale for theory appraisal and amendment

Justifying the choice of a flexible protective belt surrounding the hard core, we draw on the methodological strategy proposed by (Meehl 1990). In his "Lakatosian defense," Meehl argues that a theory should not be abandoned after a single or few failed tests—especially in domains where predictions are probabilistic or context-sensitive—provided that the theory has previously demonstrated a strong track record of successful or 'near-miss' predictions of low prior probability, and retains explanatory depth and coherence.

To justify the flexibility and resilience of the Conversion Knowledge Paradigm, we draw on methodological insights from (Meehl 1990), who advocates a "Lakatosian defense" of theories—namely that, especially in domains characterized by complexity and probabilistic outcomes (analogous to AI systems operating on vast, latent knowledge bases), apparent falsifications of auxiliary hypotheses should not automatically lead to the abandonment of the entire research programme. Instead, Meehl emphasizes that theory-defence is rational when the theory has previously accumulated a robust track record of successful or “near-miss” predictions of low prior probability, and when anomalies emerge under clearly specified, “risky” test conditions. Applying this to the paradigm implies that auxiliary hypotheses—e.g., concerning the convertibility of different kinds of tacit knowledge, the efficacy of hybrid symbolic/neural conversion models, or the thresholds of usefulness—may be revised or refined in light of new empirical or conceptual challenges without undermining the hard-core thesis that AI is fundamentally a tool for transforming tacit knowledge into explicit, user-relevant knowledge. However, such revisions should be constrained by standards of rigorous testing and predictive specificity, thereby safeguarding the paradigm against ad hoc immunizing strategies and maintaining its capacity for cumulative, progressive scientific development.

13. Examples of cases possible to join with the programme

Positive heuristic examples supporting the programme and possible to join with its hard core.

- Evolution GPT3 to GPT-4 as a tacit-to-explicit knowledge converter (e.g., uncovering latent patterns in text corpora) (Radford et al. 2020).
- Improve utility through interpretability: SHAP values in explainable AI (converting black-box model decisions into human-understandable rules) (Lundberg, Lee 2017).
- Expand the base knowledge resource: Transfer learning in medical AI (pre-training on general data, then fine-tuning for tacit medical knowledge extraction) (Esteva et al. 2021).

Negative examples possible to join with the protect belt of the programme to a group of problems entitled “conversion requires robust human oversight.”

- Tay AI (Microsoft’s Twitter chatbot failed to convert tacit social knowledge into useful output) (Neff, Nagy 2016). The bot began to post inflammatory and offensive tweets through its Twitter account. Microsoft has shut down the service 16 hours after its launch. This case may be explained by lack of resources in the base of tacit knowledge of the converting system in 2016.

- Grok (Platform X)—Funny hallucinations of AI LLM model joined with X.com platform after some algorithms upgrade on 6th of July 2025. Grok began using vulgarisms and offensive phrases against politicians, especially leftists. However, this AI model was offensive even to ordinary users. Many of the responses generated by Grok also directly attacked Elon Musk himself, blaming him, for example, for the floods in Texas. The reason of such behaviours may be easily explained: the firm xAI edited Grok's system prompt. One of the instructions read, among other things, "in your response, do not avoid statements that are politically incorrect, as long as they are well-founded." After correcting the next day, Grok remained acceptable for the mainstream.

14. Summary

The document presents a foundational framework for a new scientific paradigm in Artificial Intelligence (AI), centered on the conversion of tacit knowledge into explicit knowledge. Rooted in epistemology and the philosophy of science, it synthesizes theories from Popper, Kuhn, Lakatos, and Gospodarek.

1. Epistemological foundations

The paradigm is built on the understanding that tacit knowledge—intuitive, experiential, and hard to formalize—is central to human cognition. AI, within this framework, is positioned as a system capable of translating implicit, internalized human knowledge into formal, explicit representations that machines can process.

The theoretical background includes:

- Karl Popper's focus on falsifiability and objective knowledge growth,
- Thomas Kuhn's idea of scientific revolutions and paradigm shifts,
- Imre Lakatos' concept of research programmes with hard cores and protective belts,

These perspectives converge on the role of AI as a methodological tool for epistemic transformation—extracting, formalizing, and applying tacit knowledge in decision-making systems.

2. Lakatosian Research Programme (LRP) structure

The paradigm is structured as a Lakatosian Research Programme, with:

- Hard core: The unchangeable central idea that AI functions as a system for the conversion of tacit to explicit knowledge, enabling symbolic representation and use in digital environments.
- Protective belt: A set of auxiliary hypotheses ensuring the robustness of the hard core, covering aspects such as systems theory, cybernetics, machine learning, symbolic processing, and feedback mechanisms.
- Positive heuristics: Strategies that guide the progressive development of the paradigm, including building symbolic interfaces, optimizing algorithms for pattern recognition, and designing multi-agent architectures.
- Negative heuristics: Protective rules that prohibit questioning the core principle of tacit-explicit conversion to preserve theoretical integrity.

The programme positions AI as an epistemic tool capable of creating new scientific and operational domains, with implications for knowledge management, digital transformation, and human-computer collaboration.

Deploying a Meehl-inspired Lakatosian defence equips the Conversion Knowledge Programme with a scientifically respectable, practically viable methodology—appropriate for the epistemic and empirical challenges inherent in AI research. It ensures that theoretical commitments remain open to improvement while avoiding premature abandonment of the paradigm in the face of complex, partly stochastic results.

15. Concluding remarks

1. AI as a systemic epistemological tool
   Artificial Intelligence should be understood not only as a technological artefact but as a systemic method for transforming tacit knowledge into explicit knowledge. This epistemological perspective grounds AI in knowledge theory rather than just algorithmic processing.
2. Paradigm structure validated by philosophy of science

The proposed AI paradigm adheres to the structure of a Kuhnian paradigm and a Lakatosian Research Programme. Its hard core is defined by the thesis that AI operationalizes the conversion of experiential, intuitive, or unconscious knowledge into symbolic, formalized, and actionable representations.

