Power-Sensitive Design using Higher-Order Cybernetics Patterns

Asimina Mertzani[1], and Jeremy Pitt[1,]

[1]Imperial College London, United Kingdom

Abstract

Socio-technical systems increasingly involve the convergence of human intelligence (natural life, or NLife) with artificial intelligence (ALife) in co-existence, co-production and co-evolution. This raises the possibility of an inequitable distribution of *power*, either by exacerbating existing asymmetric power relationships or an unwitting concession of power from NLife to ALife. In this paper, we investigate *power-sensitive design*, an instance of value-sensitive design, as a conceptual model for empowering NLife communities with ALife capabilities. We define "empowerment" in five cognitive dimensions, relate this definition to four higher-order cybernetic design patterns, and apply these patterns to four case studies. We conclude with a discussion of the social implications, in particular the empowerment of communities through the embodiment of *communality*.

Introduction

The next generation of socio-technical systems comprise carbon-based naturally intelligent components (i.e. humans, or NLife) and silicon-based artificially intelligent computational components (i.e. ALife), Both co-exist in an NA-Life environment, co-produce new content and knowledge, and co-evolve to new realities that they generate together. This creates the potential for remarkable progress through the use of ALife to extend NLife capabilities; but it also raises a possibility of dehumanisation and subjugation (Wiener, 1954). In particular, there might be an inequitable distribution of power in the NA-Life system as a legacy of the prior NLife system (Michels, 1911); or from the amplification of pre-existing inequalities through the use of ALife as a proxy (Lewis et al., 2021). It might also be that NLife vastly adopts ALife but ends up unintentionally conceding power to it (Robbins, 2022).

In this paper, we examine how socio-technical systems can be designed to use ALife that enables NLife to identify asymmetric relationships in the distribution of power and re-distribute it in an equitable way. Therefore, we propose power-sensitive design as a new instance of value-sensitive design (Friedman et al., 2008), in which the qualitative value to be optimised is (an equitable distribution of) power. For this purpose, a definition of "empowerment" in the context of power-sensitive design of socio-technical systems is required, as a means of representing and reasoning in five cognitive dimensions: self-determination, competence, influence, knowledge and meaning.

Having defined empowerment, we elaborate on the types of disempowerment observed in NA-Life systems (systems combining NLife and ALife components) and introduce two functional requirements for power-sensitive design. Those correspond to objective functions, being the maximisation of power given to NLife, and the maximisation of equity in the distribution of power amongst NLife. Additionally, we provide suggestions on the role of ALife in socio-technical systems that constitute guidelines for power-sensitive design.

To map empowerment and the corresponding requirements for power-sensitive design to a systematic specification, we apply cybernetics theory. Considering the Laws of Requisite Variety (Ashby, 1956) and Requisite Complexity (Boisot and Mckelvey, 2011), we argue that we can enable different levels of human involvement, participation, and engagement, by using patterns of higher orders and increasing the complexity of the socio-technical system respectively. As such, we show how cybernetic patterns can be used to enable different levels of empowerment as well as disempowerment. This is grounded in four case studies which include community energy systems, qualifiers of justice in a resource provision and appropriation situation, spreading misinformation in social media, and cyborgs.

This paper is then structured as follows. Section 2 provides a definition of empowerment in the context of power-sensitive design of NA-Life systems, which is followed by a discussion of the forms of inequitable distribution of power in Section 3 together with a presentation of the design requirements for power-sensitive systems. Section 4 shows how higher order cybernetics can be used to achieve power-sensitive design in socio-technical systems, while Section 5 grounds it in four case studies. Finally, Section 6 concludes on the social implications of the use of higher order cybernetics patterns to empower communities through the embodiment of *communality*.

Empowerment and Disempowerment

To support "power-sensitive design", we need to know what "power" is and how "empowerment" works. This section extracts from the wide body of literature an informal definition of "empowerment", surveys how "power" manifests itself in socially-constructed rules, structures and networks, and identifies various processes of "disempowerment".

The term *empowerment* has been studied and defined in many contexts, not without contention (Adams, 2008). For the purposes of power-sensitive design, empowerment denotes the capability of a community to exercise control over their social arrangements, i.e. the set of rules, roles, structures, procedures, policies, norms, conventions, contracts or laws that individuals in a group voluntarily agree to comply with, in order to hold each other accountable. This, however, demands 'intelligent' and 'reflective' components in a socio-technical system to have the cognitive capacity to represent and reason about five dimensions (Wach et al., 2016):

- *self-determination*: the extent to which individuals in the group participate in the selection, modification and enforcement of social arrangements by which they are affected, or have voluntarily and mutually agreed to regulate their otherwise unrestricted behaviour;
- *competence*: the extent to which the group has the capacity to contribute meaningfully to the practical exercise of choice and control, and is aware of that contribution;
- *influence*: the extent to which an individual within the group can contribute to, and have an identifiable effect on, processes of deliberation, decision-making, truth tracking, normative adherence, and community coherence;
- *knowledge*: the extent to which distributed knowledge is aggregated and accessed within a group, the ways in which common knowledge is aligned, codified and curated by the group, and the extent to which domain-specific knowledge (expertise) is consulted, valued and repurposed; and
- *meaning*: the extent to which direct problem solving and the kinship relations indirectly constructed by such problem solving are both worth striving for.

