Scientific Research Methodology: Project Ideas

Contact: beniamin.bogosel@uav.ro

Choose one of the questions/themes/ideas proposed below. If you wish, you may propose a personal theme (provided it is validated; send me an email).

The project will be structured as follows:

- Present the theme, its context and the associated history.
- Formulate the associated research problem; justify whether it is researchable; state the general objective and formulate specific objectives.
- Present arguments and counterarguments (if they exist) supporting the claims to be discussed.
- For each argument, give the reference used and its type (scientific article, website, online resource). The relevance and credibility of each source will be analyzed in light of the discussions held during the course.
- The project will be handed in as a written document. The document must include the bibliographic references used.
- Send the project by email in advance for validation.
- The project is presented during the course (10–15 minutes).
- Minimum 3–4 pages (Word, 12pt). Adding images and charts is appreciated. Add references for images: state where they were taken from.

The hints and starting references below are only entry points. You are expected to search for further sources and to assess their credibility yourself.

Themes/ideas/claims

General topics

1. Scientific progress was accelerated by the invention of the printing press.

   Hint: How can a historical claim be "tested"? Compare the standard view with its critics (e.g. A. Johns, *The Nature of the Book*, 1998).

   Start with: E. L. Eisenstein, *The Printing Press as an Agent of Change*, Cambridge University Press, 1979.

2. Consumption of processed foods leads to cancer.

   Hint: Define "processed". Most evidence comes from observational studies: discuss correlation vs. causation and confounding factors.

   Start with: V. Bouvard et al., Carcinogenicity of consumption of red and processed meat, *The Lancet Oncology* 16(16), 2015; T. Fiolet et al., Consumption of ultra-processed foods and cancer risk, *BMJ* 360:k322, 2018.

3. Wearing sunglasses increases the risk of skin cancer.

   Hint: Trace where the claim comes from (blogs, social media?) and whether any peer-reviewed study supports the proposed mechanism.

   Start with: World Health Organization, Ultraviolet radiation (fact sheet), `https://www.who.int`.

4. Exposure of children to short-form video content (TikTok-like) produces major changes in brain structure.

   Hint: Look at study design (cross-sectional vs. longitudinal), effect sizes, and how media report "brain changes".

   Start with: J. S. Hutton et al., Associations between screen-based media use and brain white matter integrity in preschool-aged children, *JAMA Pediatrics* 174(1), 2020; M. P. Paulus et al., Screen media activity and brain structure in youth, *NeuroImage* 185, 2019.

5. Electrification of the vehicle fleet leads to a reduction of global CO_2 emissions.

   Hint: Life-cycle analysis: battery production and the electricity mix of each country matter.

   Start with: G. Bieker, *A global comparison of the life-cycle greenhouse gas emissions of combustion engine and electric passenger cars*, ICCT, 2021; IEA, *Global EV Outlook* (annual).

6. Access to the information available on the Internet (free and in large quantity) leads to an increase in the intelligence level of the population.

   Hint: How is intelligence measured? Look at the Flynn effect and its recent reversal.

   Start with: B. Bratsberg, O. Rogeberg, Flynn effect and its reversal are both environmentally caused, *PNAS* 115(26), 2018; B. Sparrow et al., Google effects on memory, *Science* 333, 2011.

7. Why did interest in chess increase during the Covid pandemic?

   Hint: Several candidate causes (lockdowns, the series *The Queen's Gambit*, online streamers). How could you separate their effects? Google Trends gives usable data.

   Start with: Google Trends, `https://trends.google.com`; public statistics of online chess platforms.

8. What are the effects of childhood vaccination worldwide on the spread of contagious diseases?

   Hint: Use official coverage and incidence data; discuss how modelling studies estimate lives saved.

   Start with: A. J. Shattock et al., Contribution of vaccination to improved survival and health: modelling 50 years of the Expanded Programme on Immunization, *The Lancet* 403, 2024; WHO/UNICEF immunization coverage estimates.

9. Can visual learning effectively replace classical learning?

   Hint: Distinguish the "learning styles" hypothesis (weak evidence) from multimedia learning (better supported).

   Start with: H. Pashler et al., Learning styles: concepts and evidence, *Psychological Science in the Public Interest* 9(3), 2008; R. E. Mayer, *Multimedia Learning*, Cambridge University Press.

10. The Earth is not flat.

    Hint: List independent lines of evidence anyone can check, and analyze how flat-Earth sources argue.

    Start with: C. Garwood, *Flat Earth: The History of an Infamous Idea*, 2007; J. B. Russell, *Inventing the Flat Earth*, 1991.

11. Building insulation reduces the energy consumption for heating/cooling.

    Hint: Compare engineering predictions with measured savings; look up the "rebound effect".

    Start with: M. Fowlie, M. Greenstone, C. Wolfram, Do energy efficiency investments deliver? Evidence from the Weatherization Assistance Program, *Quarterly Journal of Economics* 133(3), 2018.

