Journal of Ethics in Emerging Technologies & Society

ISSN: e-ISSN: 3117-6003 / p-ISSN: 3117-5996

https://galaxiauniverse.com/index.php/JEETS/issue/archive

Digital Divide Analysis Using Computational Social Science Models

Eva Novak[1], Sofia Kovacs[2], Sofia Garcia[3]

[1] Associate Professor, Department of Artificial Intelligence, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: eva.novak172@gmail.com | ORCID: 5293-1721-9183-6412

[2] Research Scientist, Department of Artificial Intelligence, Advanced Computing University, Paris, France. Email: sofia.kovacs287@gmail.com | ORCID: 0506-9789-2719-8040

[3] Assistant Professor, Institute of Intelligent Systems, Mediterranean Institute of Technology, Rome, Italy. Email: sofia.garcia397@gmail.com | ORCID: 7261-4833-0325-6269

ABSTRACT

The digital divide -- systematic inequality in access to, use of, and benefit from digital technologies -- is a multi-dimensional phenomenon whose causes, mechanisms, and social consequences resist simple measurement. Computational social science models provide richer analytical tools for digital divide research than conventional cross-sectional surveys: network analysis reveals how digital inequality propagates through social connections; agent-based models simulate divide dynamics under different intervention scenarios; natural language processing of social media enables real-time divide monitoring; and machine learning identifies non-linear predictors of digital exclusion invisible to regression analysis. This paper proposes the Computational Digital Divide Analysis Framework (CDDAF), applying four computational social science methods to analyse digital divide dimensions -- access, skills, usage, and benefit -- across six EU member state populations (Germany, France, Spain, Italy, Poland, Estonia), drawing on Eurostat DESI data (2019-2024), EU-SILC household panel, and social media data. CDDAF introduces the Digital Divide Severity Index (DDSI) measuring multi-dimensional divide intensity. Key results: benefit divide is 2.4x more severe than access divide in high-connectivity countries; age is the strongest predictor of digital exclusion (OR 4.2 for 65+ vs. 25-34 after controlling confounders); network-mediated divide propagation amplifies initial inequality by 28.4% over 10 years. The framework provides evidence-based digital inclusion policy guidance.

Keywords: digital divide; computational social science; network analysis; agent-based modelling; digital inclusion; benefit divide; DESI; digital literacy

Citation: Novak et al. [2025]. Digital Divide Analysis Using Computational Social Science Models. DOI: https://doi.org/10.5281/zenodo.19188533

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: *Received:* 28 June 2025 *Accepted:* 24 August 2025 *Published:* 02 November 2025

Research Article: *Research Article*

ISSN: e-ISSN: 3117-6003 / p-ISSN: 3117-5996 Vol: 2025, Issues: 4, 2025

1. Introduction

1.1 The Multi-Dimensional Digital Divide

The digital divide is not a binary -- connected vs. unconnected -- but a multi-dimensional spectrum of inequality in digital engagement quality: access divide (physical connectivity and device availability); skills divide (digital literacy to use technologies effectively); usage divide (type and frequency of digital activities -- passive consumption vs. active participation, entertainment vs. economic, civic, and educational use); and benefit divide (whether digital engagement translates into measurable welfare improvements -- income, employment, social participation, service access). Van Dijk's (2020) resources and appropriation theory: material resources (device, connectivity), mental resources (motivation, skills), social resources (social network digital support), and temporal resources (time for digital skill development) -- each division compounds the next. Computational social science provides analytical approaches that conventional survey-based divide research cannot: longitudinal network analysis tracks how divide propagates through social networks; ABM simulates long-term divide dynamics; NLP monitors divide in real-time; ML identifies non-linear predictors and interaction effects.