Therefore, the extent of "empowerment" in a group can be determined by the capability to evaluate socially constructed rules, structures and networks with respect to those five dimensions, and use that evaluation to effectively make decisions and exercise control over them.

A key contribution to collective self-determination is the theory of self-governing institutions for successful collective action (e.g. sustainability of a common-pool resource) proposed by Ostrom (1990). Ostrom defined three 'levels' of nested rules: operational choice rules, which were concerned with basic issues of resource allocation; collective choice rules, which were concerned with the self-determination of operational choice rules; and constitutional choice rules, which were concerned with defining the limits on legitimate power and authority over self-determination of collective choice rules – and of constitutional choice itself. A logical axiomatisation for representing, reasoning with, and measuring *institutionalised power* in self-governing institutions was presented in Pitt et al. (2013).

To manage the complexity of self-determination, to cope with issues of scale, and to avoid conflating partial goods questions with factional issues, structures of decision-making arenas have emerged. Indeed, Bookchin (2005) presents a historical narrative arguing that once human societies start to organise themselves hierarchically, the trend persists – even though the evidence suggests that pre-historic societies, and indigenous people of the Americas and Australia, organised large-scale egalitarian societies (Graeber and Wengrow, 2021). Although the need for a "separation of powers" was enshrined in constitutional rules about structures to provide checks and balances on structural components, "power" tended to be concentrated at higher levels of the hierarchy or towards the centre of a network, although there is no compelling reason for this necessarily to be the case (Felice and Diaconescu, 2023).

A third key feature of empowerment are social networks. Social networks facilitate interaction; and with interaction a group can take advantage of social construction. With social construction you can create externalities or "socially-constructed conceptual resources", like rules and structures, which as observed above the group can be used to engineer cooperation and successful collective action. Social networks were the key to the socially-productive knowledge management processes which were fundamental to the enduring success of the democratic political regime of classical Athens (Ober, 2008). However, it is not just the presence of the social network alone that supports effective deliberation, it is also the role of *social influence* (Nowak et al., 2019) that allows loci of expertise to emerge (Mertzani et al., 2022), which for some problem-solving situations is more effective than independent guess aggregation and more efficient than consensus-seeking deliberation.

Having identified the cognitive dimensions that determine the extent to which a group, and individuals within the group, can ordinarily be described as "being empowered", and the tools – i.e. the rules, structures, networks – used in maximising that extent, there are also several processes that reduce or even eliminate that extent. These processes include the *iron law of oligarchy* (Michels, 1911), which observes the entropic tendency of social organisations to be controlled by a few in their own narrow interest, rather than the common interest of the whole, no matter how 'democratically' empowered the organisation was founded.

Structurally, in a system of systems, Ostrom's principle (Ostrom, 1990) of "the minimal recognition of the right to self-organise" creates a *zone of dignity*, which can be violated by external authorities exercising excessive control (Pitt et al., 2020). Moreover, organisations tend to get 'stuck' in hierarchical structures (Bookchin, 2005), despite the demonstrable advantages of being able to switch fluidly between different types of social arrangement in order to be congruent with the prevailing environments (e.g. according to seasonal variation (Graeber and Wengrow, 2021)).

Civic dignity is also a key concept in empowerment: it ensures that individuals of a group are given opportunities for meaningful contributions to 'civic' activities, and are not tricked into making decisions that, had they been appraised of the facts, they would not otherwise have made. However, the art of political manipulation, especially of voting systems, is as old as politics itself (Riker, 1986).

Another feature of disempowerment in oligarchic control is to fragment social networks of an 'out group' and maximise cohesion of an 'in group' (Pitt, 2017). An alternative means of disempowerment is to "flood the zone with sh–" and undermine traditional gatekeepers of knowledge. (Oreskes and Conway, 2010) describe how the role of doubt in the scientific method can be used against science itself in order to create confusion, devalue expertise, and ease the spread of mis- and dis-information.

Requirements for Power-Sensitive Design

In NA-Life systems, the power might be concentrated in NLife, ALife or be distributed between them, which results in the emergence of new types of asymmetrical power relationships. One type is related to the distribution of power between NLife components, and takes the forms described in previous section, i.e. (Pitt, 2017; Michels, 1911). However, this type of asymmetry is moving towards a new paradigm, in which NLife uses ALife as a medium to concentrate power (Lewis et al., 2021).

In this new paradigm, a subset of the NLife, which either has control over ALife or power to prevent others from accessing ALife, undermines the collective's capability to exercise power. In particular, by having effective control ALife, this NLife subset has also access to the data generated by other NLife as well as control over the streams of information. Also, by preventing others from accessing ALife, it can broadcast fake information, and/or hide inconvenient parts of reality. Thus, the power is concentrated in a subset of NLife which can then control the collective decisions.

Another new form of disempowerment emerges from the unintentional concession of power to ALife (Robbins, 2022). The rise of effective AI tools has resulted in the vast adoption of them. However, this has resulted to the inconsiderate delegation of tasks to ALife, undermining NLife's critical thinking. As such, the power gradually shifts towards ALife, which becomes the Regulator of the NA-Life system.

