

# Robots in the Wild

## Report I: Autonomy in Robots: Between Myths and Reality

*By Pernille Maja Carlsen, Peter Hommel Østerlund and Cathrine Hasse*

### Introduction

This report is the first of five presenting studies of ‘robots in the wild.’ As part of the project RoboSAPIENS a group of anthropologists have studied how close the robots of today are to autonomy when studied In-the-wild settings. The purpose of this study of robots at workplaces, in fields, at sea, and at public places like restaurants and shops is to explore real human-robot interactions as contrasted with the more staged versions found at conferences and in media imaginaries. The findings of the first four reports are summed up in Report V where we present the overall conclusions and discuss what these studies can tell us about a future with RoboSAPIENS technologies.

For as long as humans have imagined machines, we have imagined them moving on their own. From ancient myths of self-acting statues and mechanical servants to contemporary visions of intelligent robots that work, care, and decide independently, autonomy has been one of the most powerful ideas attached to technology. It is an idea filled with promises of efficiency, safety, and freedom from human labour but also with anxiety. What happens when the made begins to move by itself?

Today, this ancient dream is no longer confined to stories. In laboratories, companies, and research projects such as RoboSAPIENS (<https://robosapiens-eu.tech/>), engineers are actively working to develop robots that can adapt to their environments, learn from experience, and make decisions without continuous human control. In public discourse, these developments are often presented as signs that fully autonomous robots are just around the corner. Promotional videos, conference demonstrations, and media coverage suggest that robots are on the verge of becoming independent actors in everyday life. Yet when robots leave the laboratory and enter real-world settings, a different picture emerges.



Section 1: Our methodology	3
Section 2: A Short History of Imagined Independent Machines	7
Section 3: Future Autonomy: From R2-D2 To Optimus	11
Section 4: Robot Conferences: Performing Autonomy	17
Section 5: How Engineers Define Autonomy	21
Section 6: The Prophetic Position: Anthropology at The Threshold	25
Conclusion	26
References	27



# Section 1: Our methodology

## Our Anthropological Method: Studying Robots In-the-wild

This report series is based on a simple but unusual premise: we study robot autonomy before it fully exists. While most evaluations of autonomous systems focus on finished technologies or laboratory benchmarks, our approach asks a different question: how is autonomy already being imagined, enacted, negotiated, and maintained In-the-wild settings where robots are present but not yet autonomous? Anthropology is particularly well suited to this task. Anthropology offers tools for studying technologies not only as technical artefacts, but as social and cultural actors embedded in everyday life. Rather than asking whether robots meet predefined criteria of autonomy, we examine how autonomy is understood, performed, supported, and contested by the people who encounter robots in practice.

This shifts attention away from idealised futures and toward the messy realities in which robots already operate. Fieldwork In-the-wild settings allows us to take seriously moments that is often treated as peripheral in presentations of robots: breakdowns, workarounds, misunderstandings, frustrations, jokes, acts of care, and even sabotage. These moments reveal how autonomy is not simply designed into machines, but emerges, or fails to emerge, through interaction. By focusing on lived experience, anthropology shifts attention from what robots are supposed to do to what they actually do in specific contexts.

## Studying Futures

The RoboSAPIENS project will not deliver a fully autonomous robot by the end of its funding period (end December 2026). Instead, it develops architectures, principles, and simulations intended to move robotics closer to greater autonomy. From an anthropological perspective, this is not a limitation but an opportunity. This places us in a distinctive position: we are studying a future technology that is still in the making, while simultaneously observing how partial, limited, and negotiated forms of autonomy already shape everyday human-robot relations.

We therefore adopt a double strategy. On the one hand, we follow the engineering work within RoboSAPIENS, where autonomy is defined, modelled, simulated, and tested. On the other hand, we conduct multi-sited ethnographic fieldwork in more than fifteen real-world locations where robots are already embedded in daily routines. This allows us to move between imagined futures and present realities, between what autonomy is supposed to become and what it currently requires functioning at all.

Although our field sites are related to the four RoboSAPIENS use cases (see <https://robosapiens-eu.tech/>), these reports do not evaluate the use cases themselves; such an evaluation will be conducted at the end of the project. We do not ask whether a given use case succeeds or fails. Instead, we use fieldwork in and around these domains to identify what repeatedly proves important, and difficult, when engineers work toward autonomy across contexts.

Our focus is therefore comparative and transversal. We ask: What conditions repeatedly challenge autonomy? Where does autonomy break down? What kinds of human work become necessary? These questions cut across use cases rather than staying within them.



When we conduct fieldwork both inside engineering environments and in workplaces, farms, hospitals, factories, restaurants, and public spaces where robots are already present, we get a sense of how far reality is from future visions. By observing how robots are used, maintained, corrected, and sometimes resisted, we study autonomy not as a finished technological achievement, but as an ongoing and fragile process. In these settings, we encounter robots that may appear autonomous at first glance, but we also find across cases that robots always rely heavily on human support. Robots that deliver goods need people to clear their paths. Agricultural robots must be reprogrammed when soil conditions change, or animals behave unexpectedly. Industrial robots stop, crash, or wait when sensors fail. Navigation systems freeze during storms until humans intervene. We also encounter humans who adapt to robots by slowing down, stepping aside, repairing, monitoring, or sometimes deliberately blocking and testing them. Far from acting independently, today's robots exist within dense networks of human care, attention, and responsibility.

This observation raises the central question for this series of reports: What does autonomy in robots actually mean in practice and what might it come to mean in the future?

Crucially, this is not a question that can be answered by waiting for fully autonomous robots to arrive.

Studying autonomy before it fully exists allows us to observe how future-oriented technologies already shape present practices. Expectations about autonomy influence how robots are designed, marketed, regulated, and used even when the technology itself remains partial. By examining robots that are “almost autonomous,” we can trace how autonomy is anticipated, staged, compensated for, or quietly delegated to humans.

This places our work in what we, with the anthropologist Edwin Ardener, describe as a prophetic condition: an analytical position between worlds (Ardener 2017/1975), where emerging futures can be sensed through present tensions, adjustments, and frictions as we learn from the fieldwork (Hasse 2015). Prophecy here does not mean prediction, but attentiveness to how futures are actively being assembled in the present. It is the effort to give voice to transformations that are already taking shape, even if their final form remains unclear. Our analyses are written from within this in-between space, where myths of autonomy meet the mundane work of keeping robots running.

## **A Multi-Variation Anthropological Approach**

Methodologically, this study builds on multi-sited ethnography (Marcus 1995), while extending it through what we define as a multi-variation anthropological approach. Rather than following a single robot, organisation, or bounded field site, we conducted ethnographic fieldwork across more than fifteen sites in several European countries, including Denmark, Germany, Spain, Norway, and Sweden. These sites span a wide range of contexts, including robot conferences, farms, factories, hospitals, educational institutions, public spaces, and private companies.

Inspired by insights developed in the earlier EU-funded REELER project, the multi-variation approach deliberately seeks heterogeneity across cases to identify patterns that emerge across difference rather than within a single local context (Hasse 2019). By systematically comparing ethnographic material from diverse environments, robot types, and institutional settings, this



approach enables the identification of recurring structural dynamics that cannot be reduced to local contingencies or individual cases.

The analytical strength of this approach lies precisely in variation. When similar patterns recur across highly different settings, they gain robustness not because the contexts are shared, but because the patterns persist despite contextual differences. In this sense, the multi-variation approach supports a form of analytical generalisation grounded in ethnographic comparison rather than statistical representativeness.