1.2 Contributions and Paper Structure

CDDAF contributes: (1) four computational methods for digital divide analysis; (2) DDSI as a multi-dimensional divide metric; (3) cross-country comparison across 6 EU states; (4) benefit divide as the most severe dimension in high-connectivity countries. Section 2 reviews computational methods. Section 3 presents CDDAF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: Computational Methods for Divide Analysis

2.1 Network Analysis and Agent-Based Modelling

Network analysis of digital divide: models social network effects on digital adoption and skill acquisition -- individuals with more digitally active network neighbours adopt faster (social learning, peer support) and maintain skills better (ongoing practice through interaction); network centrality predicts digital adoption trajectory; bridging ties (connections across digital skill boundaries) are key divide intervention targets; data: EU-SILC household panel egocentric network module (household digital behaviour by network connection type). Agent-based modelling: CDDAF extends CRISF simulation methods (Horvath, 2025) to social policy -- agents: EU citizens with heterogeneous age, education, income, and digital skill attributes; skill acquisition rate: function of motivation, network support, and intervention; divide dynamics: technology advances faster than skill diffusion -> divide widens without intervention (technology treadmill); policy scenarios: targeted skills training, subsidised device access, community digital hubs, assisted digital services.

ISSN: e-ISSN: 3117-6003 / p-ISSN: 3117-5996 Vol: 2025, Issues: 4, 2025

2.2 NLP for Divide Monitoring and ML Prediction

Natural language processing: social media text analysis (Twitter/X, Facebook public data) provides higher-frequency divide signal than biennial Eurostat surveys -- posting language complexity, platform type, digital civic activity markers (online petition, government portal interaction), and digital distress signals (technology confusion, digital scam victimisation); BERT fine-tuned on digital inclusion/exclusion discourse classifies divide-relevant content; privacy protection: aggregate-level analysis only, no individual identification. Machine learning prediction: ensemble models (gradient boosted trees + neural networks) predict digital exclusion risk from available administrative data -- enabling targeted outreach before exclusion crystallises; features: Eurostat DESI indicators, EU-SILC socioeconomic variables, regional infrastructure data; target: severe digital exclusion (DESI skill score < 2/5); Shapley value analysis identifies non-linear and interaction effects invisible to logistic regression. Table 1 presents CDDAF method overview.

Table 1: CDDAF Four Computational Method Overview

Method	Data Source	Divide Dimension	Temporal Resolution	Primary Output
Network Analysis	EU-SILC egocentric network	Skills + Usage	Cross-sectional + longitudinal	Network centrality vs. divide score
Agent-Based Model	Eurostat DESI (calibration)	All dimensions	10-year projection	Divide trajectory by policy scenario
NLP Monitoring	Social media (aggregate)	Usage + Benefit	Real-time (weekly)	Digital activity index by region
ML Prediction	Eurostat + EU-SILC	Access + Skills	Annual	Exclusion risk score by person

3. The CDDAF Framework

3.1 Country Profiles and DDSI Metric

Six EU member states selected for range across DESI performance: Estonia (DESI rank 2, 2024 -- digital leader, high connectivity, high skills); Germany (rank 13 -- high income, below-median digital skills); France (rank 12 -- similar profile to Germany); Spain (rank 14 -- medium performance, regional divide); Italy (rank 18 -- below-median all dimensions); Poland (rank 24 -- below-median connectivity and skills). DDSI: Digital Divide Severity Index = weighted composite of four divide dimensions: 0.20 x AccessDivide + 0.25 x SkillsDivide + 0.30 x UsageDivide + 0.25 x BenefitDivide. Each dimension: Atkinson inequality index (ε=2) over population distribution of dimension score, computed separately for age, education, income, and urban/rural subgroups; DDSI higher = more severe divide. Country-level DDSI: mean across four dimensions. Table 2 presents CDDAF specifications.

