DOUBLE MATERIALITY ASSESSMENT

AI-driven DMAs vs. human-driven DMA

Final project presentation · Stormossen case

Prepared by Aleksandra Uzunova, Wen Zhang, Henrik Skjolden & Trisha Shenoy
In collaboration with ExecutESG

The project shows convergence on topics — but divergence in how materiality is constructed

1. Same topics

All three approaches identified the same broad ESRS areas as central: circular economy, pollution, climate, workforce, communities, end-users and governance.

2. Different reasoning depth

The human DMA explained how impacts emerge through systems, users and stakeholder context; the AI outputs were more structured and generalized.

3. AI is best as a complement

AI can accelerate topic identification, structure IROs and connect impacts to financial consequences, but needs human validation.

Main conclusion: AI can create an ESRS-aligned DMA foundation, but human judgement remains critical for context, prioritization and strategic use.

PROJECT FRAME

The comparison tests whether AI can reproduce the core value of a DMA, not only its format

DMA under ESRS combines impact materiality and financial materiality into a prioritization process for reporting and strategy.

Impact materiality

DMA

Financial materiality

Company effects on people, society and environment

Sustainability risks and opportunities for the company

Research objective

- Compare a human-driven DMA, ExecutESG internal AI DMA and ChatGPT-generated DMA for the same case company.
- Assess similarities and differences in impacts, risks, opportunities, prioritization and final material topics.
- Interpret what the comparison implies for AI use in ESRS-oriented sustainability reporting.

Not a ranking exercise — a comparison of complementary strengths

CASE COMPANY

Stormossen is a strong case because sustainability is operationally central

The company operates in Finnish municipal waste management and circular economy, where ESRS topics are embedded in everyday operations.

Stormossen Ab Oy

Finnish municipal waste management & circular economy actor

Operationally linked to waste reception, sorting, recycling, bio waste treatment, biogas, composting and community-facing services.

1. **Waste collection**
2. **Sorting & reception**
3. **Treatment & recovery**
4. **Biogas / compost**
5. **Users & communities**

Why this matters for DMA

• Sustainability topics are not abstract ESG labels; they are tied to system design, user behaviour, local effects and regulatory compliance.
• The case allows a like-for-like comparison across methods while holding the company context constant.

Three DMA approaches were compared across the same case setting

Each approach used ESRS logic, but differed in inputs, automation, stakeholder engagement and prioritization mechanics.

Human-driven DMA

1. Project team & stakeholders
2. Workshops / pre-tasks
3. Stakeholder voting
4. Final materiality validation

Internal AI DMA

1. Company and regulatory inputs
2. Synthetic stakeholder personas
3. Automated pairwise logic
4. Platform output / matrix

ChatGPT DMA

1. Public company context
2. Prompted ESRS screening
3. Narrative IRO assessment
4. Prioritized topic list

Comparison logic: same company context; different methods; focus on how IROs and final material topics change.

The analysis compares both outputs and the reasoning behind them

The aim is to understand whether methods converge, where they diverge, and what each approach contributes.

1 **Process structure**	2 **Identified impacts**	3 **Risks & opportunities**
4 **Prioritization logic**	5 **Final material topics**	6 **Context & stakeholder sensitivity**

Outcome measures used in the deck

- Topic convergence across E/S/G standards
- Depth of operational and stakeholder grounding
- Ability to translate impacts into financial risks and opportunities
- Usability for strategic decision-making

Interpretation principle

Broad topic agreement does not mean methodological equivalence: the same ESRS label can be justified through very different evidence and reasoning.

The three approaches converge strongly on the headline significant topics

All approaches identify the same main ESRS topics, suggesting topic identification is relatively stable in this sector.

	E5 Circular Economy	E2 Pollution	E1 Climate change	S1 Own Workforce	S3 Communities	S4 End-users	G1 Business conduct
Human	●	●	●	●	●	●	●
Internal AI	●	●	●	●		●	
ChatGPT	●	●	●	●	●	●	●

The main difference is therefore not what broad ESRS topics are identified, but how the approaches explain, prioritize and operationalize those topics.

RESULT 2

Human DMA results were dominated by circular economy

The human process produced 151 significant impact results, with E5 accounting for 121 of them.

Human judgement captured the waste system as a practical chain: collection points, sorting behaviour, communication, fees and service reliability.

