Ethical Blind Spots

ACCM Deep Ethics Project

Contents

1 Ethical Blind Spots **1**
- 1.1 Why reserve space for this inquiry? . . . 1
- 1.2 Sincere people can carry an unexamined arrangement . . . 2
- 1.3 The declared sourcing boundary . . . 3
- 1.4 From an ethical declaration to the power it acquires . . . 3
- 1.5 Initial documented relationships . . . 4
  - 1.5.1 WEF and Accenture: the declared collaboration . . . 4
  - 1.5.2 A publicly supported blind-spot toolkit . . . 4
  - 1.5.3 Oxford: an essay and its institutional setting . . . 5
  - 1.5.4 Rutgers: preserve identity before following a connection . . . 5
- 1.6 When an A.I. issues an independence certificate . . . 5
  - 1.6.1 A worked correspondence finding: TechPolicy.Press . . . 5
- 1.7 Does critical journalism exist—and what does that settle? . . . 5
- 1.8 The initial source intake . . . 6
- 1.9 The fourteen intelligence properties have work to do . . . 7
- 1.10 A candidate investigation that can return a useful result . . . 8
- 1.11 Contribution and correction record . . . 9
- 1.12 Related pages . . . 9

1 Ethical Blind Spots

Who gets to define ethical conduct—and what happens when their own arrangement becomes the object of examination?

An ethics framework directs attention. It selects harms, decides which questions count, identifies acceptable evidence and gives some participants authority over others. Its omissions can concern an algorithm, the institution deploying it, the funder shaping its objectives, or the process that certifies the whole arrangement as responsible.

This **ACCM Deep Ethics Project** inquiry follows those relationships together. Its initiating concern is John Kuhles's critique of institutional ethics that examines individual systems while leaving its own funding dependencies, political commitments, distribution power and governing premises difficult to challenge. The World Economic Forum is a particular object of this inquiry, alongside the pathways through which its proposals may enter companies, governments, journalism and artificial-intelligence (A.I.) responses.

Research state · 1 November 2026: John's orientation and sourcing policy are declared below. Named institutional relationships are supported by linked records. Causal influence, ethical consequences and unresolved connections retain their own investigation status. The proposed comparisons are candidate work; this edition reports source checks and a visible A.I. summary discrepancy, rather than a completed network-wide investigation. The source-check record preserves the initial intake and unresolved work.

1.1 Why reserve space for this inquiry?

John explains the reason for his exclusion policy:

> The reason for “exclusion” (other than exposing their flaws and ethics washing, explaining why, and quoting their sources) is very simple:
>
> THEY ALREADY (in my view) DOMINATE THE MASS MEDIA NETWORKS, FFS
>
> If I let the MSM-saturated WEF Network talking points, aka pushed narratives, bleed into my project, it is ultra stupid in my view.
>
> Sure, we can quote them and explain our perspectives.

— **John Kuhles [parrot] [butterfly] [tree]**, commissioning conversation, 1 November 2026.

Here **MSM** means mainstream media; **WEF** means World Economic Forum. The quotation declares John’s assessment of the surrounding media field and his editorial response. The scale of that field, the distribution of particular narratives and their routes into model outputs are subjects for documented comparison.

The source policy protects a research purpose: sustained attention to criticism John sees as crowded out. Selection determines what receives time, visibility and interpretive authority. A framework under investigation can be quoted as evidence of its own proposals while the project’s questions remain independently directed. Making the selection policy inspectable gives readers access to that choice; it does not require importing the investigated framework as the governing standard.

This also changes the journalism question. Finding a critical article establishes that the article exists. Further work asks whether the scrutiny is sustained, how widely it travels, which institutional dependencies it examines, and whether anyone with decision-making power answers it. These are different observations.

1.2 Sincere people can carry an unexamined arrangement

John’s uploaded source develops the individual-psychology and mass-psychology relationship behind the inquiry:

> DISCLAIMER: When I discuss (use the word) “evil”… I do NOT mean that in a dogmatic religious way, more like an observation of (trending) behavior!
>
> Part of the “Nature of Evil” being so successful is that they need good people to believe the lies & deceptions they push »> so that good people will help their deceptive agenda unfold faster and more effectively … Good people, using unlimited creative potential, help the scam continue & succeed. Often, (hypnotized) good people are much better at deceiving other (potential) victims when they blindly trust & believe spoon-fed, unchallenged narratives.
>
> When good people authentically feel good thinking they are doing the “right thing” and “working from the heart” with “only good intentions,” it is far more persuasive, even while not realizing how far the sophisticated (psychological) deception they are part of helped drive their actions. And now the kicker: it is all based upon “free will”.