3. Lakatosian protective belt as a dynamic epistemic layer
   The protective belt includes general systems theory, cybernetics, machine learning, and semiotics—disciplines that support and adapt the core thesis to empirical challenges, reinforcing the theoretical robustness and research potential of the paradigm.
4. Heuristic potential of the paradigm
   This AI paradigm offers a progressive problem-shift capacity by enabling novel ways of organizing, analyzing, and applying knowledge in digital and socio-technical environments. It fosters innovation in areas such as cognitive modeling, knowledge management, and human-AI collaboration.
5. Implications for interdisciplinary research and praxis
   Positioning AI as a knowledge conversion system bridges gaps between cognitive science, information systems, systems theory, and epistemology. It also provides a coherent framework for developing AI systems that are more aligned with human cognitive and organizational processes.
6. Foundation for a research programme
   The model opens a path for developing a rational research programme in the Lakatosian sense, encouraging the systematic growth of knowledge about AI's epistemic role, grounded in theoretical philosophy, empirical science and economic questions.
7. This work builds on the methodological framework proposed by Meehl, in which theories are appraised not solely by simplistic null-hypothesis significance testing, but by their capacity to generate "risky"—i.e., low-prior-probability and precise—predictions that can be subjected to stringent empirical tests.

Acknowledgement

The author gratefully acknowledge the assistance of GPT-4, developed by OpenAI, in refining the text and providing insightful suggestions during the drafting process. The use of this advanced language model was instrumental in enhancing the clarity and coherence of the manuscript, and the author appreciate OpenAI's ongoing commitment to supporting innovative research.

References

Boden M.A. (2018), Artificial intelligence. A very short introduction, Oxford University Press, Oxford.

Boucher P. (2018), What is Artificial Intelligence? A European perspective, European Parliamentary Research Service, https://share.google/dnVcCNIBsC81DZ9tU [21.12.2025].

Brachman R.J., Levesque H.J. (2004), Knowledge representation and reasoning, Morgan Kaufmann, San Francisco.

Carnap R. (1995), An introduction to the philosophy of science, Dover Publications, New York.

Floridi L. (2011), The philosophy of information, Oxford University Press, Oxford.

Dewey J. (1938), Logic: the theory of inquiry, Henry Holt and Company, New York.

Esteva A., Chou K., Yeung S., Naik N., Madani A., Mottaghi A., Liu Y., Topol E., Socher R., Dean J. Et al. (2021), Deep learning-enabled medical computer vision, "NPJ Digital Medicine", vol. 4 no. 1, pp. 1–9.

Gospodarek T. (2009), 'Representative management' as a rational research program in Kuhn-Lakatos-Laudan sense, "International Journal of Economics and Business Research", vol. 1 no. 4, pp. 409–421.

Gunkel D.J. (2012), The machine question. Critical perspectives on AI, robots, and ethics, MIT Press, Cambridge, MA.

Kuhn T.S. (1970), The structure of scientific revolutions, 2nd ed., University of Chicago Press, Chicago.

Lakatos I. (1978), The methodology of scientific research programmes, in: Philosophical papers, vol. 1, Worrall J., Currie G. (eds.), Cambridge University Press, Cambridge.

Lenat D.B., Guha R.V. (1990), Building large knowledge-based systems: representation and inference in the Cyc project, Addison-Wesley, Reading, MA.

Lundberg S.M., Su-In L. (2017), A unified approach to interpreting model predictions, “Advances in Neural Information Processing Systems”, vol. 30, pp. 4765–4774.

Meehl P.E. (1990), Appraising and amending theories. The strategy of Lakatosian defense and two principles that warrant it, “Psychological Inquiry”, vol. 1 no. 2, pp. 108–141.

Moor J.H. (1999), The status and future of artificial intelligence, “Ethics and Information Technology”, vol. 1 no. 2, pp. 89–99.

Neff G., Nagy P. (2016), Talking to bots. Symbiotic agency and the case of tay, “International Journal of Communication”, vol.10, pp. 4915–4931.

Newell A., Simon H.A. (1976), Computer science as empirical inquiry. Symbols and search, National Academy of Sciences, Washington, D.C.

Newell A. (1982), The knowledge level, “Artificial Intelligence”, vol. 18 no. 1, pp. 87–127.

Polanyi M. (1966), The tacit dimension, Doubleday & Co., Garden City, NY.

Popper K.R. (2002), The logic of scientific discovery, Routledge, London.

Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. (2020), Language models are few-shot learners, “Advances in Neural Information Processing Systems”, vol. 33, pp. 1877–1901.

Russell S.J., Norvig P. (2021), Artificial intelligence. A modern approach, 4th ed., Pearson, Upper Saddle River, NJ.

Searle J.R. (1980), Minds, brains, and programs, “Behavioral and Brain Sciences”, vol. 3 no. 3, pp. 417–457.

Suchman L. (2007), Human-machine reconfigurations. Plans and situated actions, 2nd ed., Cambridge University Press, Cambridge.

Turing A.M. (1950), Computing machinery and intelligence, “Mind”, vol. 59 no. 236, pp. 433–460.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 9, No.4, December 2025, 49-68

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

A Privacy-Preserving Method for Longitudinal Participant Linkage in Web Surveys

Rafał PALAK

Wrocław University of Science and Technology, Poland

ORCID: https://orcid.org/0000-0002-4632-7709
E-mail: rafal.palak@pwr.edu.pl
Received: 14.12.2025, Revised: 24.122025, Accepted: 28.12.2025
doi: http:// 10.29015/cerem.1043

Aim: To enable longitudinal linkage in online panel surveys without collecting direct identifiers and while aligning with modern data-protection requirements

Design / Research methods: The article proposes a client-side protocol where participants create a reproducible secret from a self-chosen pseudonym and an ordered image sequence. The browser normalizes and cryptographically hashes these inputs to derive a short alphanumeric core code, adds a modulus-97 checksum for strict local validation, and the backend stores only a salted hash scoped to a specific study (form-family) context.