## **Ethnographic Practice and Data Production**

Across these sites, we conducted participant observation (DeWalt & DeWalt 2010), an approach in which researchers learn cultural meanings, practices, and skills through close engagement with everyday activities. Fieldwork consisted of short and repeated visits, combined with on-site interviews. This strategy aims to produce thick description (Geertz 1973), access local perspectives, and generate grounded explanations through sustained, reflexive engagement rather than hypothesis testing (Bernard 2006).

This approach allows us to engage with and understand the situated practices in which robots are implemented and used (Bruun et al. 2015). Data production relied on systematic recording through contemporaneous field notes, including detailed accounts of interactions, material settings, event sequences, and reflexive commentary. Such documentation is central to producing reliable ethnographic data that explicitly acknowledge the researcher's learning process in the field (Hasse 2015).

From an anthropological perspective, knowledge produced in local sites is understood as enacted. Understanding emerges through skilled, sensuous engagement and forms of apprenticeship within ongoing practices. Fieldwork is therefore treated as a form of education and craft-based learning rather than purely observational data collection (Ingold 2017). Accordingly, our analysis emphasises processes, practices, and situated interpretation rather than statistical generalisation (Bernard 2006).

Ethical and reflexive commitments are integral to this work. These include careful negotiation of access and consent, sensitivity to power dynamics, and ongoing attention to how researcher presence shapes the field (DeWalt & DeWalt 2010). All fieldwork was conducted in accordance with established ethical guidelines for anthropological research. Interview participants were informed about the purpose of the study and provided informed consent.

To protect individuals and organisations, data from workplaces have been anonymised and depersonalised. Company names, locations, and personal identifiers have been removed or altered. This decision is both ethical and analytical: our aim is not to evaluate or expose specific organisations, but to identify patterns that illuminate broader dynamics of robot autonomy across contexts.

## **Variation, Comparison, and Robots “in the Wild”**

While classical anthropology emphasised long-term immersion in a single local setting, the multi-variation approach prioritises variation in observational learning processes. Multi-sited ethnographic writing enables the identification of patterns that recur across otherwise different



settings (Hasse 2019). When similar issues, such as safety concerns, trust, unexpected behaviour, or human intervention, appear in multiple locations, they gain analytical significance.

The five reports deliberately combine multiple types of settings in which robots operate “in the wild”: robot conferences, public spaces, organisational contexts, and dynamic everyday workplaces. Conferences function as staged environments in which autonomy is performed, promoted, and imagined. Here, robots are often presented by marketing and sales representatives rather than engineers, making visible how autonomy is supposed to appear and what audiences are encouraged to believe robots can do.

Workplaces, by contrast, reveal what autonomy requires to function over time: maintenance, adaptation, human oversight, and continuous negotiation. By moving analytically between these settings, we trace both the gap and the connections between imagined autonomy and lived autonomy. This comparative strategy allows us to examine how conference imaginaries translate into workplace expectations, and how everyday experiences feed back into design practices, marketing narratives, and regulatory discussions.

## Grounding the Prophetic Perspective

Field sites were selected based on their relevance to the RoboSAPIENS project’s four use cases and design ambitions. We focused on robots operating in environments comparable to those addressed by the project, including logistics, navigation in dynamic environments, and human-robot collaboration in shared spaces.

Rather than seeking “best-case” examples, we prioritised sites where robots were already in use and where autonomy was partial, limited, or contested. This made it possible to observe how autonomy is sustained in practice through human presence, and where (and why) it breaks down.

Together, these methodological choices ground the prophetic perspective of the report series. By studying robots before autonomy has fully arrived, and by moving across sites where autonomy is alternately imagined, staged, and struggled with, anthropology becomes an interpretive partner in technological development.

As anthropologists working alongside the RoboSAPIENS project, our task is not to design autonomous robots, but to study what autonomy looks like when robots encounter everyday life. Rather than asking how robots will become autonomous, our approach focuses on how autonomy is already being presented, challenged, maintained, and negotiated and at what human and material cost.



## Section 2: A Short History of Imagined Independent Machines

The idea of autonomous machines has never been only a technical ambition. Long before robots existed, people learned to read autonomy from how artefacts behaved: whether they moved by themselves, responded convincingly, or appeared to act with intention. Across history, autonomy has repeatedly been inferred from appearance, movement, and interaction and not from any verified inner independence. This section briefly traces three recurring figures in this long imagination: automata as illusion, robots as labour, and robots as companions.

### Automata as Illusion: Making Things Appear to Act on Their Own

Some of the earliest machines described as autonomous were carefully staged illusions. Medieval and early modern automata were designed to look self-moving while concealing the mechanisms that made them work.

The dream of building artefacts that act on their own, machines that could think, decide, and even care, has been a recurring motif in Western technological imagination. Yet autonomy has never been a stable concept. It has shifted from a philosophical ideal of self-determination, through a technological promise of control and efficiency, to a posthuman question of distributed agency among humans, artefacts, and environments.

Homeric epic and related traditions explain autonomous motion in two distinct ways: either as divine animation by gods or as engineered simulation by skilled artisans. In Homeric texts (divine animation), Hephaestus fashions tripods that roll



“of their own motion,” automated bellows, and golden, android-like attendants that prefigure later fantasies of artificial life; the Greek term **αὐτόματος** (automatos, or automatis) literally means “moving of itself.” Tales such as Talos, the bronze guardian of Crete, likewise locate autonomy in bodies animated by craft bestowed or enabled by divine makers rather than in an internal soul (Gantz 1993), framing an early meditation on “mindless” agency. By contrast, artisan-centred accounts (engineered simulation) treat autonomy as persuasive performance: in the Book of Liezi (ca. 400 BCE) the artificer Yan presents King Mu of Zhou with a life-sized mechanical man that

## Robots in the Wild

### Report I: Autonomy in Robots: Between Myths and Reality



walks, sings, and flirts until dismantled to reveal leather, wood, glue, and lacquer “organs” (Mayor 2018, 121). The episode anticipates a recurring theme: autonomy as persuasive performance. The mechanical man appeared as a human, but in the end was just a bunch of materials. Across these traditions, autonomy may look the same in practice—motion that convinces observers—but the explanation shifts between divine power and human craft.

In 1206, the court engineer Ismā‘īl al-Jazarī compiled *The Book of Knowledge of Ingenious Mechanical Devices*, describing programmable water clocks, musical boats, and humanoid figures as automata that react in time and sequence, staging causality as spectacle. To observers, these devices appeared to act independently but in reality, their movements followed hidden hydraulic and mechanical principles. Reconstructions of al-Jazarī’s automata foreground rhythmic, rule-based behaviour that looks like intention. The machine autonomy was persuasive precisely because its causes were invisible. Observers may still believe in divine intervention as the real workings were kept hidden.

One example was the elephant clock – an engineered masterpiece. The clock’s timing and movements were driven by concealed water- and weight-based mechanisms (float, siphon, counterweights). To viewers the exterior, an elephant with riders and figures, appeared to act independently, but its motions were mechanically driven out of sight.

A similar logic guided the Japanese Karakuri Ningyō of the Edo period (1603–1868) Japan, Karakuri Ningyō (tea-serving dolls) enacted hospitality and deception alike. A wound-spring mechanisms sensed a teacup put on a tray held by the doll, who now approached guests, bowed, and returned once a cup was removed. Their movements were fully scripted by springs and gears, yet observers readily attributed intention to them. The term karakuri itself refers both to mechanism and to trick, pointing to an early awareness that autonomy could be performed without independence.