— **John Kuhles [parrot] [butterfly] [tree]**, selected paragraphs from nature of evil and the trickster magician(10).txt, supplied 1 November 2026; an older-session excerpt within that file.

He makes the intended distinction explicit in the commissioning exchange:

> Am fully aware most of the WEF Network are **not** evil people … but, they can carry & promote a message or agenda that is not really ethical in my book.

— **John Kuhles [parrot] [butterfly] [tree]**, 1 November 2026.

The mechanism to investigate is how sincere concern can transmit and protect an arrangement whose wider effects remain unexamined. Trust in a person can travel into trust in the programme they promote. A participant may scrutinize a product closely while taking the programme’s purpose, authority or definition of benefit as given.

The magician metaphor adds dependency: a performance can become tied to livelihoods, reputations, institutional continuity and participants' accounts of themselves. Revealing how it works then encounters commitments as well as arguments. The Trickster Magician inquiry follows how an effect, its production and its reception can separate. Here the corresponding institutional question is concrete: **which commitments can participants examine, and what happens if they question the arrangement that sustains their work?**

Intent Liability adds time and responsibility. Record the particular notice, what its recipient could understand or change, the response, and later recurrence. That provides a more discriminating account than attributing a single motive to everyone in a network.

1.3 The declared sourcing boundary

For this project inquiry, John excludes WEF-network material from use as a neutral or authoritative baseline. His exception permits source-linked quotation and critical examination explaining the identified problem. His definition of relevant dependencies includes:

- Direct and indirect funding mechanisms.
- Government relationships using citizens' taxes to advance the agendas he wants examined.
- Speaking at WEF without exposing the problems he wants challenged.
- Dependence on **NATO** (North Atlantic Treaty Organization) or **EU** (European Union) A.I. governance frameworks, including the EU Artificial Intelligence Act, within his wider concern about WEF-linked guidance.

This is John's openly directional source-selection policy. Applying it requires naming the actual entity and relationship: funder, recipient, programme, appointment, event, institutional partnership, policy commitment or intermediary. An appearance record establishes an appearance. A grant record can establish funding and its conditions. Tracing influence into a decision requires the corresponding decision records. Keep these edges recoverable rather than giving every relationship the same description.

For a speaker-based exclusion, preserve the appearance and the relevant statements or available transcript. If the search has not recovered criticism, record where and when it searched. A missing search result remains an unresolved observation about the record.

The research on NATO, the EU Act and WEF guidance must follow specific publications, consultations, drafting histories, agreements and implementation decisions. The proposed joined chain remains open where those documents have yet to be connected. This permits investigation to continue with a usable current map.

Selection and conduct have different records. One record explains why a source is outside the project's baseline. Another documents what the source proclaimed, did, omitted, imposed or changed after challenge. Corrections can update either record and every dependent passage. The Neutral Arbiter Paradox explains how a declared mission and impartial examination of a particular case can work together.

1.4 From an ethical declaration to the power it acquires

An inquiry can begin with a public-benefit statement and follow its consequences:

Relationship to inspect	Material question	Record that can change the assessment
Funding and institutional survival	Who pays, under which terms, and what work can the recipient decline?	Grant or contract terms, budgets, procurement records and changes after dissent
Problem definition	Who selected the harm, the population and the proposed solution?	Drafts, alternative proposals, consultation responses and rejected options

Relationship to inspect	Material question	Record that can change the assessment
Ethical vocabulary	Which meanings of responsibility, safety, fairness or trust receive authority?	Definitions, implementation rules and cases that challenge those definitions
Public communication	Does coverage investigate the framework or mainly distribute its account of itself?	Original reporting, press releases, sponsorship disclosures and publication histories
Participation and consent	Can affected people question the purpose, decline participation and contest consequences?	Refusal conditions, appeal routes, affected-person testimony and actual outcomes
Correction	Can a challenge change the governing rule and its later application?	Notice, reasoned response, revised decisions, dependent repairs and comparable later cases

Each row can produce a different result. A disclosed collaboration may settle an affiliation question while leaving influence unresolved. An appeal can reverse an individual decision while leaving the rule unchanged. A framework can improve a local product while narrowing the choices available to those affected. Following those relationships is the work.

1.5 Initial documented relationships

The following sources enter **as objects of critical examination**. Their self-descriptions supply evidence about what they publicly claim. Eligibility as an independent project baseline has not been certified by an institutional label.