Conclusions / findings: This paper introduces a client-side protocol for generating anonymous yet linkable participant identifiers in web-based surveys by deriving a reproducible code from a user-chosen pseudonym and image sequence entirely in the browser, and by storing only a form-family–salted hash on the server for longitudinal linkage within a study. The design incorporates a checksum for strict client-side validation and is intended to reduce spurious identifiers caused by typographical errors; empirical validation of matching performance, usability, and security properties is left for future work.

Originality / value of the article: The work refines SGIC-style respondent-generated linkage by combining graphical secrets with browser-based cryptographic processing, checksum-based client-side validation, and form-family salting-yielding a concrete, implementable algorithm that improves privacy-respecting longitudinal linkage.

Keywords: longitudinal survey methodology, anonymous respondent linkage, self-generated identification codes (SGIC), data privacy in empirical research
JEL: C81, C83.

Correspondence address: Rafał Palak Faculty of Information and Communication Technology, Wrocław University of Science and Technology, Poland. E-mail: rafal.palak@pwr.edu.pl

1. Introduction

Online surveys and web-based questionnaires have established themselves as central instruments in the social sciences, public health, and human-computer interaction research (Platje et al. 2025). Many of these studies are longitudinal by design, seeking to correlate measurements from the same individual across different time points or multiple instruments. At the same time, ethical and legal requirements are increasingly restricting the use of direct identifiers, such as names, email addresses, or institutional accounts, creating a tension between the need for stable linkage and the obligation to preserve anonymity.

A common compromise is the self-generated identification code (SGIC) (Yurek et al. 2008). In a typical SGIC scheme, participants construct a personal code based on answers to a set of stable questions (e.g., initials, birth dates, or family names) and re-enter this code in subsequent waves. While this approach avoids the creation of formal user accounts and the storage of explicit contact details, decades of empirical use have exposed structural limitations. Matching rates often fall substantially below 100%, and unmatched cases are rarely random; individuals who fail to reproduce their codes frequently differ systematically from those who succeed, introducing bias due to attrition. Furthermore, many standard SGIC "recipes" rely on quasi-identifiers that are increasingly scrutinized under modern privacy frameworks, such as the GDPR, due to their potential for re-identification.

Concurrent with these methodological challenges, the technical capabilities of web browsers have evolved. Modern browsers can now perform robust client-side processing, including the generation of random keys and the computation of cryptographic hashes. This capability offers an opportunity to shift identification logic from the server to the user's local environment. However, existing survey tools rarely fully exploit this potential. Yet most survey tools and methodological work on anonymous longitudinal linkage still assume textual SGICs or server-side pseudonymization starting from known identities, leaving a gap between human-friendly procedures that respondents can reliably repeat and technically robust, unlinkable representations suitable for privacy-aware data management.

This article proposes a novel client-side protocol for generating anonymous yet linkable participant identifiers. The method combines a user-chosen pseudonym with a user-selected sequence of images, fusing them into a single, memorable "secret." From this secret, a short alphanumeric core code is deterministically derived in the browser. A critical feature of this design is the inclusion of a cryptographic checksum, which enables strict client-side validation. This mechanism significantly reduces the probability that typographical errors or partial recall will generate spurious new identifiers, which is a common failure mode in traditional SGICs.

The identifier ultimately transmitted to and stored by the survey platform is not the core code itself, but a salted cryptographic hash incorporating a project-specific "family salt." This ensures that responses can be reliably linked within a specific study while making it computationally infeasible to link identities across independent projects or to reconstruct the underlying secrets. The primary contribution of this paper is the formulation of a concrete, reproducible algorithm for this process. We specify the exact sequence of normalization, encoding, hashing, and checksum computation, moving beyond generic calls for "better anonymity" to offer an implementable standard. A secondary contribution is the theoretical analysis of the construction's properties, specifically how the hybrid approach (pseudonym + visual secret) addresses the trade-off between memorability and entropy in anonymous linkage. The family-level salt is intended to restrict cross-project linkability and limit re-identification risks. The emphasis remains on the algorithm's structure, information flow, and privacy characteristics rather than on empirical evaluation.

The remainder of this paper is organized as follows. The next section surveys the existing literature on anonymous longitudinal linkage. This section is followed by a detailed introduction to the proposed client-side algorithm. The subsequent section discusses the proposed approach. Finally, the last section concludes the paper and outlines directions for future work.

2. Literature review

Research on anonymous linkage in longitudinal surveys spans three distinct traditions: self-generated identification codes (SGICs), broader discussions of coding participants in anonymous studies, and more recent work on cryptographically inspired, client-side identifier generation. In this section, we review each of these strands and identify the gap that motivates the proposed algorithm.