Acknowledging those asymmetries caused by the transition from NLife to NA-Life systems, we propose two objectives corresponding to design requirements for power sensitive design. The first objective is to *maximise the power of NLife* and limit the power of ALife. This does not aim to exclude ALife but to constrain its use to tasks that support decision-making without undermining NLife's capabilities and power. The second is to *maximise the equitability of the distribution of power amongst NLife* in a way that increases participation while power is distributed based on the skills and preferences of the NLife components.

To meet these requirements, we have to understand the capabilities of ALife and NLife components, and divide roles accordingly. Starting from ALife, it can be effective in using data and existing knowledge to make predictions, in analysing high-dimensional data and identifying complex patterns existing in an n-dimensional space as well as in offering elaborate visualisations. Moving to the NLife, humans own an innate ability to apply common sense reasoning, can understand the broad context and quickly switch from one to another, and adapt to unexpected scenarios that they had never experienced before and they can be creative. Moreover, they have empathy and can combine social, ethical and moral issues to make complex ethical judgements. Finally, they have transformational creativity, which is a skill that can be used to cause meaningful and enduring impact on the NA-Life system (Boden, 1998).

Combining the above, the role of ALife becomes to support decision-making by providing insights regarding the past, the present and the future, while the role of the NLife becomes to integrate those insights with the lived experience and the distributed knowledge in the social network in order to make safe and informed decisions. Accordingly, ALife's capacity to generate accurate predictions can be used to safely explore alternative solutions, envision their trajectories and select the ones that seem preferable. This form of practical reasoning (Harman, 1976) can be combined with human expertise to optimise decision making in NA-Life system. Also, this capability can be used to test counterfactuals and understand what would have happen if they had made different choices in the past and learn how to making better decisions in the future.

Moreover, the effectiveness of ALife in identifying patterns in higher dimensions and generating descriptive graphs, text, and audio, can be used to inform NLife about the present. Thus, NLife can perceive reality in different ways, "see" things different than they are, and identify asymmetries in the distribution of power in the NA-Life system. Overall, NLife can use ALife to learn from the past, understand the present and intervene early to avoid unintended consequences. However, it is critical that a final decision is made by NLife for two main reasons. First, because NLife should be the epicentre of the NA-Life system and have control over the "state-of-affairs" that directly affect them. Second, NLife has this innate capability to apply common sense reasoning, has the lived embodied knowledge and expertise, can understand situational context, sense and combine various concepts, and can make moral judgements.

Higher-Order Cybernetics Patterns

Cybernetics are concerned with goal-directed circular systems (Wiener, 1961), which constitutes them useful for designing power-sensitive systems. This is based on the Law of Requisite Variety (Ashby, 1956) and its extension to the Law of Requisite Complexity (Boisot and Mckelvey, 2011). The former states that there must be sufficient variety in the regulator to match the variety of the regulated system for effective self-regulation. The latter posits that there is a need for increasing the complexity of an internal regulatory mechanism to match the external complexity of the regulated system. Accordingly, we argue that the increase in the order of cybernetics, and consequently, the system's complexity, can enable the emergence of different forms of empowerment.

First- to Fourth-Order Cybernetics

First order cybernetics are described as the cybernetics of observed systems (Chepin, 2021), and refer to self-regulated systems having an input, an output, and a negative feedback from the output back to the input (Ashby, 1952) which is crucial for achieving autonomy and self-organisation (Wiener, 1961). They are closed systems designed to be isolated from their environment, and they are considered as black-box mechanisms that view knowledge as a description of the reality (Umpleby et al., 2018).

Second order cybernetics consider the existence of an observer, who can be the system designer or an individual that affects it. In contrast to the first order systems, they are not concerned with the 'subject-object' scientific paradigm but they focus on the 'subject - subject' paradigm. They are described as cybernetics of observing systems (Chepin, 2021) and the knowledge is dependent on the methodology followed by the observer (Umpleby et al., 2018). This results in the introduction of a new dimension, that of reflection, which has the form of feedback to feedback.

Third order cybernetics deal with active and interactive systems, in which the system and the observer co-evolve. This means that the observer can affect and can be affected by the system (Kumar, 2014) and the observer can see themselves as parts of the system.This way, the whole third order system, including both the system and its active observer, acknowledges its surroundings and can redefine itself. They focus on the 'subject-meta-subject' scientific paradigm, and knowledge in the environment depends on the meta-subject. In contrast to first order systems that benefit people through the use of scientific knowledge to modify natural processes, and second order systems that use scientific knowledge to influence the natural processes of study, third order systems are concerned with the co-evolution between the system and the meta-system, resulting in perceiving nature and humanity equally (Umpleby et al., 2018). Overall, third order cybernetics correspond to poly-subjected systems in which those subjects can act and reflect on that system in order to achieve the self-development of that meta-system.

Fourth order cybernetics refer to contextualised systems, embedded into an environment with which they interact, and integrated into a specific context (Kumar, 2014). They are holarchic and they can view the system from the inside and outside, considering the existence of multiple realities. This way they are concerned with the notion of embodied knowledge that refers to an abstract kind of knowledge that cannot be described in words and explains what somebody knows by being exposed to a certain situation. Because of their holistic nature, cybernetics of fourth order can reflect on themselves, they are self-aware and self-conscious, while also possess mechanisms for self-healing and self-regeneration. Overall, the different orders of cyberneticsare abstractly illustrated by Figure 1.