12. Massive deforestation leads to soil destabilization and frequent floods.

    Hint: The evidence depends on the scale (small catchments vs. large river basins). Compare the two references below.

    Start with: C. J. A. Bradshaw et al., Global evidence that deforestation amplifies flood risk and severity in the developing world, *Global Change Biology* 13, 2007; FAO & CIFOR, *Forests and Floods: Drowning in Fiction or Thriving on Facts?*, 2005.

13. Carrying a mobile phone in one's clothing pockets every day is harmful to health.

    Hint: What does the IARC category "possibly carcinogenic" (2B) actually mean? What do recent systematic reviews conclude?

    Start with: R. Baan et al., Carcinogenicity of radiofrequency electromagnetic fields, *The Lancet Oncology* 12, 2011; K. Karipidis et al., The effect of exposure to radiofrequency fields on cancer risk. . . : a systematic review, *Environment International*, 2024.

14. The computing power of computers grows exponentially over time.

    Hint: Plot the data on a logarithmic scale; distinguish transistor count, clock speed and real performance.

    Start with: G. E. Moore, Cramming more components onto integrated circuits, *Electronics* 38(8), 1965; Our World in Data, `https://ourworldindata.org`.

15. Economic growth is conditioned by the availability of cheap energy.

    Hint: Correlation between GDP and energy use is strong, but what about causality and "decoupling"?

    Start with: V. Smil, *Energy and Civilization: A History*, MIT Press, 2017; H. Haberl et al., A systematic review of the evidence on decoupling of GDP, resource use and GHG emissions, *Environmental Research Letters* 15, 2020.

Mathematics

16. The golden ratio is found everywhere in nature, art and architecture.

    Hint: Many claimed examples do not survive careful measurement. With enough freedom in what you measure, almost any ratio can be found.

    Start with: G. Markowsky, Misconceptions about the golden ratio, *The College Mathematics Journal* 23(1), 1992; M. Livio, *The Golden Ratio*, 2002.

17. Benford's law can be used to detect fraud in financial or electoral data.

    Hint: When should data follow Benford's law, and when not? Test it yourself on a public dataset.

    Start with: T. P. Hill, A statistical derivation of the significant-digit law, *Statistical Science* 10(4), 1995; M. Nigrini, *Benford's Law*, Wiley, 2012; J. Deckert, M. Myagkov, P. Ordeshook, Benford's law and the detection of election fraud, *Political Analysis* 19(3), 2011.

18. Human intuition about probability is systematically wrong (e.g. the Monty Hall problem, the birthday paradox, the gambler's fallacy).

    Hint: Combine the mathematical explanation with a simple simulation and a small survey of your colleagues.

    Start with: A. Tversky, D. Kahneman, Judgment under uncertainty: heuristics and biases, *Science* 185, 1974; D. Kahneman, *Thinking, Fast and Slow*, 2011.

19. Misuse of p-values and statistical significance is a major cause of the replication crisis in science.

    Hint: Explain what a p-value is (and is not); discuss p-hacking and publication bias.

    Start with: J. P. A. Ioannidis, Why most published research findings are false, *PLoS Medicine* 2(8), 2005; R. L. Wasserstein, N. A. Lazar, The ASA statement on p-values, *The American Statistician* 70(2), 2016; Open Science Collaboration, Estimating the reproducibility of psychological science, *Science* 349, 2015.

20. Simpson's paradox: aggregated data can lead to conclusions opposite to those from subgroup data.

    Hint: Reproduce the Berkeley admissions example with numbers; find a recent example (e.g. in medical or Covid statistics).

    Start with: P. J. Bickel, E. A. Hammel, J. W. O'Connell, Sex bias in graduate admissions: data from Berkeley, *Science* 187, 1975; J. Pearl, D. Mackenzie, *The Book of Why*, 2018.

21. Computer-assisted proofs (Four Color Theorem, Kepler conjecture, proof assistants such as Lean) are as trustworthy as traditional proofs.

    Hint: What does it mean for a proof to be "verified"? Compare human peer review with formal verification.

    Start with: G. Gonthier, Formal proof—the four-color theorem, *Notices of the AMS* 55(11), 2008; T. Hales et al., A formal proof of the Kepler conjecture, *Forum of Mathematics, Pi* 5, 2017.

22. Mathematics is discovered, not invented.

    Hint: A philosophical question: a good exercise is to discuss whether (and how) it is researchable at all.

    Start with: E. Wigner, The unreasonable effectiveness of mathematics in the natural sciences, *Communications on Pure and Applied Mathematics* 13(1), 1960; R. Hersh, *What Is Mathematics, Really?*, 1997.

23. Most researchers believe that P $\neq$ NP. What evidence supports this belief?

    Hint: Explain the problem informally; distinguish proof, evidence and expert opinion (see the polls of researchers).