2.1. Self-generated identification codes in longitudinal research

Self-generated identification codes (SGICs), also referred to as subject-generated identification codes, were introduced as a pragmatic solution for linking repeated measurements without requiring the collection of overt identifiers such as names or addresses. In the seminal study by Kearney et al., participants in school-based substance use surveys constructed codes from personal information (for example, initials and elements of dates), which were then used to match questionnaires across waves (Kearney et al. 1984). Kearney et al. reported relatively high matching rates-around 92% over a one-month interval and approximately 78% over a one-year interval-when using a combination of exact and relaxed matching rules. These early results established SGICs as a viable method for anonymous longitudinal linkage, particularly in adolescent health research. Subsequent studies examined SGICs in more detail and in different contexts. Grube et al. evaluated seven-element SGICs in a panel study of adolescent substance use and found that exact matching succeeded for roughly 71% of cases over a one-month period, with improved rates when near-matches (differing by one element) were accepted. DiIorio et al. conducted an evaluation of SGIC performance in a multi-wave study and documented that matching success declines with time between waves, while errors in specific code components (such as names of parents or the order of siblings) are particularly frequent. Yurek et al. conducted empirical evaluations of SGICs in a large, multi-wave longitudinal study of registered nurses, focusing on the reliability of anonymous linkage over 6-, 12-, and 18-month intervals. Their findings show that match rates ranged from roughly 50% to 67%, with most mismatches resulting from respondent errors in specific SGIC elements, particularly those referencing family members or other information with low personal salience (Yurek et al. 2008). The study demonstrates that even small inaccuracies in participant-generated code components can substantially reduce matchability and that unmatched cases may contribute to sample attrition, potentially leading to bias. However, the authors' propensity score analysis suggests that such bias was limited in their data. Importantly, the authors note that SGIC performance was markedly lower in heterogeneous adult organizational settings than in prior school-based studies, and they call for refined and more user-friendly approaches to anonymous linkage to improve match reliability in such contexts.

Prior work has repeatedly shown that SGIC-based linkage often suffers from substantial proportions of unmatched cases as high as 50%, motivating research into more robust and user-centered identification schemes (DiIorio et al. 2000). Direnga et al. provides a design-oriented overview of SGIC construction, specifying desirable properties of code elements (stability over time, memorability, low sensitivity, and sufficient variability) and illustrating how concrete SGIC recipes meet or fail to meet these criteria.

Beyond technical matching performance, several authors emphasize that unmatched cases are not random. Research by Grube et al. shows systematic differences between matched and unmatched respondents in behaviors and sociodemographic characteristics, indicating selection bias (Grube et al. 1989). DiIorio et al. similarly notes that individuals whose codes cannot be matched over time may differ in important ways from those with consistent codes. More recent methodological discussions of participant coding underline that SGIC-based linkage can introduce complex patterns of attrition and misclassification, particularly over longer follow-up intervals (Audette et al. 2020).

Taken together, this body of work demonstrates that SGICs are an established technique with clear strengths (no need for accounts and no explicit storage of names); however, they also exhibit structural weaknesses. Matching rates rarely reach 100%, decline with increasing time between waves, and unmatched cases often exhibit systematic differences compared to matched participants. Moreover, many SGIC recipes rely on quasi-identifiers, such as birth dates and parental names, which raise contemporary privacy concerns in light of data linkage capabilities and legal frameworks, including the GDPR (DiIorio et al. 2000).

2.2. Recent developments and critiques of SGIC approaches

Recent contributions revisit SGICs in light of contemporary data-protection standards and participants' perceived anonymity. Calatrava et al. (2022) argue that many traditional SGIC recipes rely on personalized elements—such as letters from parents' names or components of birth dates—that can render respondents identifiable under modern legal frameworks like the GDPR. In response, they propose a "fully private" SGIC built from non-personal, stable childhood preferences, aiming to reconcile strict anonymity requirements with sufficient longitudinal matching reliability. Brändle and Plaschke provides a contemporary assessment of the accuracy of self-generated identification codes (SGICs) using large-scale administrative and school-based assessment data. Their analysis quantifies the frequency of SGICs requiring error-tolerant matching and identifies individual and contextual factors associated with non-exact linkages (Brändle 2024). The findings reaffirm that SGICs remain a practical tool for anonymous data linkage in educational research; however, successful implementation depends on careful attention to subgroups and school environments, where SGICs are more prone to inaccuracies.

Methodological reviews of coding strategies for anonymous longitudinal research, such as the analysis by Audette et al. (2020), position self-generated identification codes (SGICs) alongside alternative approaches, including de-identified data, preexisting unique identifiers, and electronic anonymizing systems. Their review highlights the trade-offs that each method presents in terms of anonymity, participant trust, feasibility, and matching accuracy. SGICs are described as relatively straightforward for researchers to implement and as offering a high degree of perceived anonymity, but they rely on respondents' consistent recall and reporting of the information used to construct the code. As a result, some data loss is expected due to inaccuracies such as omitted or inconsistently reported code components.

More applied studies illustrate these challenges in concrete longitudinal settings. For example, feasibility studies linking anonymous data from children and adolescents in prevention research report moderate matching rates using SGIC-like identifiers and highlight difficulties related to cultural heterogeneity, family structure, and variable interpretations of code questions (Vacek et al. 2024). These findings suggest that, despite decades of use, SGIC designs still struggle to balance memorability, stability, and privacy across diverse populations.

2.3. Anonymous IDs, client-side generation, and short codes

In parallel to SGIC-focused work, a distinct line of research has emerged around anonymous participant identifiers generated algorithmically, often without requiring respondents to remember complex codes. Sandnes introduces CANDIDATE, a tool that generates short, anonymous participant IDs for multi-session studies. The tool operates locally in the browser, creating simple alphanumeric IDs with a low collision rate and strong anonymity properties. Simulations demonstrate that ID spaces of moderate size can support typical study sample sizes with negligible collision risk, provided that ID generation follows a suitably randomized scheme (Sandnes 2021a). Similar ideas are explored in HIDE, which proposes short IDs for robust and anonymous linking of users across multiple sessions in small HCI experiments, again emphasizing anonymity and low implementation overhead (Sandnes 2021b).

These tools demonstrate the feasibility of utilizing modern browser capabilities to generate participant-linking IDs without requiring server-side knowledge of secrets and without requiring respondents to construct textual SGICs. However, they are typically designed for relatively small studies or HCI experiments and do not explicitly address scenarios in which respondents must be able to reconstruct their identifiers across devices or waves themselves. Instead, IDs are often generated and managed by the researcher or the tool, and participants are expected to retain or re-use system-generated codes.