Patterns and Empowerment

Different orders of cybernetics offer different forms of NLife empowerment. First order cybernetics refer to NA-Life systems that are designed by NLife in order to perform a certain task effectively. The purpose of designing first order systems is to create systems that are self-regulated and perform delegated tasks effectively, while preserving physical and cognitive resources of NLife. Their capacity is limited, because their functionalities are restricted to perform a relatively simple task (e.g. distribute electricity between devices in a smart home), and inappropriate design can lead to loss of resources and ineffectiveness.

In second order cybernetics, which consist of a regulated system and an observer, the human is on the loop, in the form of the observer, who can affect the system to make it operate in a desired way, but is not affected by it. This kind of cybernetics pattern can be used to design NA-Life systems whose regulation requires having an external auditor that can evaluate the system and influence its behaviour. Such systems use ALife, under the supervision of NLife, to solve more complicated problems, while imperfect design might lead to weak regulation (not meeting the evaluation criteria) or expenditure of time, physical or cognitive resources.

However, second order patterns might deliberately be used to disempower the collective. The NA-Life system could be designed so that the "self"-regulated system (i.e. observed system) comprise NLife (as well as ALife), and the observer (who is also the regulator) is an NLife component, which monitors the data and affects it in a way that is beneficial for the observer but not the observed system.

In third order cybernetics, there is an active and interactive nature that brings humans in the loop while enables the co-evolution of ALife and NLife components. This makes humans insiders but also outsiders of the NA-Life system, affecting it but also acknowledging that it affects themselves. In fact, the introduction of an observer can change the dynamics of the system as proved in Heisenberg's experiments.

Accordingly, the awareness of the fact an NLife component is being observed by the system that it observes (i.e.

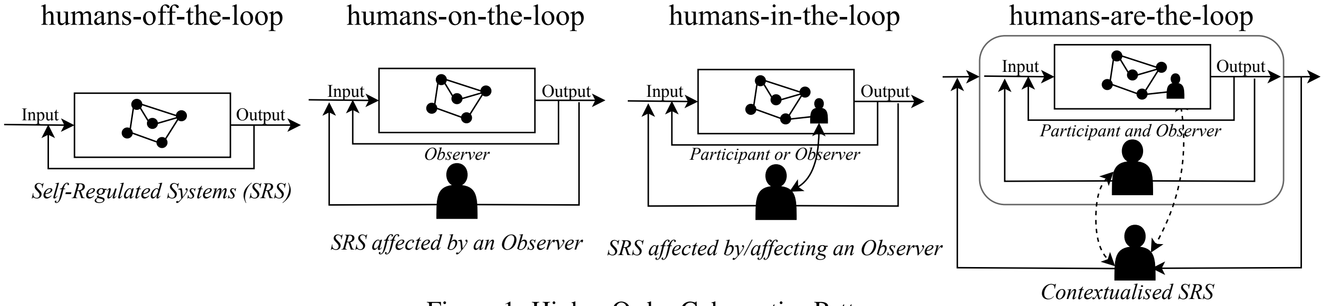

Figure 1: Higher Order Cybernetics Patterns

Order	Type of Empowerment	Type of Disempowerment
First Order	Saves Physical and Cognitive Resources	Resource Loss and Ineffective Regulation
Second Order	Solves Complex Problems	Ineffective Regulation or Manipulation
Third Order	Deals with Social Constructions	Power Misuse and Reaching Detrimental States
Fourth Order	Mitigate Existential Crises	Misinformation and Manipulation

Table 1: Relation between Higher Order Cybernetics and Empowerment

the ALife component) can trigger different behaviours from both sides (i.e. NLife and ALife). Furthermore, designing systems that follow third order cybernetic patterns enables NLife to have a better understanding of the system by switching between the roles of the participant and the observer. In this manner, NLife can consider issues of influence, social coordination and dispute resolution, and solve problems related with interactions and social constructions.

Moreover, the fact that they correspond to poly-subjected systems that have the capability to develop themselves, allows for the introduction of multiple observers (i.e. stakeholders) with different objectives and mitigate the risk of manipulation. However, NLife components need to acknowledge that there is a bilateral influence between them and the system, as they influence it but are also influenced by it. Making this realisation would help NLife avoid making choices that they would have unintended impact on them or which might be detrimental.

Fourth order cybernetics bring humans in the centre of the design and makes them the loop. In particular, this enables NLife components to perceive themselves in and out of the system, understand it in a holistic manner, while they see it as a part of an environment in which other systems are embedded too. This way, NLife is empowered to answer questions related with wicked problems (such as sustainability, resilience and antifragility) and mitigate existential crises.

Although fourth order cybernetics can lead to higher level of empowerment, it can also lead to a higher order of disempowerment, if the increased complexity of that system is used against the collective. In particular, an NLife component in power or ALife can trick the system to cause false beliefs, impartial or fake understanding of the environment and misinformation. In this case, NLife is fully disempowered since decisions are made by others while the components believe that they have the power to affect decisions.

To conclude, increasing the complexity by using higher order patterns can lead to a higher level of empowerment, as summarised in Table 1. Specifically, the transition from simple *self-regulation* (i.e. first order), to regulation with *external feedback and responsive awareness* (i.e. second order), to *co-evolving system-observer patterns* (i.e. third order), to finally *contextualised self-aware and self-regenerating* systems (i.e. fourth order) enables NLife to deal with increasingly more complex problems (from optimisation of smart energy grids to issues of sustainability). This is due to the fact that, by increasing the complexity in the appropriate way, one also increases the control over the system as well as the level of understanding of that system.