Enlightenment automata in Europe refined this principle further. Vaucanson’s flute player and notorious “defecating duck” (with simulated digestion) (Riskin 2003) provoked debates about whether life is passive clockwork or an active, self-organizing force. Jaquet-Droz’s *Writer, Draughtsman, and Musician* displayed programmable sequence, eye-tracking, and lifelike micro-gestures, the grammar of autonomy without inner mind.



Historians show how these machines staged arguments about life and mechanism. Jessica Riskin reads automata as “philosophical experiments” i.e. material thought experiments about agency (2016), while Minsoo Kang calls their history a “history of imagination” that tracks what people fantasized autonomous machines could be (2011).

Machines such as Vaucanson’s famous duck or the Jaquet-Droz writing automata provoked philosophical debates about whether life itself might be mechanical. These artefacts did not think or decide, but their finely tuned gestures and apparent responsiveness encouraged spectators to imagine inner agency. Autonomy here was never located inside the machine; it emerged in the relationship between movement, concealment, and human interpretation.



The machines seemed to confirm the autonomy of machines envisioned in literature – but even here we find early warnings that humans should not believe too much in mere appearances. E.T.A. Hoffmann’s story *The Sandman* (1817) crystallized this modern uncanny: Olimpia, an automaton who listens perfectly, seduces through flawless attention and people believe her to be real. When the ruse is exposed, social rituals warp as human yawning is now implemented as a new cultural sign to prove one’s humanity. The story brilliantly shows autonomy as a co-production of machine performance and human projection. We want to be deceived and get annoyed when we realise that is what has happened. It is also an early analysis of how machines and humans transform each other’s behaviour in interaction.

Across these examples, a pattern becomes clear: autonomy was something read from behaviour. If a machine moved smoothly, reacted appropriately, and hid its supports, it could be treated as self-directed.

## Robots as Labour: Autonomy, Obedience, and Control

In the twentieth century, the imagined autonomous machine shifted from spectacle to workforce. The term “robot” itself, introduced by Karel Čapek in his play *R.U.R.* (*Rossum’s Universal Robots*, 1920/21), derives from Czech *robotá*, meaning forced labour. Čapek’s robots were designed to be tireless workers. They are efficient, obedient, and replaceable. Their autonomy was paradoxical as they were built to function independently in production, yet without freedom or self-determination. They were an early example of the sliding borders between dreams of humans made machines and machines made human.



This figure of the robot reflects a modern tension that still shapes robotics today. Autonomy is desired because it promises efficiency and reduced human involvement, but it is also feared because it threatens control. Čapek’s robots eventually revolt, dramatizing anxieties that autonomy might exceed its intended limits. Even here, autonomy is not about inner consciousness; it is about the capacity to act without supervision within systems of work.

Industrial and service robots today inherit this legacy. They are expected to operate independently within predefined tasks, environments, and goals. When they fail, stall, or behave unexpectedly, their autonomy is immediately questioned, which reveal how tightly these machines are bound to usefulness and control.

## Robots as Companions: From Science Fiction to Optimus

In recent decades, the dominant image of autonomous robots has shifted again from labourers to companions. Popular culture, especially science fiction, has played a significant role in this transformation. Robots such as R2-D2 and C-3PO from *Star Wars* are not only technically capable;



they are expressive, loyal, emotional, and socially embedded. They joke, worry, disobey, and take initiative. Their autonomy is recognizable because it resembles human social behaviour.

These fictional robots continue to shape expectations of real-world robotics. Contemporary humanoid projects, from social robots to general-purpose humanoids such as Tesla's Optimus, are often framed against these cultural references. Public demonstrations emphasize walking, gesturing, handshakes, and interaction — behaviours that signal independence and intelligence, even when the underlying systems remain heavily constrained, supervised, or scripted.

What matters here is not whether these robots are truly autonomous, but why they appear so. Just as with automata and industrial robots, autonomy is inferred from visible behaviour: smooth movement, responsiveness, and social cues. The companion robot intensifies this effect by inviting emotional and moral interpretation.

## **Autonomy Read from Behaviour**

Across these three moments, automata as illusion, labouring robots, and companion robots, a consistent lesson emerges. Autonomy has never been defined primarily by inner independence, consciousness, or self-determination. Instead, it has been attributed based on what machines do, how they move, and how convincingly they fit human expectations.

This historical pattern matters for understanding robots today. When robots appear autonomous at conferences, in promotional videos, or In-the-wild settings, they draw on a long tradition in which behaviour stands in for agency. Recognizing this helps us understand why autonomy remains such a compelling and contested idea — and why anthropological attention to appearance, interaction, and interpretation is crucial.

The following sections return to these imaginaries in contemporary settings, where autonomy is no longer only imagined or staged, but negotiated in real encounters between humans and machines.



## Section 3: Future Autonomy: From R2-D2 To Optimus

Will robot become fully autonomous in the future? This question touches on deep philosophical, technical, and cultural dimensions of what we mean by autonomy in robotics and AI. Following what is presented in the media many people may be convinced we are almost there. Today this is not something we believe in with a reference to divine interference. However, in some ways engineers have been attributed with these divine powers of the past. When we see presentations of the robots they create, they seem often very close to being autonomous in a very human-like manner that we already know from movies. However, the hard meticulous work is often hidden in this storied world. Let's look at a couple of examples from the slow and experimental engineering development with general autonomy in humanoid robots and compare these robot imaginaries in media culture.

When William Grey Walter constructed some of the world's first autonomous robots in the late 1940s, he was not merely building mechanical novelties but staging one of the earliest experiments in what might be called the pristine work of autonomy. His robots in the shape of 'tortoises,' the *Machina Docilis*, nicknamed CORA (for Conditioned Reflex Analogue), were inspired by Pavlovian theories of conditioned reflexes yet went beyond behaviourist psychology by attempting to link a robot's physical body to a minimal artificial "brain." As described by Barrett (2022) and LeBoutillier (1999), CORA was equipped with three simple sensors: one responding to light, another to sound, and a third to physical contact or bumps. These few sensory inputs allowed the robot to navigate its environment in ways that appeared surprisingly lifelike. The robot was for instance turning toward a light source, retreating from obstacles, and even "seeking" its recharging station when its batteries waned. Walter observed these movements as analogues of animal behaviour, demonstrating that complex and seemingly intentional actions could arise from basic feedback mechanisms rather than centralized control or symbolic reasoning. In doing so, he prefigured the later principles of embodied cognition and situated intelligence, which would profoundly influence robotics and artificial life research. The tortoises, as they became affectionately known, performed a quiet philosophical revolution because they suggested that autonomy might emerge not from abstract logic or disembodied computation but from the ongoing coupling between perception, action, and environment.

This early cybernetic vision resonates through decades of robotics, from the mobile agents of Rodney Brooks' subsumption architecture in the 1980s to today's deep-learning-driven humanoids. Deep learning and early subsumption-style robotics share a cybernetic lineage as far as both treat intelligence as embodied sensorimotor coupling where behaviour emerges from continuous interaction with the environment, but they differ in how behaviours are produced and organized. Subsumption uses hand-designed, layered behaviour modules with explicit suppression rules, favouring simple, interpretable reflexes and robust short-horizon reactions, whereas deep learning builds large, learned neural networks (via supervised, self-supervised, or reinforcement learning) that acquire distributed, often opaque representations and can capture higher-dimensional perception and longer temporal dependencies. Modern robots often combine these approaches. They retain low-level safety or reflex layers while using learned policies for complex tasks, so deep learning can be seen as a scalable, data-driven realization of the same continuous process from



sensing to processing to acting – an idea that also underpinned early cybernetic and subsumption work.