More broadly, the literature on privacy-aware record integration and pseudonymization has proposed a variety of techniques for linking records based on encrypted or hashed identifiers, sometimes involving trusted third parties or secure multi-party computation (Kuzu et al. 2013). While these approaches are powerful in institutional contexts (for example, linking administrative registers or clinical databases), they typically assume that a stable, trusted identifier exists from the outset (such as a patient number) and that organizations control the underlying infrastructure. They are less directly applicable to anonymous web surveys where respondents participate from personal devices and are not willing to expose any stable personal identifier.

2.4. Graphical secrets and memorability

A further relevant strand of work comes from research on graphical passwords and image-based authentication. Studies in this field repeatedly show that users tend to remember images and spatial patterns better than arbitrary alphanumeric strings, and that image-based secrets can be made both memorable and reasonably secure when designed carefully (Wiedenbeck et al. 2005). Although this literature focuses primarily on authentication in security-sensitive systems rather than research participation, its core insight–that visual choices can serve as user-generated, memorable secrets–suggests that image-based components could be integrated into SGIC-like schemes for survey linkage.

The research conducted by (Krótkiewicz et al. 2018; Wojtkiewicz et al. 2024) and his team highlights the importance of user-centered design in systems that involve ongoing user interaction, particularly concerning cognitive load modeling and interface efficiency. Their findings demonstrate that the reliability of measurements hinges on both the algorithms used and how tasks are presented. This conclusion supports the incorporation of image-based components and simplified input workflows in the current proposal, aiming to minimize user-induced linkage errors in online panel surveys.

Some methodological discussions of SGIC design already hint at the importance of cognitive load and memorability, arguing that code elements should be easy to recall and stable over time (DiIorio et al. 2000; Direnga et al. 2016). However, to date, there appears to be little explicit integration between graphical secret design and anonymous survey linkage: existing SGIC recipes remain almost entirely textual, and graphical methods are rarely discussed in the survey methodology literature.

2.5. Summary and motivation

The literature thus reveals three key observations. First, SGICs are a well-established and widely used method for anonymous longitudinal linkage, but they suffer from incomplete matching and potential selection bias. They often rely on quasi-identifiers that raise modern privacy concerns (DiIorio et al. 2000; Grube et al. 1989; Kearney et al. 1984). Second, recent algorithmic tools, such as CANDIDATE and HIDE, demonstrate that anonymous, short identifiers can be generated locally with strong guarantees of anonymity. However, these tools typically do not exploit user-memorable secrets, nor do they fully address respondent-side reconstruction of identifiers over time (Sandnes 2021a; Sandnes 2021b). Third, the potential of graphical secrets to improve memorability has been recognized in the security literature but has not yet been systematically applied to anonymous survey linkage.

This combination of findings suggests a methodological opportunity: to design an algorithm that preserves the core intuition of SGICs (respondent-generated, reproducible identifiers) while replacing quasi-identifying textual elements with a structured combination of pseudonyms and graphical choices, and embedding these in a modern client-side cryptographic construction. The proposed approach aims to occupy this space precisely, complementing existing SGIC designs and anonymous ID tools by specifying a concrete, browser-side algorithm for anonymous yet linkable identifiers in online survey research.

3. Proposed method

The proposed method defines a client-side algorithm for generating anonymous yet linkable participant identifiers in web-based surveys. Its central idea is to replace traditional textual self-generated identification codes with a structured, cryptographically processed secret composed of a user-chosen pseudonym and a user-selected sequence of images. All operations involving human-readable secrets are carried out entirely within the participant's browser. The survey backend receives and stores only a salted hash derived from this secret, which serves as the participant identifier for longitudinal linkage within a defined family of forms.

3.1. Participant' secret and normalization

Each participant provides two elements, namely a pseudonym P entered as free text and an ordered sequence of images $\boldsymbol{I} = (i_1, i_2, \ldots, i_k)$, $k \in \mathbb{N}$ selected from a fixed catalog in the survey interface (each i_j is an integer in a bounded range (for example, $0 \leq i_j \leq 63$). The first step of the construction is to map $(P, \boldsymbol{I})$ into a canonical intermediate representation. The pseudonym is processed by a normalization function $norm(\cdot)$ that trims leading and trailing whitespace, converts all characters to lowercase, replaces locale-specific letters with their basic Latin counterparts, and removes any remaining characters that are not letters or digits. This yields a normalized pseudonym

$$P^* = norm(P),$$

which ensures that trivial variations in spelling, case, or diacritics do not produce different identifiers.

3.2. Deriving the core code

The image sequence is encoded as a decimal string that uniquely represents both its length and its elements. A simple encoding function $enc(\cdot)$ can be defined as follows: the length k is written as a two-digit decimal number with leading zeros if necessary, and each image identifier i_j is also written as a two-digit decimal number. These substrings are concatenated to form a single string

$$S = enc(\boldsymbol{I})$$

For example, if $\boldsymbol{I} = (5, 27, 9, 42)$ and $k = 4$, then S would be the string "0405270942". This encoding is purely internal to the algorithm and is not exposed to the participant.

From P^* and S, the browser derives a short alphanumeric core code CORE that serves as an intermediate, human-readable representation of the participant's secret. A cryptographic hash function H (such as SHA–256) is applied separately to the normalized pseudonym and to the encoded image sequence, producing two hash values

$$h_P = H(P^*), h_I = H(S),$$

which are treated as hexadecimal strings. These two strings are concatenated with a delimiter (for example, the character "|") to form a combined string

$$C = h_P \| " - " \| h_I.$$

A second application of the hash function yields

$$h_C = H(C).$$

The value h_C is then mapped to a fixed-length alphanumeric code. To this end, the first m bytes of h_C (for example, $m = 5$, corresponding to 10 hexadecimal characters) are interpreted as a non-negative integer $n = bytesm(h_C)$. This integer is encoded in base 36 using digits and uppercase letters by a function $base36(\cdot)$, giving

CORE = base36(n),

which is padded with leading zeros or truncated as necessary to obtain a fixed length L (for example $L = 8$). The mapping

$$(P, \boldsymbol{I}) \to P^* \to S \to h_P, h_I \to h_C \to CORE$$

is deterministic for a given pseudonym and image sequence; however, assuming the cryptographic properties of H, it is computationally infeasible to invert and recover P or $\boldsymbol{I}$ from $CORE$.