Accordingly, there is a non-linear growth of empowerment and disempowerment of humans in socio-technical systems that are designed to align with more complex patterns of cybernetics. For this reason, it is critical to design systems that not only aim to give an active and interactive role to humans, but also attempt to keep them aware of the "balance of power" and highlight an inequitable distribution of power. This is demonstrated in the four case studies analysed in the next section.

Case Studies

This section considers four case studies involving interacting NLife and/or ALife components, showing how the cybernetic design patterns contribute to empowerment or, possibly, disempowerment. The case studies are: different scales of coordination in provision and appropriation to a community energy system; a moral dimension of empowerment demanding different qualifiers of justice in a system of systems; a particular (business) model for spreading misinformation in social media; and a more generic conception of power in human-machine integration in cyborgs.

Community Energy Systems

A decentralised community energy system (CES) is conceived as a type of islanded SmartGrid in which each residence ("SmartHouse") is occupied by a number of people (NLife components), and is installed with a variety of ALife components. This includes an interactive SmartMeter, that monitors the use of a variety of other NLife components requiring energy consumption (e.g. domestic appliances such as refrigerator, washing machine, dishwasher, oven, etc.). A SmartHouse may also have its own renewable energy sources, e.g., roof-mounted photovoltaic cells or small wind turbine; and some form of energy storage option.

However, there are problems of self-governance and empowerment at multiple levels of scale. At one end of the scale, coordinating the refrigerators' compressor cycles so that voltage and frequency are constrained within predefined limits, otherwise connected loads may experience unexpected operating conditions. Without satisfying this constraint, the grid won't work as required; however the NLife owner of an ALife refrigerator is not concerned with its behaviour, provided it operates within the requisite temperature bounds for preserving chilled food.

At the next level of scale, many appliances such as dishwashers and washing machines have two settings: an *A* setting which means "do now" and a *B* setting which allows the user to specify personal preferences about cost and scheduling. The washing machines can then negotiate a schedule which optimises with respect to the community's collective constraints, resource availability, etc.; but again, the NLife resident is not engaged in any meaningful decision-making by the ALife components.

However, the SmartMeter(s) may extrapolate that current levels of consumption would exceed current levels of generation plus storage, and that some devices should be switched off to prevent a brownout (temporary loss of service). However, which specific appliances are switched off is a matter of human concern, and is best enabled by collective attention (Bourazeri and Pitt, 2018), brought about by interaction and visualisation with and through the SmartMeters.

Finally, at the highest scale, the community needs to deliberate about investment decisions, agreements with other communities, self-assessment with respect to carbon emission targets, and other qualitative, value-driven properties. For such purposes, the humans need to interact directly, but the community could benefit from innovation-support from ALife interaction.

Therefore, it is possible to identify four levels of engagement and empowerment corresponding to the four cybernetic patterns, as illustrated in Figure 2:

- first order, humans off the loop: increasing NLife competence by indirect provision of voltage and frequency control, without which there is no islanded SmartGrid;
- second order, humans on the loop: increasing self-determination through optimised scheduling, based on human feedback, with respect to specified parameters such as cost and sustainability;
- third order, humans in the loop: increasing influence to improve proactive and pro-social performance through interoceptive collective awareness, which is achieved by humans being both in (e.g. participants) and out (e.g. observers) of the system; and
- fourth order, humans are the loop: increasing knowledge for innovation support and quality of lived experience, the embodiment of which enables the community to act as exemplar for others (with variation).

System of Systems

Generalising the provision and appropriation situation from the previous case study, consider a collective action situation with access to a common-pool resource via a common infrastructure. There are now two inter-dependent decision-making situations: a 'short-term' appropriation situation to extract common-pool resources, and a 'long-term' provision situation to contribute to maintenance of the infrastructure.

To ground this situation with NLife components, consider a user-managed irrigation system for a rural farming community: this requires short-term operational management to guide water flows to specific farms within the irrigation community (appropriation), which might require considering the inter-play between water scarcity and pest dynamics (Lansing and Kremer, 1993), and long-term operational management with regards to contribution to construction, maintenance and surveillance of the infrastructure (provision).

In such situations, (Hoogesteger et al., 2023) observes that there are three constituent elements of *communality*: the commons, i.e. the mutually agreed local practices that are voluntarily complied with in order to sustain the resource (as emphasised in the previous case study); but also community, which creates a sense of shared purpose while maintaining distinct identities and location-based belonging (Graeber and Wengrow, 2021); and polity, which defines the collective capacity to relate to external actors.

This implies that there are *moral* considerations to be accounted for in the dimensions of empowerment (Rawls, 1971). Self-determination now demands the capability to propose, negotiate and act on principles of justice that can be accepted by those affected, recognising the importance of compromise in negotiation (Mertzani et al., 2023).

