Self-governance by Cyberconferencing

A Path to Direct Democracy via the Behavioral Sciences

Francis Mechner

Mechner Foundation

Preface

The impetus for this undertaking requires a few words of explanation. The author is a behavioral scientist who is watching humanity drift toward what could be the destruction of its only habitat by climate change and wars waged with weapons of mass destruction.[1]

Humanity is in a quandary. There has been much said and written about multiple looming threats, but the proffered solutions for confronting them are invariably insufficient, naive, or fatalistic.

To a behavioral scientist, a constructive approach must identify the problem and describe it objectively, naturalistically, and without preconceptions about how it might (or might not) be solved. It entails acceptance that we might not have an answer when faced with a dangerous situation.

But that is where we are. If we start, as we must, with the premise that human behavior is responsible, it follows that it is on this behavior that scientists must focus. But on whose and which behavior?

The power node concept

We know that human societies generally structure themselves into groups, however those groups are defined. We also know that there are groups that inevitably wield more power than others.[2] We will refer to groups that possess outsized power and influence as *power nodes*. Prominent examples of power nodes in present-day societies are political parties, corporations, wealthy families, organized religions, military groups, heads of state, and groups defined by particular belief systems. Power nodes may also fuse into enlarged units like nations that, in turn, function as power nodes within societies of nations.

---

[1] In recent years, I've witnessed something I had not seen since 1938. In that year, the Nazis took over my native Austria, followed by WW2 and the destruction of European civilization as we had known it. More recently, I've recognized echoes of the ominous precursors of the events of the 1940s, this time threatening *all* of humanity.

[2] REFERENCES

The concept is defined in the present context by the attributes that power nodes share and the ways in which they differ. A defining feature that all power nodes share is that they act to further their own interests, sometimes at the expense of the society as a whole. This feature often leads to problems and conflicts, including wars, within a nation's boundaries and across international boundaries. Power nodes usually have leaders at their helm, and under the helm, there may be hierarchies of additional power nodes.

Power nodes differ widely in size and the forms they take, depending on the culture, era, and type of society they inhabit. The power of some is mainly economic while that of others coalesces around various types of bonds—ideological, ethnic, racial, cultural, political, military, or religious.

The elusiveness of the concept. The power node concept is not an obvious one, despite its central importance. The diversity of power nodes can obscure their commonalities, while their familiarity, proximity, and omnipresence make it easy to "miss the forest for the trees." It is also a counterintuitive concept because it encompasses society's most benevolent and most malign entities, its most beloved leaders and most despised villains. But once power nodes are identified as a valid concept, behavioral scientists can move on to studying their omnipresence in human societies.

Comparing power nodes to tribes. Power nodes and human tribes both exhibit internal cohesion, mutuality, loyalty, and cooperation. Both tend to vie for power within the societies they inhabit. Both seek to direct assets and resources to themselves and their members. Both grant superior rights and privileges to their leaders. Both may act with antagonism, and sometimes violence, toward non-members. And both act in accordance with temporal discounting (more on that in Chapter 1).

Then, there are ways in which tribes and power nodes differ. The origins of power nodes are usually socio-economic or military, rather than ancient as in the case of tribes. The life spans of power nodes are therefore usually shorter and their internal bonds weaker.

Societal norms that surround power nodes. A norm that pervades human societies is a general acceptance of the self-benefitting behavior of power nodes and the superior entitlements of the most powerful. This norm may be traceable to primordial times, when authoritarian leadership structures of tribes often aided survival though also often led to conflict.[3] In modern times, elements of the industrial revolution gradually transformed the superior entitlement norm into societal *isms*—socialism, capitalism, communism, authoritarianism, and fascism—whose governance structures are differentiated primarily by the roles and locations of their power nodes within their respective societies. Yet all of these *isms* share the norm that power nodes have superior rights—quasi parasitic rights. In today's world, this norm takes on an ominous new meaning and entails serious dangers.

The danger power nodes can pose to their societies.

The main danger power nodes pose is their power itself, as power corrupts. The main mechanism by which power corrupts is temporal discounting. Corrupt acts usually bring immediate or near-term reinforcement, while their possible punishment is delayed and uncertain, and therefore likely to be discounted by the perpetrator. When the discount causes the punishment to be outweighed by the immediate reinforcement, the result is a corrupt act.

An important way power nodes assert their special interests or agendas is by achieving influence or control over governance processes. For instance, when scientists began to issue dire warnings of the long-term impact of certain activities on the human habitat (e.g., burning of fossil fuels and forests), many affected power nodes used their influence to oppose societal responses. They did whatever they could to protect their near-term interests while discounting the long-term dangers to society and the human habitat.

[3] Governance and legal codes are attempts to resolve these conflicts without resorting to force. As the size of human societies increased, ancient rulers like Hammurabi created legal codes to make the settlement of disputes more efficient. In subsequent centuries, even though there arose the concept of legal precedent and institutions like judiciaries and legislatures, conflict among power nodes persisted.