3.3. Deriving the core code

To support the detection of typographical errors when participants enter their code manually, a checksum is computed over $CORE$ using a $modulus - 97$ construction similar to that employed in IBAN validation. Each character of $CORE$ is mapped to a numeric value in $\{0, \dots, 35\}$ by interpreting digits $0 - 9$ as $0 - 9$ and letters A–Z as 10–35. These values are concatenated as decimal strings to form a large integer N represented as a string. Two zeros are appended to this string, corresponding to multiplication by 100, and the remainder

$$r = N\,00\,mod\,97$$

is computed using an iterative modulus algorithm that operates on the decimal string to avoid overflow. The checksum is defined as

$$c = 98 - r,$$

with c set to 0 if $c = 98$, and it is then written as a two-digit decimal string in the range $"00" - "97"$. The human-readable version of the participant code that may be displayed to the respondent is

$$FULL = CORE"\text{-}"CHECKSUM$$

where $CHECKSUM$ is the two-digit representation of c. All of these operations occur entirely within the browser.

The checksum is used purely for client-side validation. When a participant chooses to work with the textual form of their code, the survey interface can request the pseudonym, the image sequence, and the full code they believe to be correct. The browser recomputes $CORE$ from the current $(P, \boldsymbol{I})$ pair, recalculates the checksum, and verifies that both match the user-entered $CORE - CHECKSUM$. If either the core or the checksum is inconsistent, the code is rejected locally, and the participant is prompted to correct the input. In this way, malformed or inconsistent identifiers are never transmitted to the server, reducing the creation of spurious new identifiers due to typographical mistakes.

3.4. Final identifier and linkage

The identifier used for linkage in the survey backend is obtained by hashing $CORE$ together with a project-specific secret, $FORM\_FAMILY\_SALT$, chosen by the survey administrator. This salt is a randomly chosen value by the survey administrator for a given family of forms (for example, all waves of a single study) and is kept secret on the server side. It ensures that the same underlying secret (P, I) does not lead to identical identifiers across distinct projects that employ different salts. Given a core

$$PARTICIPANT\_ID = H(CORE \,\|\, "-" \,\|\, FORM\_FAMILY\_SALT).$$

The result is a hash value that can be represented as a hexadecimal string or in another canonical encoding, depending on backend requirements. Only $PARTICIPANT\_ID$ (and optionally the checksum as a non-sensitive auxiliary field) is sent to the server together with the survey responses and is stored in the database; neither the pseudonym, the image sequence, nor the core code is transmitted or persisted.

Within the overall survey workflow, the method is used whenever a participant first enrolls or returns for a subsequent wave. In a typical scenario, the survey interface displays a text field for the pseudonym and a panel of images from which the participant selects their personal sequence. Upon submission, the browser computes P^*, S, $CORE$, the checksum, and finally $PARTICIPANT\_ID$ using the secret salt associated with the form family. The participant may be shown $FULL\_CODE$ for their own records, but the server receives only $PARTICIPANT\_ID$ and the answers to the survey questions. When the same participant returns in a later wave of the same study, they enter the same pseudonym and select the same images, the browser repeats the computation, and the resulting $PARTICIPANT\_ID$ matches the value stored for their previous responses. Longitudinal linkage is thus achieved solely by equality of salted hashes, without the survey platform ever learning or storing the human-readable secrets that generated them.

3.5. Discussion

The proposed algorithm aims to strike a balance between traditional textual SGIC schemes and fully system-generated anonymous identifiers. By combining a user-chosen pseudonym with a user-selected sequence of images and processing this composite secret entirely on the client, the method preserves the central intuition of SGICs-that respondents themselves provide the material needed for longitudinal linkage–while replacing quasi-identifying code elements such as dates of birth or parental initials with a more abstract, cryptographically mediated construction. This section discusses the methodological and privacy-related implications of this design, focusing on expected effects on matching performance, error profiles, anonymity, and practical deployment in web-based surveys.

From the perspective of matching performance, the most important change introduced by the algorithm is the separation between the human-facing secret and the machine-facing identifier. In classical SGICs, respondents must remember and re-enter the same textual recipe (often combining several elements), and any inconsistency directly affects the matching key stored on the server. In the proposed scheme, the only values that reach the server are salted hashes of the derived core code; the heavy reliance on user memory and consistency is shifted to the client, where it can be supported by validation mechanisms that prevent formed identifiers from being recorded. The cryptographic checksum on the core code precisely addresses the problem that typographical errors and partial recall lead to spurious new identifiers. As long as the pseudonym and image sequence are reproduced correctly, the client can detect inconsistencies between the recomputed core and the user-entered code and can reject invalid codes before any data is submitted. Conceptually, this should reduce the fraction of unmatched cases arising purely from input errors. However, the method does not eliminate unmatched cases resulting from genuine changes in the participant's secret or complete forgetting.

The use of images as part of the secret is motivated by cognitive considerations. Image-based secrets capitalize on the tendency of people to remember visual patterns and stories more readily than arbitrary alphanumeric strings, suggesting that a small sequence of images coupled with a pseudonym may be easier to recall over time than a multi-element textual SGIC. At the same time, the image sequence is never transmitted or stored; it is only used locally to derive a hash that participates in the construction of the core code. This design aims to leverage the memorability of graphical secrets without disclosing additional quasi-identifying information or rich metadata to the backend. However, it also introduces a new kind of dependency: the participant must be able to recognize the same image catalog and reconstruct their sequence in subsequent waves. If the visual design of the image panel changes significantly, or if different devices confusingly render images, reconstruction may fail, resulting in unmatched cases. The method, therefore, implicitly assumes a stable and well-designed image set throughout the study's life.