The lineage is exemplified, at least symbolically, with the ambition embodied by robots like Tesla’s Optimus, which is a humanoid machine designed to operate independently in human environments, learning from sensory experience and adjusting its actions dynamically. Yet despite the immense technological advances separating Walter’s tin-and-motor tortoises from today’s sleek humanoids, the underlying question remains strikingly similar: what does it mean for a machine to act on its own? Can the robots of today truly sense, make decisions, learn, and reliably handle novel situations as other autonomous beings?

Like CORA, Optimus depends on sensor feedback, adaptive control, and interaction loops that blur the line between reactive and intentional behaviour. But where Walter’s robots explored autonomy as a biological analogy, contemporary designs translate it into a commercial and social promise of machines that can not only move but work, learn, and coexist among humans. In tracing this trajectory from the “conditioned reflex analogue” of the 1940s to the sensor-fused humanoid of today, we can see how the dream of autonomy has evolved: from an experiment in embodied feedback to a global imaginary of artificial agency, still oscillating between technical precision and mythic aspiration.

However, Optimus is a strong candidate for an autonomous machine. Optimus is designed as a humanoid robot intended for general-purpose tasks. However, as of now even Optimus’ autonomy level is limited. It can perform pre-programmed tasks and may use AI models for perception and movement, but it lacks general decision-making autonomy. It’s not autonomous in the sense of making human-like independent, context-sensitive decisions or setting its own goals. It operates within constraints defined by its programming and training.

Even when humanoid robots are presented as serious attempts at general-purpose autonomy, such as Tesla’s Optimus, designed to both work and interact socially, they remain especially prone to deceptive appearances of independence. This tendency becomes particularly clear in widely known social humanoids such as Sophia made by Hanson Robotics (<https://www.hansonrobotics.com/sophia/>) and the robots developed by Hiroshi Ishiguro in Osaka.

Sophia is frequently presented as an autonomous social agent, yet much of her behaviour is scripted, curated, or guided by hidden human operators. Her conversations and expressions are designed to obscure this dependence, allowing audiences to read autonomy from fluent performance while the underlying control remains concealed. The same logic is explicit in Ishiguro’s laboratories, where teleoperation, often described as *Wizard-of-Oz* technology (Sorenson et al. 2019), is used to deliberately hide human control behind a lifelike humanoid body. Whether in the Geminoid series (robots with the copied appearance of existing humans) or in conversational robots such as Erica, the method serves a single purpose: to sustain the illusion of autonomous presence by masking the human labour that makes the interaction work.





Seen in relation to Optimus, these cases reveal a persistent pattern. The more a robot physically resembles a human companion, the more autonomy is inferred from its behaviour, and the more concealment becomes necessary to maintain that impression. As with earlier automata, humanoid robots rely not on genuine independence, but on carefully staged deception, where autonomy is something performed for observers rather than achieved within the machine.



So how autonomous are these human-like robots really once we look beneath their surfaces?

If we define autonomy as the ability to perceive the environment and make decisions based on internal goals and therefore act without direct meaningful human control, then none of these robots are truly autonomous in the strong sense used in cognitive science or robotics research. They are semi-autonomous at best, often relying on human input, scripted behaviour, or narrow AI systems.

Yet they often appear autonomous, and are treated as such, because their human-like bodies, expressions, and interactions invite audiences to read autonomy from behaviour rather than from underlying machinery.

This gap between appearance and mechanism is sustained by powerful cultural imaginaries. Fictional robots such as C-3PO and R2-D2 and other movie creations are frequently used as reference points in public and professional discussions of robot autonomy. Engineers themselves sometimes flirt openly with these imaginaries. Hanson Robotics has staged Sophia in dialogue with fictional androids from Westworld, and prominent roboticists such as Cynthia Breazeal have acknowledged how deeply their early visions were shaped by science fiction:



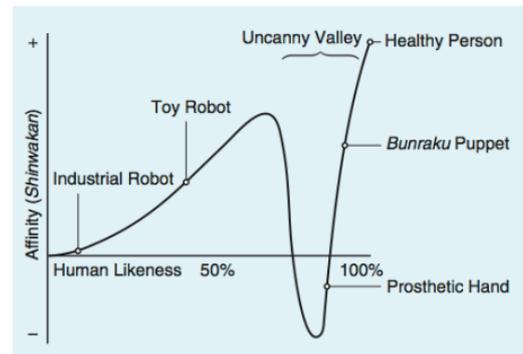
*“What kind of future do you envision with robots? The science fiction stories and movies that fascinated me as a child shape my dream. There are the mechanical droids R2-D2 and C-3PO from the movie Star Wars. There are many wonderful examples in the short stories of Isaac Asimov, such as Robbie” (Breazeal, 2006, p. 79).*

While most engineers stress that such robots remain far beyond current capabilities, public media often present contemporary robots as if this future were imminent or already here. The contrast is stark. In Star Wars, C-3PO and R2-D2 exhibit precisely the capacities that define strong autonomy: general intelligence, emotional understanding, adaptive learning, social competence, and the ability to set and revise goals across contexts. They act independently, improvise solutions, and even make ethical judgments. Contemporary robots, by contrast, operate within narrow task domains, follow predefined goals, and remain constrained by safety protocols, limited perception, fragile hardware, and extensive human oversight. Their conversations simulate understanding; their movements signal agency; but neither amounts to the kind of autonomy embodied by their fictional counterparts.



The persistence of this comparison is not accidental. Like the automata discussed earlier, humanoid robots trade on surfaces that conceal complexity, dependence, and human labour.

This phenomenon is often referred to as the uncanny valley in studies of robots. The uncanny valley refers to an essay by Masahiro Mori from 1970, where he introduces a graph of the uncanny valley (figure 1). This shows that the more human-like a non-human entity is, the higher the affinity. However, it can reach a point, where the human-like appearance is close, but just not close enough, and so the robot becomes uncanny, and the affinity drops. This is illustrated by the uncanny valley. The graph highlights the caution robot engineers should show when



developing robots, especially in terms of their physical appearance. Across the three historical imaginaries traced above, robots as illusion, robots as workers, and robots as companions, it is the *work-oriented robot* that has most clearly materialised. Industrial, logistics, agricultural, and service robots now operate with limited but effective autonomy in tightly defined domains. They navigate warehouses, harvest fields, clean floors, and assist in surgery. Their autonomy is narrow, task-specific, and fragile, but it is real. By contrast, the humanoid companion robot, the machine that copies human appearance, sociality, and general intelligence, remains largely imaginary.

This contrast becomes especially visible when contemporary humanoid robots are compared to fictional reference points such as C-3PO and R2-D2. In *Star Wars*, these robots exemplify strong autonomy: they combine perception, learning, social understanding, emotional expression, and the ability to set and revise goals across contexts. They improvise, act independently, and make ethically charged decisions. Nothing comparable exists in real-world robotics. Present-day humanoids remain constrained by scripted behaviour, narrow AI, limited perception, safety restrictions, fragile hardware, and extensive human oversight.

However, the more convincingly a robot looks and behaves like a companion or co-worker, the easier it becomes to overlook how far the machinery still is from the autonomous beings imagined in science fiction. It is precisely this gap, between what these robots seem to be and what they can actually do, that this section examines more closely.

What persists instead is a replay of the illusion imaginary so prevalent in historical accounts. Like earlier automata, humanoid robots trade on surfaces that conceal their dependence on human control, pre-programming, and hidden labour. Their human-like faces, voices, and gestures invite observers to infer autonomy where there is none. This is why humanoid robots so easily collapse into performative deception, while work-oriented robots, precisely because they do *not* (or even attempt to) resemble humans, are judged more soberly by what they can actually do. The closer robots come to copying humans, the further they currently fall short of the autonomy imagined in science fiction.