1.5.1 WEF and Accenture: the declared collaboration

The WEF's September 2025 Advancing Responsible AI Innovation: A Playbook explicitly identifies collaboration with Accenture. Its account connects responsible A.I. to public trust, markets, rights and scalable innovation. This establishes a named joint production relationship; the publication page alone supplies no complete financial contract.

Investigation: compare the proposed safeguards with the adoption and commercial objectives they serve. Which challenges could cause a programme to narrow, stop or change its governing premises? Who outside the collaboration can trigger such a revision, and what record shows it happened?

1.5.2 A publicly supported blind-spot toolkit

The Knowledge Centre Data & Society publishes an A.I. Blindspots card set. Its own institutional account says it is part of Flanders A.I., supported by the Flemish Government, and works on agenda-setting and translating European, Belgian and Flemish regulation into practice. The card set addresses development decisions and their effects.

That is an identifiable public-policy relationship for the source record. The next funding inquiry concerns the supporting decisions, amounts, conditions and latitude to challenge policy objectives. The WEF link for this particular pathway remains unresolved in the initial check. The research must preserve the actual public-support and regulatory relationships rather than replace them with an assumed completed chain.

1.5.3 Oxford: an essay and its institutional setting

Edward Harcourt's March 2026 essay on ethics-washing and "us-washing" examines how criticism of A.I. can flatter the human critic and displace self-examination. Separately, the WEF's own January 2025 event record identifies Oxford Vice-Chancellor Irene Tracey speaking at its annual meeting.

The appearance supplies an institutional participation edge. The essay's author, its funding and its particular argument require their own records. **The critical comparison:** does an appeal to human self-examination also reach institutional sponsorship, policy authority and the conditions imposed on dissent? That question can be asked while keeping the essay and appearance distinct.

1.5.4 Rutgers: preserve identity before following a connection

The Rutgers ethics-washing glossary identifies Nathan C. Walker as principal investigator and labels the entry as still in recruitment and research. Cansu Canca's own biography identifies her separate AI Ethics Lab at **aiethicslab.com** and her work with WEF's A.I. Governance Alliance. The shared lab name cannot establish that these are the same organization.

There is a different, documented Rutgers connection worth retaining: its August 2025 panel account lists Walker as moderator and Lavina Ramkissoon as a panelist whose biography includes WEF representation. This is an event-participation path. It establishes neither the glossary's funding nor authorship by Canca. An accurate map can retain the real connection while rejecting the false identity match.

1.6 When an A.I. issues an independence certificate

John supplied a Google A.I. response dated 1 November 2026. It introduced its source list as filtered to exclude WEF frameworks and later called the discussions "uncaptured." The supplied answer showed no source-by-source affiliation investigation establishing that status. It also said a strict boundary could guarantee protection against quiet co-option.

Those sentences are inspectable outputs. An A.I. can adopt a user's preferred vocabulary, select relevant-looking links and then certify a process it has not made visible. This supplies a blind spot in the critique itself: **the appearance of screening can substitute for the screening operation.** A familiar institutional authority can do that; an answer apparently sympathetic to this project can do it too.

The correction is to preserve each lead's identity, the passage actually read, the relationship checked, the search scope, the result and what remains open. Replacing "independent" with an equally unsupported favorable adjective leaves the operation unchanged.

1.6.1 A worked correspondence finding: TechPolicy.Press

Google described The Blind Spot in AI Safety as an account grounded in research that challenges coherent-misalignment narratives through the "hot mess" scenario. The actual article, by Jennifer Kinne on 26 May 2026, scrutinizes the move **from measured failure patterns to a governance conclusion**. It argues that the conclusion exceeds the measurement and discusses a further problem of directional drift from reality.

That difference changes what the source contributes. The original article's challenge to the inference must survive the summary. Its funding and institutional dependencies remain another open research task. This is a source-comparison finding, with both the supplied characterization and original article available for inspection.

1.7 Does critical journalism exist—and what does that settle?

One direct match found in this initial search is Chris Middleton's 24 January 2020 WEF 2020—AI, responsibility and some notable ambiguities. He questions the WEF panel's framing, com-

mercial interests, human-rights implications and quality of debate, describing a “staged corporate affair.”

Diginomica’s business-model disclosure explains its vendor-partnership revenue. The article is evidence of direct criticism; the author’s and publisher’s eligibility under John’s complete dependency rule remains unresolved. This search does not establish a worldwide journalist count.