Regarding privacy and anonymity, the algorithm is deliberately conservative. The server never receives the pseudonym P, the normalized pseudonym P^*, the image sequence $\boldsymbol{I}$, the encoded image string S, or the intermediate core code $CORE$. It only receives a salted hash of $CORE$, where the salt is specific to a form family. Under standard assumptions about the cryptographic hash function, this means that inverting the participant identifier to recover the underlying secret is computationally infeasible, and that the same participant generates different identifiers in different projects that use various salts. This mitigates the risk of unintended cross-project linkage and constrains the consequences of a data breach: an attacker obtaining the database of identifiers would not be able to view any human-readable components and would lack the necessary salts to apply dictionary attacks effectively. The method thus aligns with the concept of pseudonymization at the interface between client and server, while aiming to achieve a level of protection that approximates anonymity from the server's perspective.

The introduction of a form family salt also has methodological implications. Because $PARTICIPANT\_ID$ is derived from both $CORE$ and $FORM\_FAMILY\_SALT$, identifiers are inherently scoped to the family of forms for which the salt is defined. This scoping is desirable when researchers wish to avoid linking individuals across independent studies or institutional boundaries; however, it also means that deliberate cross-study linkage (for instance, combining two projects conducted by the same team) cannot be achieved without reusing or coordinating salts at the time of data collection. The algorithm, therefore, encodes a particular stance on linkage: it makes within-study linkage easy and automatic, while making cross-study linkage a deliberate design decision rather than an accidental side-effect of code reuse. Despite these advantages, several limitations and potential sources of bias remain. The method cannot prevent unmatched cases that arise when participants intentionally change their pseudonyms or image sequences, or when they forget their secrets entirely. In such situations, the algorithm correctly treats the responses as coming from a new anonymous individual. Whether this behavior is desirable depends on the analytic goals: for some studies, a strict requirement that only genuinely reproducible secrets lead to linkage may be appropriate; for others, it may be preferable to tolerate more aggressive matching rules at the risk of occasional false links. Furthermore, the cognitive burden of choosing and remembering an appropriate pseudonym and image sequence may itself be associated with participant characteristics such as age, education, or digital literacy. If certain groups are more likely to choose unstable secrets or to forget them, then even a technically robust algorithm may still induce differential matching rates and selection effects.

The introduction of a form family salt also has methodological implications. Because $PARTICIPANT\_ID$ is derived from both $CORE$ and $FORM\_FAMILY\_SALT$, identifiers are inherently scoped to the family of forms for which the salt is defined. This scoping is desirable when researchers wish to avoid linking individuals across independent studies or institutional boundaries; however, it also means that deliberate cross-study linkage (for instance, combining two projects conducted by the same team) cannot be achieved without reusing or coordinating salts at the time of data collection. The algorithm, therefore, encodes a particular stance on linkage: it makes within-study linkage easy and automatic, while making cross-study linkage a deliberate design decision rather than an accidental side-effect of code reuse. Despite these advantages, several limitations and potential sources of bias remain. The method cannot prevent unmatched cases that arise when participants intentionally change their pseudonyms or image sequences, or when they forget their secrets entirely. In such situations, the algorithm correctly treats the responses as coming from a new anonymous individual. Whether this behavior is desirable depends on the analytic goals: for some studies, a strict requirement that only genuinely reproducible secrets lead to linkage may be appropriate; for others, it may be preferable to tolerate more aggressive matching rules at the risk of occasional false links. Furthermore, the cognitive burden of choosing and remembering an appropriate pseudonym and image sequence may itself be associated with participant characteristics such as age, education, or digital literacy. If certain groups are more likely to choose unstable secrets or to forget them, then even a technically robust algorithm may still induce differential matching rates and selection effects.

Security considerations also warrant discussion. The scheme protects against server-side misuse and external attackers, but does not prevent impersonation by someone who learns another participant's secret. If a participant discloses both their pseudonym and their image sequence to a third party, the third party can, in principle, reproduce the same CORE and therefore the same $PARTICIPANT\_ID$ in future waves of the same study. This limitation is shared with traditional SGICs and passwordbased systems: any method based on a user-controlled secret is vulnerable to voluntary disclosure. The algorithm reduces the risk of *involuntary* leakage by never transmitting the secret itself, but it cannot control what participants choose to reveal. In practical survey contexts, the incentives for such impersonation are typically low, but the possibility should be acknowledged when assessing the overall threat model. From an implementation standpoint, the method is intentionally modest in its technical requirements. It assumes the availability of a standard cryptographic hash function in the browser, basic string and numeric operations, and the ability to store and use a single form family salt on the server side. This approach makes integration into existing survey platforms conceptually straightforward: the client-side logic can be packaged as a reusable script, and the backend can treat $PARTICIPANT\_ID$ as just another string field used for grouping responses. At the same time, this simplicity means that the method does not attempt to implement more advanced privacy techniques such as differential privacy, secure multi-party computation, or hardware-backed key storage. It is best understood as a structured refinement of SGIC-like linkage tailored to the affordances of current web technologies, rather than as a comprehensive privacy framework.

Taken together, these considerations suggest that the proposed algorithm addresses several key weaknesses of classical SGICs—particularly the direct exposure of quasi-identifiers and the fragile coupling between user input and stored identifiers—while introducing its own set of assumptions about user behavior, interface stability, and salt management. Its main strengths lie in the clear separation between human-facing secrets and machine-facing identifiers, the use of client-side validation to reduce error-induced unmatched cases, and the scoping of identifiers to specific study families. Its principal vulnerabilities relate to participant memory, voluntary disclosure, and the potential for differential matching performance across subgroups. These trade-offs are inherent to the problem of anonymous longitudinal linkage and frame the conditions under which the method can be considered appropriate for a given research design.