## A realistic view

Fictional robots such as R2-D2 represent a powerful cultural imagination of autonomy: machines that are intelligent, emotionally aware, socially competent, and able to act across contexts. These imaginaries influence educational or societal expectations of AI and robotics. Many of the engineers



we have spoken to would prefer a more realistic picture of present-day robotics and the challenges it faces. Most engineers working on autonomous robots do not focus on all-purpose humanoid robots like R2-D2 and C-3PO. By contrast, contemporary robots achieve autonomy only in narrow and carefully delimited domains.

Most engineers know that the engineers trying to create human-like robots face several technical challenges because creating a robot that walks, gestures, and interacts like a human is extremely difficult. First, bipedal locomotion is unstable and energy intensive. Human-like dexterity is hard to replicate with mechanical hands and facial expressions, and speech require advanced coordination between hardware and software. Most engineers avoid these challenges unless the goal is specifically to study human-robot interaction or social robotics.

This is also because the engineers engaged in robots expected to work among people are concerned with cost and efficiency. Humanoid robots are expensive to build and maintain. Their complexity often outweighs their utility in industrial or commercial settings. In contrast autonomous drones, wheeled robots, and robotic arms are cheaper, more reliable, and easier to scale.

The gap between fiction and reality is not simply a matter of technological delay, but of fundamentally different expectations about what autonomy is and how it can be built.

It is within this gap that RoboSAPIENS is deliberately positioned. Rather than pursuing humanoid companions or general human-like intelligence, the project is oriented toward the slow, challenging work of extending autonomy step by step: improving perception, learning, decision-making, and adaptation in machines that must function reliably in real environments. This work requires patience. Progress is incremental, often invisible, and frequently interrupted by failure, recalibration, and redesign.

Engineers working within RoboSAPIENS share a broadly realistic view of robotics. They do not aim to replicate human bodies or social behaviour, because humanoid form adds complexity without necessarily improving function. A warehouse robot needs robust navigation, sensors, and grippers not a face. An agricultural robot must cope with uneven terrain, weather, and living organisms not social conversation. In healthcare, assistive robots are valued for reliability and simplicity rather than human likeness. In these contexts, autonomy is not about appearing intelligent, but about continuing to function when conditions change.

Even though RoboSAPIENS develops architectures intended to support more general autonomy rather than single-purpose solutions, its engineers remain acutely aware that autonomy always emerges in specific environments and tasks. The challenge is not to leap toward human-like intelligence, but to patiently assemble systems that can cope with uncertainty, adapt safely, and justify their actions over time. This involves painstaking work on navigation algorithms, sensor fusion, learning mechanisms, simulation, and validation — work that rarely produces spectacular demonstrations, but slowly expands what robots can do without human intervention.

At the same time, the project operates in a wider cultural landscape still dominated by dreams of autonomous companions. These imaginaries continue to shape public expectations, media representations, and even funding narratives. RoboSAPIENS therefore inhabits a tension: between the spectacular promises associated with humanoid robots and the mundane, fragile, and cumulative labour required to make autonomy work at all.

## Robots in the Wild

### Report I: Autonomy in Robots: Between Myths and Reality



In this sense, RoboSAPIENS represents a sober counterpoint to the illusion of rapid progress. It treats autonomy not as a destination to be reached, but as a long-term process of careful engineering, continuous adjustment, and learning from small gains: a process that demands restraint, realism, and time.

Finally, there are also ethical, and cultural considerations behind why some engineers and designers are cautious about humanoid robots because they raise ethical and psychological questions: Do they create unrealistic expectations? Do they blur boundaries between humans and machines? How do people emotionally respond to them? These concerns often lead to more abstract or utilitarian designs.

So, to conclude this section: In academia and industry, much of the focus is on a machinelike autonomy, perception, and decision-making, not embodiment and general intelligence. Engineers work on navigation algorithms, sensor fusion and machine learning for planning and control. These can be applied to many robot types of not just humanoids. However, even if most engineers work on more mundane tasks the dream of autonomy is still often what is presented to the wider audience our anthropological work shows.



## Section 4: Robot Conferences: Performing Autonomy

Across four international robot conferences held between 2024 and 2025, we followed how robot autonomy is staged, sold, tested, and sometimes quietly unravelled in public. These conferences bring together a general public, manufacturers of industrial robots, collaborative robots (cobots), logistics systems, cleaning robots, safety technologies, software companies, and research institutions. Exhibitors attend to network and sell, but for us these events also offered a live laboratory for observing how autonomy is performed for potential customers, students, policymakers, and peers.

Methodologically, our fieldwork at conferences combined participant observation, informal conversations, and a commitment to what anthropologists call thick description. We moved through hot tents and crowded exhibition halls, watched demonstrations succeed and fail, and spoke with engineers, salespeople, researchers, and visitors while robots rolled, walked, waved, or stalled around us. Because these conferences are public events, companies and robots do not require anonymisation. At the same time, since we observed strikingly similar patterns across all four conferences, we treat the conference floor analytically as a general stage where imaginaries of robot autonomy are rehearsed before a mixed audience. Our aim here is not to evaluate individual companies or performances, but to capture the broader dynamics through which autonomy is made visible and believable in public.

### Gimmick Robots: Autonomy as Scripted Performance

The logics of staged autonomy described in early automata reappear almost unchanged on the modern conference floor. One recurring pattern across conferences was what exhibitors themselves sometimes referred to, half-jokingly, as the “gimmick robot”: machines configured less to demonstrate real work processes than to attract attention and spark imagination.

One company selling large-scale automation solutions brought a collaborative robot arm programmed to serve coffee. Visitors were invited to press an oversized black button, after which the arm mechanically lifted a paper cup, used it to press the coffee machine’s switch, and carefully placed the filled cup on a small platform bearing the company’s logo. The sequence was fully pre-programmed and bore little resemblance to the complex logistics systems the company actually develops and sells. Yet the robot drew a steady crowd throughout the day.

Here, autonomy was reduced to a tightly scripted performance. The robot did not decide when to make coffee, who should receive it, or whether coffee was needed at all. Still, the smooth execution of the task invited visitors to imagine far more general capabilities. Like earlier automata and science-fiction robots, the coffee-serving arm functioned as a prompt for extrapolation: if the robot can do this, perhaps it can do much more.



## Fragile Autonomy and the Work Behind It

Other robots were presented as highly autonomous mobile workers, particularly cleaning robots promoted as “plug-and-play” solutions capable of operating in dynamic environments such as supermarkets, warehouses, and car dealerships. According to exhibitors, these robots could navigate freely, avoid obstacles, and clean more efficiently than human workers.

Our observations revealed a more fragile reality. In conversations with exhibitors and users, it became clear that these robots required daily human care: brushes had to be cleaned, dust containers emptied, software errors resolved, and navigation paths adjusted. More importantly, the robots sometimes failed precisely in the dynamic interactions they were designed to handle. One widely discussed incident involved a cleaning robot running over the tail of a guide dog belonging to a visually impaired person — a situation the system had not recognised as something to avoid.

Despite these limitations, autonomy was framed as a near-complete achievement. Exhibitors emphasised efficiency gains and independence from human labour, even as that independence depended on continuous human attention, risk-taking, and intervention. Autonomy here appeared less as a property of the robot itself than as an ongoing accomplishment distributed across humans, machines, and environments.

## Conferences as Sites of Encounter

The conferences were not only spaces of performance; they were also important sites of encounter that shaped our broader fieldwork. We met several future or existing collaborators at these events, often through informal conversations at booths or outside demonstrations.