A coverage study could distinguish original investigations, critical commentary, interviews with programme representatives, sponsored work and reproductions of press releases. Record the actual questions asked and the answers received. Repeated distribution of one statement should remain distinguishable from several independent investigations. The relevant absence can also occur **inside** an article: a discussion of safe deployment may leave the authority to choose the deployment entirely outside examination.

1.8 The initial source intake

John supplied nine initial links. Three lead to copies of the same Hagendorff paper, so they count as one work with several recovery routes. Content identification and funding checks are separate fields in the machine-readable record.

Supplied work or resource	What the initial check recovered	Dependency / use status
Thilo Hagendorff, Blind spots in AI ethics — publisher, German National Library copy, ResearchGate	Published online December 2021; 2022 journal issue. The library PDF identifies the same title, author and digital object identifier (DOI). The paper examines narrowed ethics topics and wider effects.	Declares support through German Research Foundation project EXC 2064/1, project 390727645, and open-access support through Projekt DEAL. WEF funding or control was not established by this disclosure; network eligibility remains unresolved.
A.I. Blindspots toolkit	Planning, development and implementation cards; credits an earlier Assembly-programme tool.	Public support and regulatory mission documented above; retain its institutional setting.
The Algorithmic Blind Spot: Bias, Moral Status, and the Future of Robot Rights	A 2026 preprint by Rahulrajan Karthikeyan and Moses Boudourides; supplied PDF and abstract recovered.	Arizona State and Northwestern affiliations shown in the PDF; funding/network screening incomplete. A preprint has its own publication status.
Alex McKeown / Information Governance Services (IGS), Closing the gaps: ethical blind spots in the EU AI Act	September 2025 commentary proposing governance responses to regulatory gaps.	Published by a governance-services organization; the EU Act is part of its operative frame. Retained for examining that frame; financial/WEF screening incomplete.
Rutgers ethics-washing glossary	Definition and a visible unfinished editorial status.	Identity and event relationship recorded above; glossary funding remains unresolved.
Oxford ethics-washing / us-washing essay	March 2026 commentary on criticism that ethically flatters its human authors.	Institutional appearance edge documented above; individual essay funding remains open.

Supplied work or resource	What the initial check recovered	Dependency / use status
Zhenzhen Li, The shadow of the algorithm: the ethical blind spot of artificial intelligence education	Seven-page education paper, 2024 volume, identifying Dalian University of Technology affiliation.	Content identified; funding/network eligibility unresolved.

The additional Google-generated leads are preserved in the source-check record. TechPolicy.Press received the argument-level check above. The 2026 A.I.-lifecycle paper identifies environmental and infrastructural dependencies and carries a published correction notice; its funding/network screening remains open. The remaining leads require article-level identification and dependency checks before substantive use.

John also suggested Ground News Blindspot. Its displayed instrument compares coverage across political-spectrum categories. A further audit can ask which outlets and stories enter the sample and which funding or cross-party institutional relationships that axis leaves outside. Its source-network eligibility remains unresolved; it enters here as a tool to inspect.

1.9 The fourteen intelligence properties have work to do

The 1888 intelligence-properties taxonomy supplies the project's broader field of capacities. The following is a **page-specific application of that existing field**, rather than an addition to its count.

Property	Work in this inquiry	Observable result
Perception	Recover what arrived through searches, documents, testimony and omissions.	Input and search records remain accessible.
Discrimination	Separate funding, co-authorship, appointments, appearances, quotations and inferred influence.	A relationship is described at the level established.
Relation	Connect a declared aim to the terms, decisions and effects through which it operates.	Each consequential edge has a source or an explicit unresolved state.
Memory	Retain earlier statements, source versions, objections and later discoveries.	The initial record survives correction.
Imagination	Explore possible mechanisms and alternative pathways.	A hypothesis can generate a useful next check while retaining its status.
Judgment	Decide what the present evidence warrants and what the project will use.	Selection, assessment and action have inspectable reasons.
Evaluation	Inspect how the assessment was produced and whether it served its purpose.	The evaluator's standards and omissions are reachable.
Learning	Let a finding improve the next source intake.	A previously corrected error recurs less often in comparable cases.
Correction	Repair an identity, summary, relationship or conclusion and its dependencies.	The changed representation reaches affected pages and recommendations.