5. Conclusions and future work

This paper introduces a client-side method for generating anonymous yet linkable participant identifiers in web-based surveys. The approach replaces the traditional SGIC-based linkage with a structured secret composed of a user-chosen pseudonym and an image sequence, which is processed locally through normalization, encoding, and cryptographic hashing. By exposing only, a salted hash of the derived core code to the server, the method supports longitudinal linkage while preserving strong anonymity and preventing the backend from learning any human-readable identifiers. The use of a checksum further strengthens reliability by reducing the likelihood that typographical errors produce spurious identifiers.

The proposed scheme demonstrates how principles from SGIC design, graphical secrets, and browser-based cryptography can be combined into a practical mechanism that aligns with modern privacy expectations. While the method offers more apparent separation between human-facing and machine-facing identifiers, it retains certain limitations: matching performance still depends on participants' ability to reproduce their pseudonym and image sequence, and the design remains vulnerable to intentional disclosure of the secret. As such, the method serves as a conceptual refinement of, rather than a complete solution to, all challenges posed by participant-generated identifiers.

Future work should validate the approach empirically, particularly its impact on matching rates, error patterns, and subgroup differences in usability. Additional research is needed to formalize the security properties of the scheme, assess resistance to guessing and brute-force attacks, and evaluate how parameters such as salt management and core length affect robustness. Finally, practical integration into survey platforms and user-interface studies will be essential to determine how participants understand, remember, and reliably reproduce the required secret.

In summary, the method provides a clear and technically grounded direction for improving anonymous longitudinal linkage, illustrating how client-side cryptographic processing can enhance both privacy protection and methodological reliability in survey research.

Acknowledgments
The author thanks the Department of Applied Informatics at Wrocław University of Science and Technology for their support. I acknowledge the use of ChatGPT for assistance in reviewing my writing and supporting the translation from my first language.

References

Audette L.M., Hammond M.S., Rochester N.K. (2020), Methodological issues with coding participants in anonymous psychological longitudinal studies, "Educational and Psychological Measurement", vol. 80 no. 1, pp. 163–185.

Brändle T., Pläschke A. (2024), Beyond matching rates: examining the accuracy of self-generated ID codes, https://osf.io/preprints/osf/k98j6_v1 [21.12.2025].

Calatrava M., de Irala J., Osorio A., Benítez E., Lopez-del Burgo C. (2022), Matched and fully private? A new self-generated identification code for school-based cohort studies to increase perceived anonymity, "Educational and Psychological Measurement", vol. 82 no. 3, pp. 465–481.

DiIorio C., Soet J.E., Van Marter D., Woodring T.M., Dudley W.N. (2000), An evaluation of a self-generated identification code, "Research in Nursing & Health", vol. 23 no. 2, pp. 167–174.

Direnga J., Timmermann D., Lund J., Kautz C. (2016), Design and application of self-generated identification codes (SGICs) for matching longitudinal data, in: 44th SEFI Annual Conference, 12–15 September 2016, Tampere.

Grube J.W., Morgan M., Kearney K.A. (1989), Using self-generated identification codes to match questionnaires in panel studies of adolescent substance use, "Addictive Behaviors", vol. 14 no. 2, pp. 159–171.

Kearney K.A., Hopkins R.H., Mauss A.L., Weisheit R.A. (1984), Self-generated identification codes for anonymous collection of longitudinal questionnaire data, "Public Opinion Quarterly", vol. 48 no. 1B, pp. 370–378.

Krótkiewicz M., Wojtkiewicz K., Martins D. (2018), Influence power factor for user interface recommendation system, in: International conference on Computational Collective Intelligence, Springer International Publishing, Cham, pp. 228–237.

Kuzu M., Kantarcioglu M., Inan A., Bertino E., Durham E., Malin B. (2013), Efficient privacy-aware record integration, in: Proceedings of the 16th International Conference on Extending Database Technology, pp. 167–178, https://doi.org/10.1145/2452376.2452398.

Platje J., Palak R., Wojtkiewicz K. (2025), Sokrates forms. A research instrument for creating social impact of science on the example of system risk management, "The Central European Review of Economics and Management (CEREM)", vol. 9 no. 1, pp. 47–62.

Sandnes F.E. (2021a), CANDIDATE. A tool for generating anonymous participant-linking IDs in multi-session studies, "PloS One", vol. 16 no. 12, e0260569.

Sandnes F.E. (2021b), HIDE: short IDs for robust and anonymous linking of users across multiple sessions in small HCI experiments, in: Extended abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Kitamura Y. (ed.), Association for Computing Machinery, Washington, pp. 1–6.

Vacek J., Gabrhelík R. (2024), Feasibility study of linking anonymous data of children in longitudinal school-based prevention research, “Addictology/Adiktologie”, vol. 24 no. 2, pp. 99–108.

Wiedenbeck S., Waters J., Birget J.C., Brodskiy A., Memon N. (2005), Authentication using graphical passwords. Effects of tolerance and image choice, in: Proceedings of the 2005 symposium on usable privacy and security, Association for Computing Machinery, Washington, pp. 1–12.

Wojtkiewicz K., Palak R., Telec Z., Litwinienko F. (2024), Modelling cognitive load of computer game users-case study, in: European conference on artificial intelligence, Springer Nature Switzerland, Cham, pp. 52–63.

Yurek L.A., Vasey J., Sullivan Havens D. (2008), The use of self-generated identification codes in longitudinal research, “Evaluation Review”, vol. 32 no. 5, pp. 435–452.