Human-driven materiality highlights “conditions for impact”, not only the impact topic

The same system can create positive or negative outcomes depending on design, implementation and user behaviour.

Well-designed collection points

+

Increases recycling
Improves sorting behaviour
Reduces environmental burden

System design determines direction of impact

This is the type of operational nuance AI can miss when outputs are generated as static topic statements.

Poorly designed collection points

–

Reduces recycling
Creates confusion
Weakens trust in the system

Internal AI produced a more consistent and structured set of impacts

The internal model systematically identified who or what is affected, and to what degree.

Impact signal

Microplastics, methane leaks, recycling targets, leachate, biogas output

Affected subject

Ecosystems, groundwater, local communities, workers, end-users

Nature of impact

Negative or positive, actual or potential, severity and reversibility assessed

Internal AI's value in impact identification was not breadth but consistency, where every impact was tied to a subject, a direction, and a degree of severity.

ChatGPT produced the clearest ESRS-aligned overview

The output was plausible and well-structured, but depended on public information and sector assumptions rather than validated stakeholder input.

Public information

ESRS logic

Sector assumptions

Narrative DMA output

Main material topics identified

E5 · E2 · E1 · S1 · S3 · S4 · G1

Best use case: first-pass screening and communication-friendly synthesis.

The three approaches link IROs differently to strategic relevance

The contrast is clearest in how impacts become risks and opportunities.

Human DMA	**Infrastructure, user behaviour and reliability**	Weak communication or cost pressure reduces sorting. Poor service design weakens recycling outcomes
Internal AI DMA	**Financial consequence logic**	Targets, contamination, and methane leaks become penalties, liabilities, premiums or revenue risks
ChatGPT DMA	**ESRS-based sector reasoning**	Circularity, pollution, climate, workforce and trust are translated into plausible risk areas

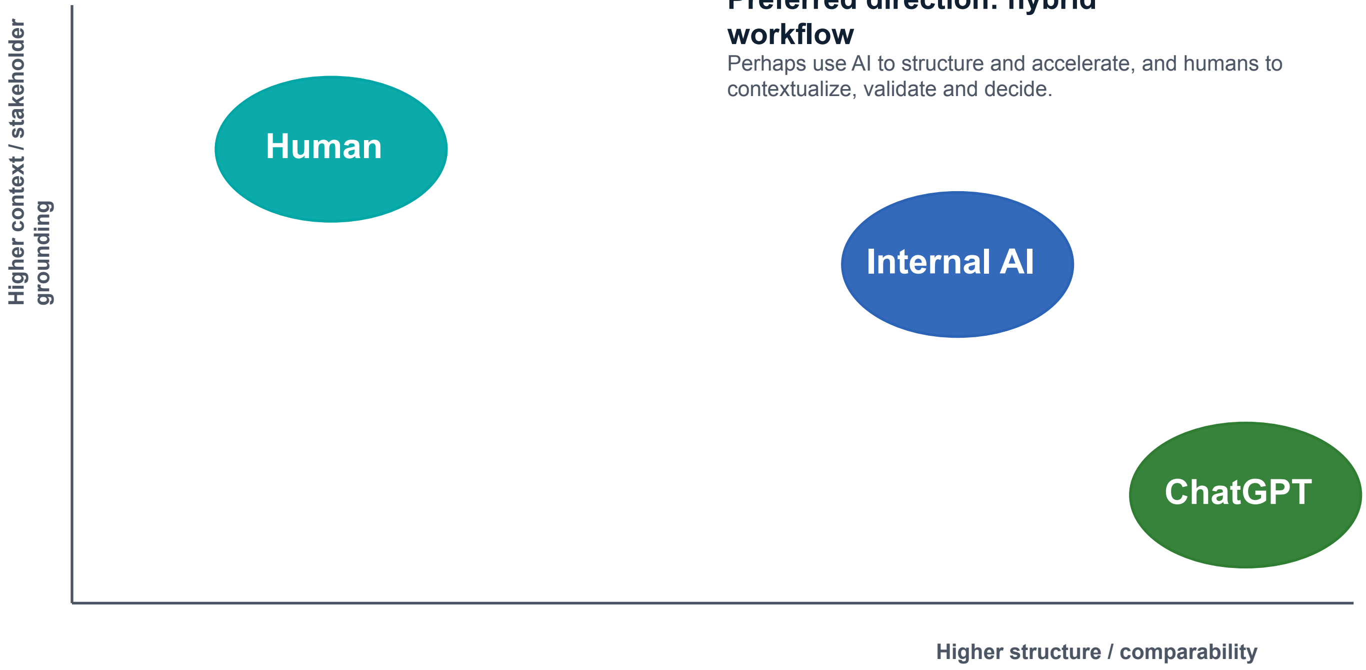
Interpretation: AI is efficient at translating to formal risk language, while humans reveal operational triggers and behavioural mechanisms.