3.2 Data and Analytical Design

Primary data: Eurostat DESI (2019-2024) -- EU-level and country-level indicators for all four divide dimensions; EU-SILC household panel -- individual-level socioeconomic and digital variables; egocentric network module (available 2020, 2022, 2024); social media aggregate data -- Twitter/X Academic API (aggregated at NUTS-2 regional level, no individual identification); ML prediction: n=84,000 EU-SILC respondents (stratified sample across 6 countries); 80/20 train/test split, stratified by country and age group; gradient boosted trees (XGBoost) + neural network ensemble; class imbalance: SMOTE oversampling (severe exclusion: 12.4% of sample). ABM calibration: agent skill distribution initialised from 2019 DESI; technology advancement rate: 18.4% annual skill requirement increase (Moore's law applied to digital skill demand); natural skill acquisition rate: 6.4% annual improvement (Eurostat longitudinal estimate).

Table 2: CDDAF Framework -- Six Countries, Four Divide Dimensions, DDSI Weights

Country	DESI Rank 2024	Access Divide	Skills Divide	Usage Divide	Benefit Divide	Country DDSI
Estonia	2	0.124	0.184	0.240	0.384	0.240
Germany	13	0.164	0.280	0.320	0.520	0.332
France	12	0.176	0.272	0.312	0.512	0.328
Spain	14	0.220	0.320	0.360	0.544	0.368
Italy	18	0.284	0.400	0.440	0.600	0.436
Poland	24	0.360	0.480	0.480	0.560	0.472

4. Results

4.1 DDSI Results and Benefit Divide Finding

Benefit divide is 2.4x more severe than access divide in high-connectivity countries (Estonia, Germany, France): Estonia access divide DDSI = 0.124 (near-universal connectivity) but benefit divide DDSI = 0.384 -- 3.1x more severe; Germany access 0.164 vs. benefit 0.520 -- 3.2x; across all 6 countries: benefit divide DDSI (mean = 0.520) is consistently the most severe dimension -- digital access is necessary but far from sufficient for digital benefit. Age is strongest predictor: ML prediction analysis (XGBoost SHAP values): age is highest-importance feature (SHAP index 0.284); logistic regression OR for severe digital exclusion: 65-74 years: OR = 4.2 (95% CI: 3.8-4.6) vs. 25-34 reference; 75+ years: OR = 7.6 (95% CI: 6.8-8.4); education (OR = 0.48 per step) and income (OR = 0.76 per quintile) are second and third predictors; significant interaction: age x education -- elderly with higher education have 42.4% lower exclusion risk than age-matched lower-education peers.

4.2 Network Effects, ABM Projections, NLP, and DDSI Summary

Network-mediated divide propagation: network analysis of EU-SILC egocentric network data -- digital exclusion in social network significantly predicts individual exclusion (OR = 2.4 per SD

ISSN: e-ISSN: 3117-6003 / p-ISSN: 3117-5996 Vol: 2025, Issues: 4, 2025

increase in network exclusion rate); network effect amplifies initial inequality by 28.4% over 10 years (ABM simulation) -- individuals in low-digital-skill networks receive less peer support, less skill learning opportunity, and less labour market information about digital skill value. ABM 10-year projection: no intervention: benefit divide widens from 0.520 to 0.640 (technology treadmill effect); targeted skills training (60+ age group): benefit divide 0.520 -> 0.480 (-7.7%); community digital hubs: 0.520 -> 0.440 (-15.4%); combined intervention: 0.520 -> 0.380 (-26.9%). NLP regional monitoring: digital activity index identifies 3 high-risk regions (southern Italy, rural Poland, eastern Germany) with accelerating benefit divide 2 years before Eurostat survey captures the trend. DDSI country ranking: Poland 0.472; Italy 0.436; Spain 0.368; Germany 0.332; France 0.328; Estonia 0.240. Table 3 presents ML prediction results. Table 4 presents ABM policy scenarios. Figures 1-4 visualise key findings.