In the United States, power nodes are able to garner control or influence through political means—they can fund the installation of intermediaries or align their actions with those of similarly motivated power nodes or political parties. They can disseminate disinformation to divide their enemies and sow chaos to weaken them. In extreme situations, when threatened with a loss of power, influence, and resources, they might even perpetrate coups or insurrections.

Aggression, striving for dominance, avarice, and violence are widespread behavioral characteristics of individuals and human groups. Their origins may date back to our evolutionary ancestors' competition for food, water, shelter, and procreation. These characteristics evolved in tandem with human intelligence and the development of major scientific and technological achievements. Some of these achievements were weapons of unprecedented destructiveness that pose threats to the human habitat. The accumulation of weapon arsenals in the hands of numerous power nodes augments the danger of their actual use.

The dilemma

Despite the threats they can pose, power nodes cannot be eradicated as they are the connective tissue of the societies that harbor them. Yet they now threaten not only their host societies but also the entire habitat of the human species, and perhaps its very survival. Is there a way to resolve this dilemma?

Homer's adage "We cannot change the wind but we can change the setting of the sails'' suggests an approach. Like the wind, power nodes cannot be changed, but there may be a governance architecture that prevents them from asserting their influence and self-interest. The present work describes an approach by which the behavioral sciences might be able to address this dilemma: making the society self-governing.

Chapter 1

The increasing danger that certain power nodes pose necessitates a protective governance architecture. Biologists could point out that such architectures already exist and are prevalent in all of Earth's autonomous life forms. Beehives, ant colonies, coral reefs, plants, aggregates of cells, and our own bodies are all leaderless, and in that sense self-governing.

One reason for the success of autonomous, self-governing life forms is the effectiveness of its architecture. It provides protection against attacks and intrusions and simultaneously the freedom to take available advantageous actions. These life forms possess physiologically encoded features that enable autonomous functioning.

Leaderless self-governance

The present work addresses if and how self-governance is applicable to large human aggregates like societies, nations, or even the entire world. It explores whether such aggregates can function without leadership or a power node at the helm. If so, the aggregates could then be described as a new life form. Unlike biological autonomous life forms, however, they would lack a physiological self-governance machinery. In order to achieve self-governance, they would have to be provided with an artificial, technology-based machinery.[4]

The model discussed here seeks to show that this type of artificial self-governance machinery is within the reach of recently emerged technologies. They include the internet and digital platform-managed cyberconferencing, and would be based on an amalgam of AI and all of the life sciences, with the behavioral sciences playing a special role.

To help determine whether self-governance of large human aggregates is possible, we have drafted a model that can be scrutinized. Regardless of

[4] An obvious concern is that the use of AI in governing functionalities might lead to mis-governance, but such concerns have no more basis than concern about a self-driving car going rogue. In the design of AI-type self-governance software, all decisions about goals or objectives would be made and monitored exclusively by the populace. (REFs)

whether this "*in vitro*" model is ever implemented, its existence permits a systematic assessment of viability.

Following Chapter 1, which describes the self-governance model in some detail, there are overviews of subsequent chapters, one of which explores the applicability of certain platforms software components to contemporary societal needs. The design of components would present novel challenges to the field of computer linguistics and the science of verbal behavior. These challenges can draw that science into the new domain of verbal discourse in cyberconferencing, specifically, goal-directed cyberconferencing in which the objective is to solve a problem, address an issue, obtain information, or reach consensus or agreement.

Another chapter describes the Mechner Foundation's current goal of building a team that will identify, analyze, develop, field test, and implement the highest priority software modules. A final chapter discusses how a self-governance technology might resuscitate conversation about the theoretical possibility of world self-governance.

Power nodes and temporal discounting. Since most wise governance involves long-range planning that balances near-term costs against far-off rewards, temporal discounting is always in play. For instance, scientists have trumpeted distant dangers like climate change. Temporal discounting by power nodes often discounts these warnings in favor of the immediate reinforcement they can glean by burning fossil fuels and forests. Governance on behalf of an entire society's interests involves the suspension of biologically driven temporal discounting. It can mean, at times, lower prioritization of near-term interests and taking distant dangers more seriously—actions that would be unpalatable to legislators susceptible to biologically normal temporal discounting.[5]

[5] *Temporal discounting* is a biological adaptation to a statistical reality of the natural world: the longer an event is delayed, the less likely that it will occur, because delay provides time for intervention from nature's perpetual stream of random events. All species evolved in continuous contact with this property of the universe, and most higher species adapted to it by evolving the behavior termed *temporal discounting*. REFs. the Critchfield book and two other refs to the temporal discounting literature.

Platform-managed cyberconference panels are devoid of the biological characteristics of human groups. They can therefore make far-sighted decisions when necessary and govern rationally and on behalf of the entire society's long-term interests. This is in contrast to human groups, which often value rewards according to their proximity (prioritize immediate consequences and discount threats according to their distance while ignoring their severity). We have only to consider political parties and corporations that scramble to satisfy the short-term goals and demands of voters, investors, or aggrieved supporters.