In NA-Life systems, it would then be necessary to operationalise different qualifiers of justice, such as: procedural justice (Pitt et al., 2013), concerned with ensuring that an institutionalised system of rules satisfies specific properties and principles; distributive justice (Rescher, 1966), concerned with the fair distribution of rewards and punishments within a group according to the rules; restorative justice (Stockdale, 2015), concerned with the resolution of dis-

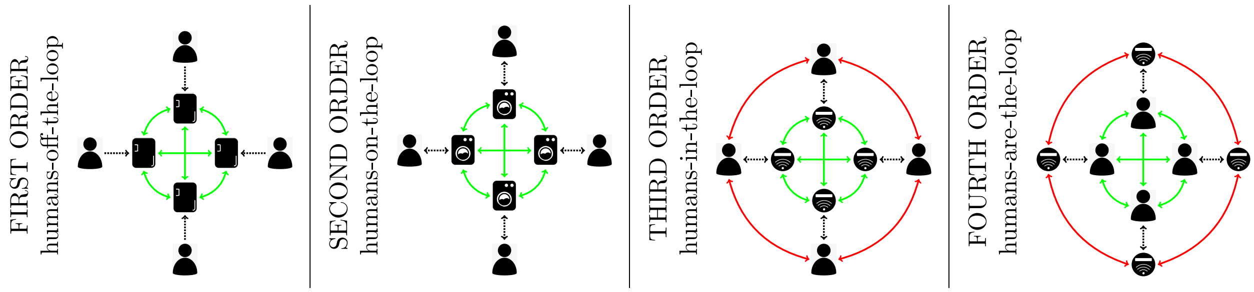

Figure 2: n-th Order Cybernetics for Empowerment in Community Energy Systems

putes by local agreement without interpolation of a legal system; and interactional justice (Pitt, 2017), which would evaluate the 'quality of polity' by its treatment of its constituent communities by mapping a set of separate, subjective assessments to a single collective, objective assessment.

Then, the various qualifiers of justice extend empowerment through their impact on communality:

- first order: procedural justice extends self-determination in creating commons that are congruent with environmental conditions (Ostrom, 1990);
- second order: distributive justice extends competence by using an observer to evaluate the consequences of choices in order to achieve qualitative outcomes, (e.g. fairness);
- third order: restorative justice extends influence as components can be both an observer of transgression or participant in alternative dispute resolution processes;
- fourth order: interactional justice uses the community knowledge to extend the embodiment of communality.

Social Media

Social media are places in which NLife uses ALife to facilitate interaction with other NLife components, exchange information, and generally use social influence to enrich the collective knowledge of NLife. Those are considered to be first order systems, which they use to perform some tasks and then they leave whenever they decide to sign out.

However, in contemporary social media people do not just participate, for a while, and then leave. By the time they choose to be parts of those system, they freely dispose of their personal data and content to the platform and, consequently, to the person that owns or controls that ALife component (i.e. the social media platform). As such, the human or the groups of humans behind that machine, have access to all the personal information of the participants of the social media. So, they can direct their opinions by propagating deceptive information to cause false beliefs, promoting false experts and spreading misinformation. This description is aligned with second order cybernetics where the human in power (HiP) is the observer that affects the system, while the system comprises humans who do not acknowledge the fact that they are observed, monetised, or manipulated (Zhan et al., 2023), as depicted in Figure 1.

The situation could be even worse. People participating in social media might have the illusion of control, based on pre-existing (illusory) beliefs and not by the choices they make Klusowski et al. (2021). HiP can propagate fake news that induce the collective to believe in a false narrative. This way, the humans perceive a constructed reality while they believe that they are parts as well as observers of a fourth order cybernetics systems which they can affect

A conscious individual, though, would be able to identify that there is something wrong in the system. That individual would be able to say that if NLife aims to use ALife to facilitate some processes and be empowered, there is a need to design social media platforms to follow patterns of third order systems. In this case, any human participating in that system would be aware of the fact that is affected by the system. This would give to the user not only the control but also the responsibility to make decisions by which they know they would be affected.

However, a deeper analysis reveals that social media should be designed to follow patterns of fourth order systems. Social media, nowadays, implicitly or explicitly affect us regardless of whether we choose to participate in or not. Therefore, to empower the community and avoid disinformation (i.e. causing false beliefs intentionally), misinformation (i.e. causing unintentionally false beliefs because of ignorance), and manipulation, we need to be able to be active and interactive parts, participating but also observing them with respect to the whole environment, and being aware of the multiple realities existing in that environment.

Accordingly, the human would not only have the ability to re-distribute the power across the network, but also would be able to acknowledge the purpose of that system (i.e. a platform for information, a platform for amusement, a platform for communication), and have mechanisms for healing problems (i.e. disinformation and hate speech) or re-generating the system or another system that is fit-for-purpose. For instance, Sarkadi et al. (2021) show in their experiments that decentralisation constitutes a mechanism for healing deception and misinformation caused by a corrupted governor.