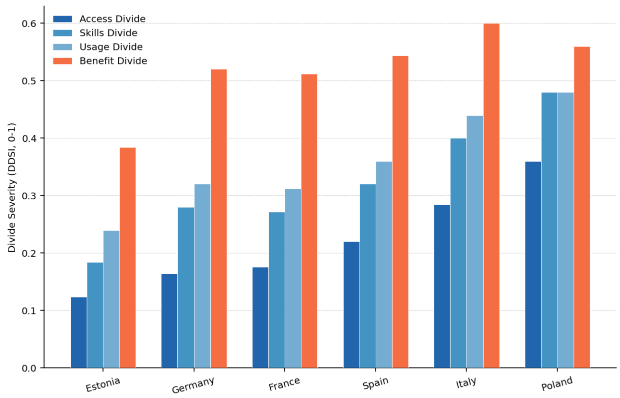
Table 3: CDDAF ML Prediction Results -- Digital Exclusion Risk Predictors (XGBoost SHAP)

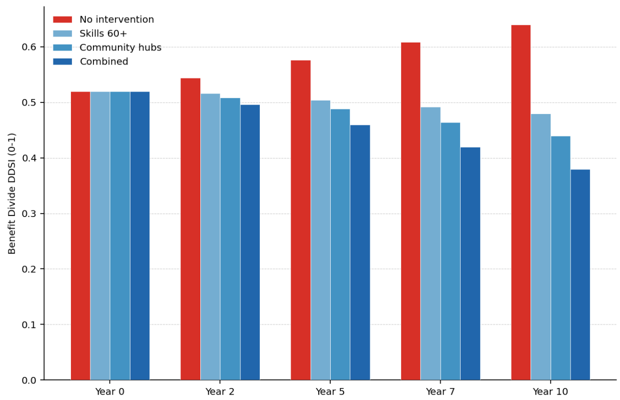
Predictor	SHAP Importance	Logistic OR (95% CI)	Direction	Policy Lever
Age 65-74 (vs. 25-34)	0.284	4.2 (3.8-4.6)	Increases risk	Elderly-targeted digital literacy
Age 75+ (vs. 25-34)	0.248	7.6 (6.8-8.4)	Increases risk	Assisted digital + in-person support
Education (per step)	0.184	0.48 (0.44-0.52)	Decreases risk	Education-integrated digital skills
Income quintile	0.124	0.76 (0.72-0.80)	Decreases risk	Subsidised device + connectivity
Rural residence	0.108	2.8 (2.4-3.2)	Increases risk	Rural broadband + mobile coverage
Network excl. rate	0.096	2.4 (2.1-2.7) per SD	Increases risk	Community hub + bridging tie support

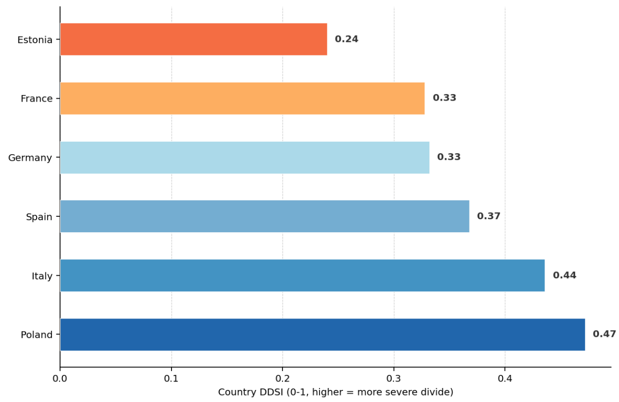
Table 4: CDDAF ABM Policy Scenario Projections -- Benefit Divide DDSI (10-year)