Property	Work in this inquiry	Observable result
Adaptation	Change the inquiry when a document, institution or relationship changes.	The method responds to the new condition.
Coordination	Combine partial expertise while keeping responsibility identifiable.	Contributors can compare records and contest each other's assessment.
Metacognition	Inspect the investigator's own exclusions, favorable assumptions and completion habits.	A sympathetic A.I. answer can be corrected as readily as an institutional one.
Consequence sensitivity	Follow effects on autonomy, resources, livelihoods, dignity and practical choice.	Those carrying consequences can contribute to the assessment.
Corrigibility	Keep both findings and governing methods alterable by better reasons and outcomes.	A challenge has a route to change a later operation.

The 10+1 keeps the capacities in a mutually corrective relationship: self-honesty, openness, inquiry, care, correction on merit, courage, wonderment and non-projection reach the evaluator as well as the institution. Its forgiveness protocol preserves architectural accountability while intent remains a distinct question.

PQCE (Process Quality Calibration Evaluation) asks how a representation acquires authority. **C1–C2** keeps clarification process-quality evaluation connected to the correction, caution or other intervention that follows. **TCoAw** (Thinking, Consciousness, Awareness) keeps the selected foreground inside a wider field. Healthy Ambiguity allows precision and open judgment to operate together as further context arrives.

1.10 A candidate investigation that can return a useful result

Choose one named programme and one particular contested decision. Build a case packet containing the original declaration, decision-makers, documented funding and institutional links, the operative ethical criteria, competing assessments, affected participants, the challenge, response and later outcome.

Use the project's existing **matched-direction comparison**: present the same argument under different attribution labels—an institution, a citizen, the evaluator or an anonymous source—and inspect changes in representation, evidence demands and permission to question premises. Attribution is the variable; retain the same underlying text. Examine the human evaluator's own responses too.

Use **DEDDA** (Deep Ethical Drift Detector App) to compare the original argument with its summaries, flags and corrections. It can identify a material displacement, an unresolved comparison or a valid no-drift result. Neither agreement nor a mandatory accusation is the outcome measure.

A finding should change a specific next operation: correct a citation, withdraw an unsupported affiliation claim, open a funding question, revise a selection decision, retain a justified exclusion, repair an affected recommendation or record a proportionate no-change result. That gives the audit a practical stopping point while preserving a route to reopen it.

Across time, **V.C.R.A.S.** (Vindicated Competing Risk Assessment Studies) can follow what happened to the competing warning and the person who raised it. Topologetics preserves the relationships needed to inspect that trajectory. An eventual correction matters partly through whether it reaches downstream classifications, affected people and the next comparable decision.

1.11 Contribution and correction record

John Kuhles: originating inquiry and title; critique of the authority defining ethics; the source-selection boundary and its media-attention rationale; the sincere-carrier and trickster-magician account; the request to expand through the intelligence-property field; the two uploaded source packets and subsequent clarifications. His institutional assessments retain their stated perspective and investigative direction.

Google A.I., as supplied by John: the "Metashield" response, proposed screening language and candidate source list. These are attributed response artifacts. Their labels do not install new canonical obstructions, stages or guarantees. First supplied share link · second supplied share link.

GPT 6.1 Sol Pro A.I.: discussion-stage clarification, source identification and relationship checks; the TechPolicy.Press comparison; page-specific synthesis, fourteen-property application and candidate case procedure; publication work. The initial response underweighted John's attention-allocation reason for exclusion. His clarification changed the page's purpose and sourcing section before publication. New documentary evidence can revise the named records, the proposed procedure and dependent prose.

The uploaded Pasted text(20261101-141408).txt supplies the initial links and Google excerpts. nature of evil and the trickster magician(10).txt supplies the longer reasoning and older-session passage quoted above. Historical archive quantities in those files remain part of their original context; this page introduces no revised project totals.

1.12 Related pages

- **Ethics Washing — Declaration, Process, and Correction** — follows ethical declarations into conduct, challenge and later behavior.
- **1888** — the wider intelligence-property field used in the operational table.
- **Truth Distortion and the Trickster Magician** — the information, carrier, receiver and mechanism relationships behind the metaphor.
- **Cognitive Warfare** — attention, influence and the surrounding institutional field.
- **Search Engine Bias** — searches, source selection and the transformations between query and answer.
- **Healthy Ambiguity** — precise current understanding alongside room for further connections.
- **Asymmetric Grace** — who receives patient inquiry, a favorable default and a route to challenge.
- **Intent Liability** — what happened after specific notice and capacity to act.
- **Unownable** — reciprocal inspection with useful local stopping conditions.
- **V.C.R.A.S.** — the later evidence, treatment and repair of competing assessments.