    Start with: L. Fortnow, The status of the P versus NP problem, *Communications of the ACM* 52(9), 2009; S. Aaronson, P $\stackrel{?}{=}$ NP, in *Open Problems in Mathematics*, Springer, 2016; W. Gasarch, The P=?NP poll, *SIGACT News*, 2002/2012/2019.

Computer Science

24. Moore's law has come to an end.

    Hint: Which "Moore's law"? Transistor density, cost per transistor and single-core performance behave differently. Links to theme 14.

    Start with: T. N. Theis, H.-S. P. Wong, The end of Moore's law: a new beginning for information technology, *Computing in Science & Engineering* 19(2), 2017; C. E. Leiserson et al., There's plenty of room at the Top, *Science* 368, 2020.

25. Quantum computers will break the cryptography currently in use (RSA, elliptic curves) within the next decade.

    Hint: Compare the resources required by Shor's algorithm with the size of current quantum computers; look at how these estimates changed over time.

    Start with: P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM Journal on Computing* 26(5), 1997; C. Gidney, M. Ekerå, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits, *Quantum* 5, 2021; NIST post-quantum cryptography standards (FIPS 203–205), 2024.

26. The choice of programming language influences the number of bugs in software.

    Hint: An excellent case study in methodology: a well-known study was later reproduced and most of its conclusions did not hold.

    Start with: B. Ray et al., A large scale study of programming languages and code quality in GitHub, *FSE* 2014; E. D. Berger et al., On the impact of programming languages on code quality: a reproduction study, *ACM TOPLAS* 41(4), 2019.

27. Open-source software is more secure than proprietary software.

    Hint: "Given enough eyeballs, all bugs are shallow" — does this hold? Use Heartbleed or Log4Shell as case studies.

    Start with: E. S. Raymond, *The Cathedral and the Bazaar*, 1999; Z. Durumeric et al., The matter of Heartbleed, *ACM Internet Measurement Conference*, 2014.

28. Forcing users to change their passwords periodically improves security.

    Hint: Official recommendations changed on this question. What evidence led to the change?

    Start with: NIST Special Publication 800-63B, *Digital Identity Guidelines*; Y. Zhang, F. Monrose, M. K. Reiter, The security of modern password expiration, *ACM CCS*, 2010.

29. Using AI assistants (ChatGPT and similar) improves students' learning outcomes.

    Hint: Distinguish performance while using the tool from learning measured afterwards without it. The two studies below point in different directions.

    Start with: H. Bastani et al., Generative AI without guardrails can harm learning: evidence from high school mathematics, *PNAS*, 2025; G. Kestin et al., AI tutoring outperforms in-class active learning, *Scientific Reports*, 2025.

30. Text generated by AI can be reliably detected.

    Hint: Discuss false positives and their consequences (e.g. for students accused of cheating).

    Start with: V. S. Sadasivan et al., Can AI-generated text be reliably detected?, arXiv:2303.11156, 2023; W. Liang et al., GPT detectors are biased against non-native English writers, *Patterns* 4(7), 2023.

31. Training and using large AI models has a significant impact on global energy consumption and CO_2 emissions.

    Hint: "Significant" compared to what? Put the numbers in context (data centers overall, other industries).

    Start with: E. Strubell, A. Ganesh, A. McCallum, Energy and policy considerations for deep learning in NLP, *ACL*, 2019; A. S. Luccioni, S. Viguier, A.-L. Ligozat, Estimating the carbon footprint of BLOOM, *JMLR* 24, 2023; IEA, *Energy and AI*, 2025.

32. Recommendation algorithms on social networks create "filter bubbles" and increase polarization.

    Hint: Popular claim vs. experimental evidence; note who funded or gave data access for the studies.

    Start with: E. Pariser, *The Filter Bubble*, 2011; E. Bakshy, S. Messing, L. Adamic, Exposure to ideologically diverse news and opinion on Facebook, *Science* 348, 2015; A. M. Guess et al., How do social media feed algorithms affect attitudes and behavior in an election campaign?, *Science* 381, 2023.

33. AI systems can diagnose diseases from medical images as well as doctors.

    Hint: Look at how comparisons with doctors are designed and whether results hold on external data.

    Start with: A. Esteva et al., Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542, 2017; X. Liu et al., A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging, *The Lancet Digital Health* 1(6), 2019.

34. Results reported on machine learning benchmarks overestimate the real performance of the models.

    Hint: Discuss overfitting to test sets, data leakage and test-set contamination.

    Start with: B. Recht et al., Do ImageNet classifiers generalize to ImageNet?, *ICML*, 2019; S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science, *Patterns* 4(9), 2023.

35. Remote work increases the productivity of programmers.

    Hint: How is productivity measured (lines of code, commits, promotions)? Compare surveys with randomized experiments.

    Start with: N. Bloom, J. Liang, J. Roberts, Z. J. Ying, Does working from home work? Evidence from a Chinese experiment, *Quarterly Journal of Economics* 130(1), 2015; N. Bloom, R. Han, J. Liang, Hybrid working from home improves retention without damaging performance, *Nature* 630, 2024.