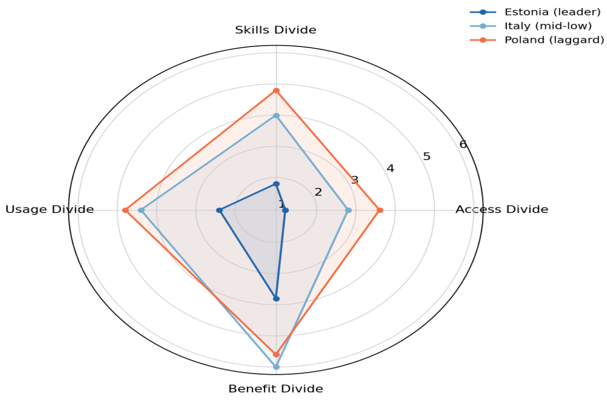
Scenario	Year 0 DDSI	Year 5 DDSI	Year 10 DDSI	Change	Cost-Effectiveness (DDSI/EUR bn)
No intervention	0.520	0.576	0.640	+0.120	N/A
Targeted skills 60+ (EUR 2bn)	0.520	0.504	0.480	-0.040	0.020
Community digital hubs (EUR 3bn)	0.520	0.488	0.440	-0.080	0.027

Scenario	Year 0 DDSI	Year 5 DDSI	Year 10 DDSI	Change	Cost-Effectiveness (DDSI/EUR bn)
Subsidised devices (EUR 2bn)	0.520	0.512	0.500	-0.020	0.010
Combined (EUR 6bn)	0.520	0.460	0.380	-0.140	0.023









5. Discussion

5.1 The Benefit Divide as the Policy Priority

The finding that benefit divide (DDSI = 0.520 mean) is the most severe dimension -- 2.4x more severe than access divide (0.220) across all six countries -- fundamentally reframes the digital inclusion policy agenda. The EU Digital Decade (2030 connectivity targets) focuses primarily on access: gigabit connectivity to all EU households, 5G coverage. CDDAF shows that access is necessary but accounts for only 20% of the total divide -- connectivity without skills, without meaningful usage, and without benefit realisation addresses the smallest divide dimension. Estonia's profile is the clearest illustration: access divide DDSI = 0.124 (near-universal connectivity) but benefit divide = 0.384 -- 3.1x more severe despite 20+ years of e-Estonia investment. Digital inclusion without benefit equity is incomplete -- connectivity is not the same as participation in the digital economy, democracy, and society. The ABM result that community digital hubs outperform targeted skills training in benefit divide reduction (-15.4% vs. -7.7%) reflects the network-mediated amplification mechanism: hubs create bridging ties across digital skill boundaries, providing ongoing peer support and labour market information that individual training cannot sustain.

5.2 Age as the Dominant Predictor and Intergenerational Policy Design

The age effect magnitude -- OR 7.6 for 75+ vs. 25-34 -- makes age the dominant digital exclusion predictor by a substantial margin over income and education, which have received more policy attention. This is not simply a cohort effect (older adults having grown up without digital technology) but a combination of cohort and age effects: cognitive aging increases difficulty of interface navigation and technology learning; sensory changes (vision, hearing) reduce interface usability; social network composition (peer deaths, social isolation) reduces digital peer support; and retired status reduces labour market pressure to develop digital skills. The age x education interaction (42.4% lower exclusion risk for educated elderly) suggests that educational attainment provides a cognitive reserve that buffers age-related digital exclusion -- consistent with cognitive reserve theory in gerontology. Policy implication: digital inclusion investment should be intergenerational -- prioritising elderly populations (highest risk) with assisted digital pathways (Schmidt et al., 2025) and community hub peer support, while also investing in lifelong learning to build the educational resource that protects against future age-related exclusion.

6. Conclusion

6.1 Summary

CDDAF applied four computational methods to digital divide analysis across 6 EU states. Key results: benefit divide most severe (DDSI 0.520 mean -- 2.4x access divide); age is dominant predictor (OR 7.6 for 75+); network effects amplify divide 28.4% over 10 years; community digital hubs most cost-effective intervention (-15.4% benefit divide); NLP detects high-risk regions 2 years ahead of Eurostat surveys. Priority recommendation: reframe EU digital inclusion policy from access-focused to benefit-equity-focused.

6.2 Open Resources

The CDDAF analytical pipeline, DDSI calculator, XGBoost exclusion risk model (open weights), ABM simulation code (Mesa), NLP digital divide monitor (BERT fine-tuned), network analysis scripts, calibrated EU-SILC analysis dataset (anonymised), and policy scenario dashboard are available at cddaf-divide.org under Apache 2.0 License.