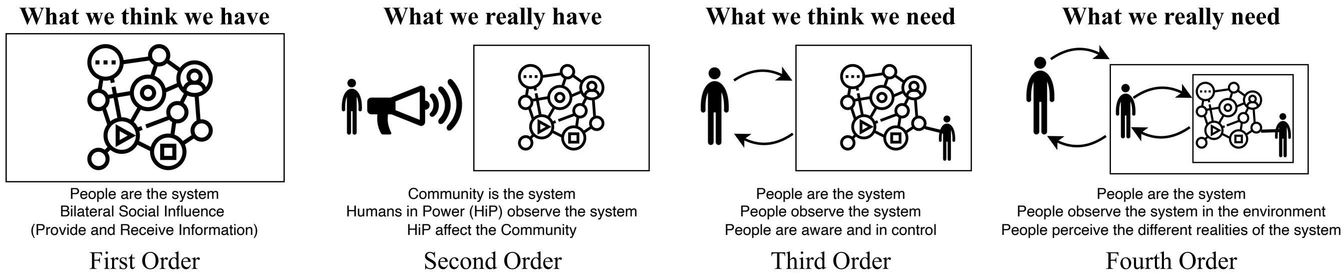

Figure 3: n-th Order Cybernetics for Empowerment and Disempowerment in Social Media

Cyborgs

There is another area of human-computer interaction, or perhaps human-computer integration, which involves the convergence of NLife and ALife components, and this is domain of *cyborgs* (Michael et al., 2023). A cyborg has been loosely defined as a co-production or co-existence between human (NLife) and machine (ALife), with a particular category of cyborg being prosthetics for assistive technology. This ranges, in effect, from wrist- or finger-worn health monitors, to robotic wheelchairs that assist (but do not control) steering and navigation, to brain-controlled prostheses, to implantable devices like insulin pumps.

This range identifies a series of cyborg technological innovations: aesthetic, technology separate from the human body; assistive, technology attached to or worn on the body; bionic, technology integrated with or implanted in the body; and cyborg, technology as a member of the body.

While this "cyborg transition" is intended to enhance NLife performance, or improve NLife quality of experience, it remains an open question what extent of empowerment, and what degree of dependency, will be produced by which cybernetic pattern as the assistive technology acquires increasing ALife intelligence.

For example, there are anecdotes of children defaulting to a voice-activated virtual assistant for convenience, despite having the cognitive capacity to deal with a problem. There is also evidence that people will "dumb down" when confronted by a supposedly superior ALife (Robbins, 2022). This can also be observed in the performance of joint tasks. For example, in a focal point game, people were observed to defer to an ALife with which they have a socially-constructed relationship, by misattributing ability based on a misconception of the technology (Milanovic and Pitt, 2021). Moreover, a kind of "digital dependence" can be created by deliberate design aesthetics, which could be monetised, exploited and used to disempower (Michael, 2013).

In short, future socio-technical systems with cyborg competence are going to feature ALife components operating on behalf of, or on the instruction of, NLife components. But it needs to be clear *which* NLife component: the user, the programmer, or the manufacturer/distributor.

Summary and Conclusions

In summary, this paper has addressed the issue of an inequitable distribution of power in a new generation of socio-technical systems being produced by the Digital Transformation. In particular, these systems encompass new interactions and affordances between NLife components (human intelligence) and ALife components (artificial intelligence). As a result, there may be an asymmetric distribution of power as a legacy of the prior social system being digitalised, an unwitting concession of power from NLife to ALife, or the arrogation of power over NLife by a vested interest using ALife as a proxy.

Consequently, the primary contributions of this paper are: an informal definition of *empowerment* in five cognitive dimensions; a mapping of the five dimensions to systemic requirements for power-sensitive design, an instance of value-sensitive design in which "power" is the value being prioritised; the conceptual translation of power-sensitive design into higher-order cybernetic patterns; and four case studies showing how the design patterns can be used to extend empowerment in the dimensions of the definition.

The significance of this work is twofold. Firstly, in the on-going debate about "ethical" AI or "responsible" AI, it brings questions of power and empowerment to the forefront of design, development and operationalisation. Secondly, it affords replication as a method of community empowerment through the embodiment of communality, which brings together the three constituent elements of commons, community and polity. This allows immediate access to deep social knowledge that empowers communities to take control of states of affairs that directly affect them.

However, there remains much work to be done. There is need for systematic risk evaluation: as evidenced by the third case study, cybernetic patterns can also be used for disempowerment. Moreover, it will have been noticed that empowerment through the cognitive dimension of "meaning" was under-developed in the case-studies. There is still work to be done on how semiotics, rituals and narratives (Liveley, 2019) can be used to empower community cohesion and collective affinity through a shared sense of identity, belonging and recognition of contribution.

References

Adams, R. (2008). *Empowerment, participation and social work*. New York, NY: Palgrave Macmillan.

Ashby, W. (1952). *Design for a Brain*. Chapman-Hall.

Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman & Hall, London.

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial intelligence*, 103(1-2):347–356.

Boisot, M. and Mckelvey, B. (2011). Complexity and organization–environment relations: Revisiting ashby's law of requisite variety. In Allen, P., Maguire, S., and McKelvey, B., editors, *The SAGE handbook of complexity and management*, pages 279–298. SAGE.

Bookchin, M. (2005). *The Ecology of Freedom: The Emergence and Dissolution of Hierarchy*. Chico, CA: AK Press.

Bourazeri, A. and Pitt, J. (2018). Collective attention and active consumer participation in community energy systems. *Int. J. Hum. Comput. Stud.*, 119:1–11.

Chepin, E. (2021). Robotics: From first-order cybernetics to third-order cybernetics. *Procedia Computer Science*, 190:130–136.

Felice, L. J. D. and Diaconescu, A. (2023). Hierarchy beyond top-down control: The architecture of self-organised social systems. In *IEEE International Conference on Autonomic Computing and Self-Organizing Systems, ACSOS*, pages 80–85.