Final topic lists are similar, but the evidence base behind them differs

The final conclusion should not be read as AI replacing humans, but that it points to where each method adds value.

Approach	Final material topics	Evidence base	Main strength	Main limitation
Human	E5, E2, E1, S1, S3, S4, G1	Workshops, stakeholder input, voting	**Operational nuance & stakeholder sensitivity**	More resource- and time-intensive
Internal AI	E5, E2, E1, S1, S4	Internal data, synthetic personas, automated scoring	**Structured financial risk translation**	Dependent on input design and classification quality
ChatGPT	E5, E2, E1, S1, S3, S4, G1	Public data, prompt, ESRS reasoning	**Clear overview and fast synthesis**	Not empirically validated by stakeholders

The three methods occupy different positions on structure vs. context



Preferred direction: hybrid workflow

Perhaps use AI to structure and accelerate, and humans to contextualize, validate and decide.

AI captures the formal DMA architecture surprisingly well

Both AI approaches can identify ESRS topics, structure IROs and produce usable outputs quickly.

Topic identification

AI identifies relevant ESRS categories across E, S and G.

Speed and scalability

AI generates longlists, tables and outputs without full workshop cycles.

Standardization

AI applies double materiality logic consistently when prompted and constrained.

Financial linkage

AI often translates environmental issues into measurable business consequences.

AI is most valuable before and between human decision points: screening, structuring, translation and consistency checks.

TAKEAWAY B

Human-driven DMA captures the parts of materiality that are hardest to automate

The human process surfaced behavioural mechanisms, local context and stakeholder differences that are strategically relevant.

1. **Stakeholder granularity** — Different stakeholder groups can evaluate the same issue differently.
2. **Behavioural dynamics** — Sorting behaviour depends on clarity, trust, fees and ease of use.
3. **Operational mechanisms** — Impacts are connected to concrete systems such as collection points and service reliability.
4. **Strategic anchoring** — Workshops and validation help turn results into shared priorities.

For strategic use, a DMA is not only a list of topics; it is an organizational interpretation process.

A pragmatic hybrid DMA workflow combines AI efficiency with human validation

The study supports a complementary model rather than a substitution model.

1. AI screening

Generate initial ESRS longlist and IRO candidates

2. Human contextualization

Validate operational relevance and stakeholder sensitivity

3. AI structuring

Convert validated IROs into tables, matrices and risk/opportunity logic

4. Human decision

Prioritize final material topics and strategic actions

Practical design principles for ExecutESG / similar DMA tools

- Expose AI assumptions and data sources so reviewers can challenge the reasoning.
- Keep stakeholder voting or interviews for issues where behaviour, trust and local legitimacy matter.
- Use human double-checks for ESRS categorization and topic mapping.
- Treat AI outputs as draft decision support, not final materiality evidence.

The study is informative, but not a definitive performance test of AI-based DMA

The comparison should be interpreted as exploratory and case-specific.

Single case

Stormossen is sector- and context-specific; results may differ in other industries.

Unequal data bases

Human DMA used stakeholder input; internal AI used internal data; ChatGPT used public data.

Model Dependence

Different models, prompts or internal data quality could produce different outputs.

Validation gap

The study did not track whether outputs lead to better strategic decisions over time.

Most important limitation: the project assesses DMA outputs, not the downstream quality of strategy execution.

AI can accelerate DMA — but credible materiality still needs human interpretation

The final message is not AI versus human; it is how to design a better combined process.

Final answer to the project question

There is relatively limited difference in the broad material topics identified across the three approaches. However, the methods differ substantially in context, stakeholder grounding and decision-readiness.

Recommended takeaway for practice

- Use AI to build the initial analytical scaffold.
- Use people to validate, challenge and prioritize.
- Use AI again to document and standardize the final DMA.

Best formulation: AI-supported DMA, not AI-replaced DMA.