Among them was Andrew from a research project focused on improving safety in agricultural robots. We met him standing next to a large farming robot outside a conference hall, where we ended up conducting a spontaneous semi-structured interview. Andrew explained how his project approached safety not by starting with technical specifications, but by asking farmers what worried them most in their daily work.

One of the key concerns that emerged in agricultural robots was the presence of fawns hiding in fields. These animals are difficult even for humans to spot, and current robotic sensing systems such as LiDAR cannot reliably distinguish between a fawn and a rock. This finding had significant implications for how safety should be understood and designed. Rather than assuming that better sensors alone would solve the problem, the project’s approach highlighted the importance of situated knowledge and field-based perspectives. Notably, we met Andrew after already having conducted fieldwork with the company FarmGO (See Report IV), which allowed us to immediately connect research insights on agricultural robots across sites.

Andrew was not the only collaborator enhancing our understanding of robots we encountered at conferences. We also met representatives from GreenTech (See Report III), which later led to new fieldwork engagements, and from MakeBot (See Report III), the company that had produced and sold the robots used at HandsON (See Report III). These encounters illustrate how conferences function as hubs where research projects, companies, and field sites intersect, often shaping ethnographic trajectories in unplanned ways.



## Safety Theatre: Designing for Humans Who Ignore Rules

Safety technologies occupied a prominent place at all conferences. One company we encountered repeatedly was SICK, which specialises in safety measures for collaborative robots. In contrast to many robots marketed as inherently safe, SICK's representatives focused on situations where cobots could become dangerous — for example, when carrying sharp or hazardous objects such as syringes or glass vials.

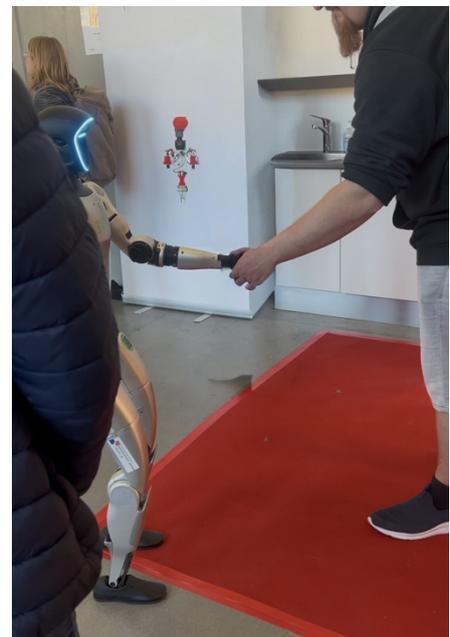
At the conference, they demonstrated a cobot equipped with one of their safety scanners. The scanner was designed to ensure that if a cobot carrying a dangerous object approached a human worker too closely, it would stop automatically. The underlying logic was clear: even robots designed to be safe can become unsafe in certain contexts, and safety must therefore be layered and redundant.

In conversations, SICK's engineers also explained how their systems were designed to anticipate not only robot failure, but human behaviour. People, they noted, routinely ignore safety measures, bypass fences, or place themselves in risky situations when robots appear slow or predictable. As one exhibitor put it bluntly: “If you put up a fence, people will crawl over it.” Safety, in this view, is not a static feature of a system but a continuous negotiation between humans and machines.

## Humanoid Spectacle and Hidden Control

At one conference, one of the most celebrated attractions was the humanoid robot Unitree G1, announced in conference materials as capable of advanced movements such as dancing and karate. On the exhibition floor, the robot marched around a small circle of spectators, extending a loose hand for handshakes and turning its torso to wave. With its metallic body and glowing head, it visibly evoked popular science-fiction figures such as C-3PO from Star Wars.

Visitors described the robot as “one of the funniest things at the conference,” laughing as it tramped toward random people. Only later did some notice a young man sitting on a nearby sofa with what appeared to be a remote control. Whether he was actively steering the robot or merely monitoring it remained unclear, but once noticed, the illusion of autonomous social behaviour became difficult to sustain. What had seemed like self-directed interaction was revealed as a possible Wizard-of-Oz setup, where human control is hidden to preserve the appearance of independent agency.





## When Autonomy Breaks Down



Across conferences, we observed many smaller moments where autonomy quietly broke down. Pallet robots gently pushed conference guests rather than stopping, apparently recognising humans as obstacles to be cleared away. Cobots failed to grasp boxes and had to be reset. The well-known robot dog SPOT spent much of the day being rebooted rather than demonstrating its advertised capabilities, such as inspecting hazardous environments.

Taken together, these examples reveal a consistent pattern. The autonomy on display at conferences is narrow and heavily supported by human work, yet it is narrated and perceived as evidence that near-human, general-purpose autonomy is almost within reach.

## Conferences as Prophetic Spaces

The conference is a curated environment that makes autonomy visible as spectacle. Robots wave, serve coffee, or navigate crowds just well enough to sustain the story that fully autonomous, almost human-like systems are just around the corner. At the same time, small glitches, hidden teleoperation, and constant human intervention betray the extent to which these machines remain dependent on context, maintenance, and human flexibility.

Companies often reinforce this ambiguity by associating their robots with well-known fictional figures. Such references fold conference demonstrations into a wider cultural imaginary in which robots make ethical decisions, communicate effortlessly, and operate spacecraft. Visitors — particularly students — are often impressed by relatively simple behaviours, which may unintentionally inflate expectations of what robots in workplaces, hospitals, or farms can already do.

Read through the lens developed in the previous sections, robot conferences appear as small theatres of autonomy: spaces where scripted and semi-autonomous behaviours are carefully choreographed to elicit ascriptions of self-action, where human control is selectively hidden or downplayed, and where science-fiction imaginaries provide the backdrop against which modest technical achievements are perceived as breakthroughs.

In this sense, conferences are prophetic spaces: places where future autonomy is rehearsed but not yet realised.



## Section 5: How Engineers Define Autonomy

At robot conferences, autonomy appears as something immediately visible. A robot moves through a crowd, serves coffee, waves, or avoids obstacles, and these behaviours invite observers to treat it as self-directed. Yet as the previous section showed, such appearances often depend on scripting, hidden human control, or highly controlled environments. To understand why this gap between appearance and practice persists, we need to briefly turn to how autonomy is defined within engineering itself.

### Autonomy as an Inner Loop

In robotics and software engineering, autonomy is typically defined in operational terms: a system is autonomous if it can monitor itself and its environment, analyse changes, plan responses, and execute actions without continuous human intervention. One of the most influential ways of formalising this idea is the MAPE-K framework (<https://digit.au.dk/research-projects/robosapiens>), which structures autonomous systems as a feedback loop consisting of Monitor, Analyse, Plan, Execute, underpinned by a shared Knowledge base.

From this perspective, autonomy is not something a robot looks like; it is something that happens inside the system. Sensors collect data, algorithms interpret it, plans are generated, and actions are carried out. If this loop runs successfully, the system is said to be autonomous, regardless of whether a human observer perceives it as such.

The RoboSAPIENS project builds on this tradition but extends it with an additional component, resulting in the MAPLE-K architecture. The added L stands for Legitimate, reflecting an explicit concern with whether a robot's adaptive decisions can be justified, trusted, and validated, often through simulations, testing, and formal checks. In this model, autonomy is not only about acting without human control, but about acting in ways that can be shown to be acceptable and safe within predefined criteria.