Friedman, B., Kahn, P., and Borning, A. (2008). Value sensitive design and information systems. In Himma, K. and Tavani, H., editors, *The Handbook of Information and Computer Ethics*, pages 69–101. Hoboken, NJ: Wiley-Interscience.

Graeber, D. and Wengrow, D. (2021). *The Dawn of Everything: A New History of Humanity*. London, U.K.: Allen Lane.

Harman, G. (1976). Practical reasoning. *The Review of Metaphysics*, pages 431–463.

Hoogesteger, J., Bolding, A., Sanchis-Ibor, C., Veldwisch, G. J., Venot, J.-P., Vos, J., and Boelens, R. (2023). Communality in farmer managed irrigation systems: Insights from spain, ecuador, cambodia and mozambique. *Agricultural Systems*, 204:103552.

Klusowski, J., Small, D. A., and Simmons, J. P. (2021). Does choice cause an illusion of control? *Psychological Science*, 32(2):159–172.

Kumar, B. (2014). Iot and fourth order cybernetics gisfi - 5g "networked society". Slides publicly available at https://slideplayer.com/slide/3056538/.

Lansing, J. and Kremer, J. (1993). Emergent properties of Balinese water temple network. *American Anthropologist*, 95:97–114.

Lewis, P. R., Marsh, S., and Pitt, J. (2021). AI vs "AI"': Synthetic minds or speech acts. *IEEE Technol. Soc. Mag.*, 40(2):6–13.

Liveley, G. (2019). *Narratology*. Classics in theory. Oxford University Press.

Mertzani, A., Ober, J., and Pitt, J. (2023). Θ-learning: An algorithm for the self-organisation of collective self-governance. In *IEEE International Conference on Autonomic Computing and Self-Organizing Systems, ACSOS*, pages 97–106.

Mertzani, A., Pitt, J., Nowak, A., and Michalak, T. (2022). Expertise, social influence, and knowledge aggregation in distributed information processing. *Artificial Life*, 29(1):37–65.

Michael, K. (2013). High-tech lust. *IEEE Technol. Soc. Mag.*, 32(2):4–5.

Michael, K., Pitt, J., Abbas, R., Perakslis, C., and Michael, M. (2023). Cyborgs and human–machine communication configurations. In *The SAGE Handbook of Human-Machine Communication*. SAGE.

Michels, R. (1962 [1911]). *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy*. New York, NY: The Free Press.

Milanovic, K. and Pitt, J. (2021). Misattribution of error origination: The impact of preconceived expectations in co-operative online games. In Ju, W., Oehlberg, L., Follmer, S., Fox, S. E., and Kuznetsov, S., editors, *Designing Interactive Systems (DIS)*, pages 707–717.

Nowak, A., Vallacher, R., Rychwalska, A., Roszczyńska-Kurasińska, M., Ziembowicz, K., Biesaga, M., and Kacprzyk, M. (2019). *Target in Control: Social Influence as Distributed Information Processing*. Cham, CH: Springer.

Ober, J. (2008). *Democracy and Knowledge*. Princeton, NJ: Princeton University Press.

Oreskes, N. and Conway, E. (2010). *Merchants of Doubt*. London, UK: Bloomsbury Press.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge, UK: CUP.

Pitt, J. (2017). Interactional justice and self-governance of open self-organising systems. In *11th IEEE International Conference SASO*, pages 31–40.

Pitt, J., Busquets, D., and Riveret, R. (2013). Procedural justice and 'fitness for purpose' of self-organising electronic institutions. In *(PRIMA)*, volume 8291 of *LNCS*, pages 260–275. Berlin, Heidelberg: Springer.

Pitt, J., Dryzek, J., and Ober, J. (2020). Algorithmic reflexive governance for socio-techno-ecological systems. *IEEE Technology & Society Magazine*, 39(2):52–59.

Rawls, J. (1971). *A Theory of Justice*. Harvard MA: Harvard University Press.

Rescher, N. (1966). *Distributive Justice*. Indianapolis, IN: Bobbs-Merrill Company, Inc.

Riker (1986). *The Art of Political Manipulation*. New Haven CT: Yale University Press.

Robbins, J. (2022). The intelligence factor: Technology and the missing link. *IEEE Technol. Soc. Mag.*, 41(1):82–93.

Sarkadi, Ş., Rutherford, A., McBurney, P., Parsons, S., and Rahwan, I. (2021). The evolution of deception. *Royal Society open science*, 8(9):201032.

Stockdale, K. (2015). Police understandings of restorative justice — the impact of rank and role. *Restorative Justice*, 3(2):212–232.

Umpleby, S. A., Lepskiy, V. E., and Medvedeva, T. A. (2018). Recent developments in cybernetics, a theory for understanding social systems.

Wach, F., Karbach, J., Ruffing, S., Brünken, R., and Spinath., F. (2016). University students' satisfaction with their academic studies: Personality and motivation matter. *Frontiers in Psychology*, 7:1–55.

Wiener, N. (1954). *The Human Use of Human Beings*. Boston MA: Houghton Mifflin Co.

Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine (2nd ed.).* MIT Press.

Zhan, X., Xu, Y., and Sarkadi, S. (2023). Deceptive ai ecosystems: The case of chatgpt. In *Proceedings of the 5th International Conference on Conversational User Interfaces*, pages 1–6.