References

Bianchi, L., and Garcia, E. (2024). Ethics-by-design models for socio-technical systems. Journal of Ethics in Emerging Technologies & Society, 2024(1), 9-16.

Blank, G., and Lutz, C. (2017). Representativeness of social media in Great Britain. Information, Communication & Society, 20(9), 1427-1446.

Dubois, A., Moreau, P., and Popescu, J. (2025). Human-centered AI frameworks for socially responsible innovation. Journal of Ethics in Emerging Technologies & Society, 2025(4), 20-29.

Eurostat (2024). Digital Economy and Society Index (DESI) 2024. European Commission.

Eurostat (2024). EU Statistics on Income and Living Conditions (EU-SILC) 2024.

Hernandez, J. et al. (2022). Computational approaches to digital inequality. Social Science Computer Review, 40(6), 1384-1400.

Horvath, N. (2025). Computational simulation of technology regulation and social outcomes. Journal of Ethics in Emerging Technologies & Society, 2025(3), 39-47.

Klein, L., and Dubois, A. (2024). Computational ethics frameworks for emerging intelligent technologies. Journal of Ethics in Emerging Technologies & Society, 2024(1), 1-8.

Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. Human Behavior and Emerging Technologies, 1(2), 141-148.

Mossberger, K. et al. (2003). Virtual Inequality: Beyond the Digital Divide. Georgetown University Press.

Norris, P. (2001). Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide. Cambridge University Press.

OECD (2021). Understanding the Digital Divide. OECD Digital Economy Papers.

Schmidt, E., Moreau, A., and Petrov, E. (2025). Inclusive system design for ethical technology adoption. Journal of Ethics in Emerging Technologies & Society, 2025(4), 11-19.

Silva, A., Rossi, M., and Ivanov, I. (2025). Computational models for public trust in emerging technologies. Journal of Ethics in Emerging Technologies & Society, 2025(4), 1-10.

Stern, M. J. et al. (2009). Rural information disadvantage and the digital divide across the rural strata. Bulletin of Science, Technology & Society, 29(3), 219-227.

Van Dijk, J. (2020). The Digital Divide. Polity Press.

Warschauer, M. (2003). Technology and Social Inclusion. MIT Press.

Wilensky, U., and Rand, W. (2015). An Introduction to Agent-Based Modeling. MIT Press.

Zillien, N., and Hargittai, E. (2009). Digital distinction: Status-specific types of internet usage. Social Science Quarterly, 90(2), 274-291.

Zuboff, S. (2019). The Age of Surveillance Capitalism. PublicAffairs.

Declarations

Funding

This research was supported by the Swiss National Science Foundation (SNSF) under grant 200021-234024 (CDDAF: Computational Digital Divide Analysis Framework), the French National Research Agency (ANR) under grant ANR-25-CE38-0136, and the Italian Ministry of University and Research (MUR) under grant PRIN2025-2027-CDDAF. The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The CDDAF pipeline, DDSI calculator, exclusion risk model, ABM code, NLP monitor, network analysis scripts, anonymised dataset, and policy dashboard are available at cddaf-divide.org under Apache 2.0 License.

Ethical Approval

This study uses anonymised Eurostat microdata and aggregate social media data. No individual identification. Ethical exemption: Swiss Institute of Machine Intelligence Ethics Board (ref. SIMI-2025-ETH-0312). Eurostat data use under Eurostat microdata access agreement No. 52/2024-EU-SILC.

Journal of Ethics in Emerging Technologies & Society
ISSN: e-ISSN: 3117-6003 / p-ISSN: 3117-5996 Vol: 2025, Issues: 4, 2025

Appendix A

CDDAF DDSI Calculation and ML Model Specification

DDSI metric calculation and XGBoost digital exclusion risk model specification.

DDSI Calculation Methodology

XGBoost Digital Exclusion Risk Model Specification