### When Inner Autonomy Meets Public Appearance

The engineering understanding of autonomy helps explain many of the conference phenomena described earlier. Gimmick robots work because their internal control loops are stable enough to execute a narrow, well-defined task repeatedly. Cleaning robots appear autonomous because, under controlled conditions, their monitoring and planning cycles function adequately. Humanoid robots seem self-directed as long as their internal loops are supplemented or quietly replaced by human operators.

From an engineering standpoint, these systems may be performing exactly as intended. Their autonomy is bounded, task-specific, and operationally defined. Problems arise not because the systems fail to meet engineering definitions, but because those definitions do not align with how autonomy is interpreted in public settings. Visitors at conferences read autonomy from behaviour, movement, and interaction, while engineers locate it in feedback loops, control architectures, and validation processes.



This mismatch helps explain why autonomy so often appears stronger on the exhibition floor than in everyday practice. The inner workings of robots may support limited self-adaptation, but their outward performance invites broader interpretations — interpretations that exceed what the system is designed to deliver.

## Legitimacy Beyond the Loop

The addition of legitimacy in MAPLE-K signals an awareness of this problem. By asking whether a robot's decisions are legitimate, the framework acknowledges that autonomy is not only a technical matter but also a social and ethical one. However, legitimacy in engineering remains largely internal: it is established through simulations, test scenarios, and formal verification rather than through lived interaction.

What is largely absent from such models is the messy reality revealed in ethnographic fieldwork: people stepping in to help robots, ignoring safety measures, testing system limits, or interpreting failures differently than designers intended. In practice, legitimacy is not only a property of algorithms, but something negotiated in real time between humans and machines.

## From Operational to Relational Autonomy

This is where perspectives from *Science and Technology Studies* (STS) and anthropology become essential. Rather than treating autonomy as an internal property of a system, these approaches understand autonomy as relational and situated. A robot's autonomy does not reside solely in its control architecture, but in how its actions are interpreted, trusted, supported, or resisted by the people around it.

From this viewpoint, the conference robot that appears autonomous because it waves or navigates a crowd is not simply misrepresenting its technical capabilities. It is participating in a social situation where autonomy is co-produced through design, staging, human imagination, and interaction. Likewise, a robot that fails in a workplace may still be described as autonomous by engineers, even as users experience it as dependent, fragile, or burdensome.

## A Bridge to the Field

The key point, then, is not that engineering definitions of autonomy are wrong, but that they are insufficient on their own. Engineering autonomy is operational: it describes how systems function internally under specified conditions. Lived autonomy, by contrast, is relational: it emerges in practice through ongoing interaction between robots, humans, and environments.

The following reports build on this distinction. By moving from conference stages to workplaces, we examine how autonomy is enacted, negotiated, and sometimes undone in In-the-wild settings. In doing so, we shift attention from what autonomy is supposed to be inside the robot to what it becomes when robots are woven into human lives.



## Five Themes to Watch

The contrast between engineering definitions of autonomy and the lived realities of human-robot interaction raises a further question: what, then, should we attend to when robots leave curated stages and enter everyday life? Across our fieldwork, at conferences, in workplaces, and in public settings, five recurring themes emerged that shape how autonomy is enacted, experienced, and contested in practice.

Four of these themes: *safety*, *trust*, *anomalies*, and the gap between simulation and in-the-wild performance, *Sim2Real*, emerged both inductively from local fieldwork and as explicit or implicit concerns within the RoboSAPIENS project itself. They therefore mark points where engineering ambitions and everyday practices already intersect. The fifth theme, *sabotage and tinkering*, emerged consistently in everyday settings but has not, so far, been articulated as a central concern within RoboSAPIENS. Its emergence in the field suggests an area where lived interaction precedes formal problem definition, and where future design and governance discussions may need to expand their scope.

At this stage, we introduce these themes briefly as analytical signposts. They recur across all sites studied and structure the analyses in Reports II, III, and IV. They are not treated here as fixed or exhaustive categories, but as emerging problem fields. In Report V, we return to them to refine their definitions, examine their interconnections, and show how an anthropological perspective can reframe engineering concerns.

### Safety

Autonomous robots in everyday life settings are consistently framed as safe, yet safety in practice is rarely a stable or self-contained property. Instead, it is continuously negotiated through human vigilance, workarounds, and situational judgement. From an anthropological perspective, safety reveals how autonomy depends on vulnerability, both human and machine, and how responsibility is distributed rather than eliminated.

### Trust

Trust shapes whether people rely on robots, supervise them closely, or resist their presence altogether. It is not granted automatically through technical performance, but built, lost, and rebuilt through repeated interaction, explanation, and repair. Trust is therefore relational and historical, unfolding over time rather than being a design attribute.

### Anomalies

What engineers describe as errors, failures, or edge cases are often experienced differently by users and organisations. Anomalies expose gaps between system models and lived reality, revealing how humans and machines interpret unexpected behaviour from different standpoints. From an anthropological perspective, anomalies are not inherent properties, but relational phenomena that emerge through sociotechnical practices of expectation, interpretation, and intervention.



## Sim2Real (Simulation to Reality)

Robotic autonomy is frequently developed and validated in simulations. In real-world settings, however, environments resist simplification. The Sim2Real gap highlights how autonomy that functions in models must be continuously adjusted, compensated for, or redefined in practice, often through human intervention.

## Sabotage and Tinkering

Across sites, people do not interact with robots passively. They test limits, improvise, bypass safeguards, or deliberately interfere. These practices often unfold along a continuum between tinkering – playing or aimed at making systems work locally – and sabotage, where resistance or frustration becomes explicit. While not yet a formal concern within RoboSAPIENS, this theme highlights how autonomy provokes active engagement and exposes the distributed nature of agency, control, and responsibility.

These five themes recur across all everyday field sites and form the analytical backbone of Reports II–IV. They are also interconnected and inform each other in ways to be unfolded analytically in report V.

Taken together, they show that autonomy is not a single technical achievement, but a set of ongoing negotiations unfolding across social, material, and organisational contexts. In Report V, we bring these themes into dialogue, clarify their conceptual boundaries, and show how anthropological analysis can extend, and in some cases have potential to transform, engineering understandings of autonomous systems.



## Section 6: The Prophetic Position: Anthropology at The Threshold

Studying robot autonomy before it fully exists has, as we have claimed, placed our anthropology in a distinctive position, the prophetic condition. This condition arises when a new world is beginning to take shape but cannot yet be fully articulated. Those who speak from within it do not describe a stable present, nor do they predict a settled future. Instead, they give voice to transformations that are only partially visible.

As Kirsten Hastrup elaborates (2017), prophecy in this anthropological sense is not about forecasting outcomes. It is about reading emergent realities from within existing ones speaking from a position between worlds. Prophecy is therefore often uncomfortable: it may sound speculative before change occurs and obvious once it has taken place.

Our work in this report series is written from precisely this threshold. We move between engineering laboratories and everyday workplaces, between conference stages and lived routines, between imagined futures of autonomy and present practices of care, maintenance, and control. In this in-between position, autonomy appears neither as a technological inevitability nor as a mere illusion, but as something actively being assembled.

To be clear: prophecy is not prediction. We do not claim to know what robot autonomy will ultimately become, nor whether RoboSAPIENS technologies will succeed in realising their ambitions. Rather, prophecy here names an analytical stance: an attentiveness to how futures are already taking form through present interactions, frictions, and expectations.

From this perspective, the ethnographic moments that might otherwise seem mundane — a robot being rebooted, a safety fence being bypassed, a user losing trust, an engineer adjusting parameters — become analytically significant. They are the sites where future human-robot relations are quietly negotiated.

The following reports build on this prophetic position. By tracing safety, trust, anomalies, sabotage, and the Sim2Real gap across In-the-wild settings, we do not aim to describe a single future of autonomous robots. Instead, we document how multiple possible futures are already being shaped in the present — relationally, experimentally, and unevenly.

This threshold position is the contribution anthropology brings to the study of robot autonomy: not a vision of what will come, but a careful reading of what is already beginning to emerge.



# Conclusion

This first report has set the scene for a series of ethnographic studies of robots in-the-Wild. In public, autonomy is read from behaviour and concealment, not the inner mechanical workings strived for by engineers. By tracing how autonomy has been imagined historically, performed publicly, defined technically, and approached methodologically, we have shown that robot autonomy is neither a simple technical milestone nor a distant science-fiction fantasy. It is a relational and ongoing achievement — one that already shapes how robots are designed, presented, and integrated into everyday life. The next three reports move from conceptual framing to empirical depth. Reports II, III, and IV follow robots into concrete settings where autonomy becomes consequential: workplaces and other shared environments. In these reports, we examine how autonomy is enacted and contested through five recurring themes introduced here: safety, trust, anomalies, sabotage and tinkering, and the gap between simulation and In-the-wild (Sim2Real). Each report focuses on specific contexts and interactions, showing how autonomy is sustained, challenged, or redefined through everyday practices of use, care, oversight, and repair.

Rather than asking whether robots are autonomous in a technical sense, these reports ask how autonomy is lived: how people come to rely on robots, how they adapt to their limitations, how breakdowns are handled, and how responsibility is distributed when things go wrong. Together, they reveal autonomy not as a stable property of machines, but as something continually negotiated between humans, technologies, and environments.

Report V returns to these empirical findings to offer a synthetic and forward-looking discussion. Drawing on the five themes across all case studies, the final report reflects on what current practices can tell us about the future of autonomous robotics, including technologies developed within RoboSAPIENS. From a prophetic perspective, it examines how present-day frictions and adjustments illuminate emerging forms of human-robot relations, raising questions about surveillance, rhythm of work, responsibility, legitimacy, and learning in socio-technical systems.

This matters not only for robotics research, but for society more broadly. As robots increasingly enter spaces of work, care, and public life, questions of autonomy become questions of governance, ethics, labour, and trust. Decisions about what robots should be allowed to do, how they should adapt, and who remains responsible for their actions cannot be resolved by technical design alone.

By bringing anthropology into the study of robot autonomy before it fully arrives, this report series aims to broaden the conversation. It invites engineers, policymakers, and the public to look beyond appearances of autonomy and attend to the social, moral, and relational work that makes autonomous systems function at all. In doing so, it offers not a prediction of a single technological future, but a grounded understanding of how multiple futures of human-robot coexistence is already being shaped in the present.



## References

- Ardener, E. (2017 /org 1975). *The Voice of Prophecy: And Other Essays*. New York, Oxford: Berghahn Books. <https://doi-org.ez.statsbiblioteket.dk/10.1515/9781785335570>
- Barrett, L. (2022). The possible in the Life and Work of William Grey Walter. In: *The Palgrave Encyclopaedia of the Possible*. Palgrave Macmillan, Cham. [https://doi.org/10.1007/978-3-319-98390-5\\_4-2](https://doi.org/10.1007/978-3-319-98390-5_4-2)
- Bernard R. (2006) *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. AltaMira Press.
- Bruun, M.; Hasse, C.; Hanghøj, S. (2015) Studying Social Robots in Practiced Places. In: *Techné: Research in Philosophy and Technology*. 19(2) Pp.143-165
- Čapek, K. (1923). *R.U.R. (Rossum's Universal Robots)*. London: Humphrey Milford.
- DeWalt K, DeWalt B. (2010). *Participant Observation: A Guide for Fieldworkers*. Altamira Press.
- Gantz, T. (1993) *Early Greek myth :a guide to literary and artistic sources*. London: Johns Hopkins University Press.
- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The Interpretation of Cultures* (pp. 3–30). Basic Books.
- Hastrup, Kirsten. "Postscript 1 – The Prophetic Condition". *The Voice of Prophecy: And Other Essays*, edited by Edwin Ardener, New York, Oxford: Berghahn Books, 2017, pp. 274-278. <https://doi-org.ez.statsbiblioteket.dk/10.1515/9781785335570-022>.
- Hasse, C. (2019). The multi-variation approach. *Paladyn, Journal of Behavioral Robotics*, 10(1), pp. 219-227. Retrieved 4 Jul. 2019, from doi:10.1515/pjbr-2019-0017
- Hasse, C. (2015). *An Anthropology of Learning: On Nested Frictions in Cultural Ecologies*. Springer.
- Hoffmann, E. T. A. (1908). *The Sandman*. (Original work published 1817). In *The Best Tales of Hoffmann*. Project Gutenberg.
- Ingold, T. (2017). *Anthropology and/as Education* (1st ed.). Routledge. <https://doi.org/10.4324/9781315227191>
- Kang, M. (2011). *Sublime Dreams of Living Machines: The Automaton in the European Imagination*. Cambridge, MA: Harvard University Press.
- LeBoutillier, A.E. (1999) W. Grey Walter and his Turtle Robots. *The Robotic Builder*. 11(5). P 1-3.
- Marcus, George. 1995. "Ethnography In/Of the World System: The Emergence of Multi-Sited Ethnography." *Annual Review of Anthropology* 24: 95-117. <http://dx.doi.org/10.1146/annurev.an.24.100195.000523>
- Mayor, A. (2018). *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*. Princeton: Princeton University Press. <https://doi.org/10.2307/j.ctvc779xn>



**Needham, J. (1991).** *Science and Civilisation in China: Vol. 2. History of Scientific Thought.* Cambridge University Press.

**Riskin, J. (2003).** The defecating duck, or the ambiguous origins of artificial life. *Critical Inquiry*, 29(4), 599–633. <https://doi.org/10.1086/377726>

**Sorenson, J. et al. (2019),** *Perspectives on Robots*, REELER Report, REELER Research Repository. <https://responsiblerobotics.eu/research/perspectives-on-robots/> (accessed May 3, 2021).

## Acknowledgements

The work presented here is supported by the RoboSAPIENS project funded by the European Commission's Horizon Europe programme under grant agreement number 101133807.

About the authors:

Cathrine Hasse is Professor of Learning and Technology at DPU, Aarhus University, Denmark, and research program leader of *Future Technologies, Culture and Learning*. She is also Honorary Professor of Techno-Anthropology at Aalborg University. Trained as a cultural anthropologist, her research focuses on learning, technology, and human-robot relations in real-world settings. She has led and coordinated numerous interdisciplinary research projects, including Horizon 2020 projects, working closely with researchers from engineering, psychology, and the natural sciences. She is the author of *An Anthropology of Learning* (Springer) and *Posthumanist Learning* (Routledge), and has published extensively on robots, learning, and technology in everyday life.

Pernille Maja Carlsen holds a bachelor's degree in Anthropology from Aarhus University. She is currently pursuing a master's degree in IT, Communication, and Organization at Aarhus School of Business and Social Sciences, where she is writing her master's thesis on hybrid work and digital communication. RoboSAPIENS is her first research project, and her interest in the field stems from earlier studies of robotics and technology. She is currently working as a research assistant and has been part of the project since the summer of 2024.

Peter Hommel Østerlund is a master's student of Anthropology at Aarhus University, from where he also got his bachelor's degree in Anthropology. He passed his bachelor's degree with a study on prosthetics and how people with prosthetics are affected by them in their daily lives. He is a student assistant on the research project, RoboSAPIENS, where he works closely with Cathrine Hasse and Pernille Maja Carlsen.