Decision making by random samples of the populace. In the model, a random sample of an expert-assisted populace, uninfluenced by power nodes, achieves balance among long-term and near-term interests. This feature accommodates long-term coexistence of power nodes in non-predatory relationships with each other, with the rest of society, and with Earth's environment. Can a model that excludes power nodes from governance produce such a society? Is it achievable even in principle?

The present chapter explores these questions by describing a governance model that can be examined and evaluated. The model is built on digital platforms that could potentially imbue it with self-governance features and functionalities through which the populace can engage in direct democracy and govern without intermediaries.

A novel technology: platform-managed cyberconferencing

Platform-managed *cyberconferencing* already exists as an amalgam of the internet, digital technology, and distributed networking. To adapt it for governance use, the model enhances it with digital platforms that have capabilities for reading and understanding text. Contributions from computer linguistics and the science of verbal behavior enable digital platforms to moderate the verbal discourse aspects of cyberconferencing.

The model places all governance functions in the hands of panels of *randomly selected members of the populace,* beyond the reach of power node influence. A key to such populace panels' ability to govern is the

expert assistance they receive when it comes to technical and complex matters. Another key is insulating them from influences. Panelists are dispersed (they participate from where they live), don't know each other, and remain unidentified. All communication among panelists is in writing (text or audio input) and platform-mediated so as to prevent private interactions that could lead to the formation of alliances, political parties, personal relationships, and entrainment,[6] with consequent distortion of the governance process.

Text-only communication creates a record that panelists can study, review, and respond to at their convenience. For instance, if panelists have a job or pressing personal obligations, they can attend to their governance-related work as it fits into their agreed-upon personal schedules.

Cyber panelists function separately. Cyberconference panelists do not engage with each other directly, and they have no access to information about anyone's identity. They are not aware of each others' associations, affiliations, interests, history, face-to-face contact, or opportunity to collude. If they were informed of or exposed to such influences, their objectivity could become compromised and subject to entrainment. Cyber panels must be free, too, of temporal discounting, which can interfere with responding appropriately to distant threats or with engaging in long-range planning.

Advantages of governance by direct democracy. Direct democracy by cyberconferencing places decision making firmly in the hands of the populace. There, the society's long-term interests can be balanced against the near-term interests of any adversarial power nodes, without interference. Societal interests include long- and short-term ones that are subjectively defined but quantified "quality of life" features, economic interests, safety concerns, security, and basic rights (property rights,

[6] *Entrainment* refers to influences that members of groups can have on each other. In face-to-face situations they can cajole, pressure, bully, imitate, or follow the lead of others, thereby interfering with independent thinking. Entrainment is seen in rallies and mob behavior, as well as in herds, flocks of birds, schools of fish, and swarms of insects.

freedom), and so forth. Populace panels can balance these interests more objectively than uninsulated and unsequestered governance bodies ever could. For instance, they can control and restrict activities that endanger the environment, health, or public safety, or that create strife, including disinformation activities that distort public understanding of the real world.

The architecture of self-governance

The cyberconferencing system of digital platforms supports the governance functionalities of the populace panels. The platforms are complemented by advisory agencies (perhaps hundreds) whose rotating staffs are vetted, appointed, and supervised by the populace panels.

The advisory agencies are the society's repository of knowledge and expertise in the areas that make a society function, like energy, agriculture, health, defense, transportation, education, immigration, monetary policy, justice, environment, and so forth. They are also an important source of petitions (see below).

There are advisory agencies that perform physical functions, like air traffic control, coast guarding, mining, wildlife preservation, and infrastructure maintenance, in addition to advisory functions. Advisory agencies function under the supervision of populace panels, thus they are prevented from developing direct relationships with power nodes.

The Petition Platform

Petitions are demands for specific governance actions, and they function as the drivers of the governance agenda. Petitions may demand the enactment or repeal of a law, the performance of an act, a response to a complaint, or the resolution of an issue. All of these are collectively referred to as "Actions." Emergency petitions consist of alerts or warnings of identified dangers or attacks for which emergency Actions are required.

Petitions can be originated by citizen groups, advisory agencies, special interest groups, scientific, educational, cultural, industrial, international,

and other organizations, or by power nodes. Power nodes are, after all, impacted by Actions and therefore entitled to have their legitimate interests considered, like anyone else. No segment of a society's population is excluded from submitting petitions, or from appointing advocates to represent their interests in the ensuing cyberconferences.

Impactees: entities affected by Actions. Actions, laws, or decisions have different impacts on different members of the society. Every matter a cyberpanel considers will have a unique profile of *impactees***,** regardless of whether the impacts are positive or negative, beneficial or harmful. A system of governance by the people and for the people must consider any governance Action's impact on the diversity of interests that prevail in a diverse society, including the society's power nodes.

Some impactees of a potential Action(s) may not become apparent until the panel has considered the petition's ramifications. The identification and appointment of the potential impactees' advocates is therefore an ongoing iterative process: as the range of impactees becomes clearer, so does the list of required advocates.

The Petition Platform ascertains that every petition meets the criteria for acceptance,[7] and refers qualified petitions to a cyberconference dedicated to resolving them. For every petition it launches, it makes an initial estimate regarding the required number of panelists, a number subject to revision as the petition's implications become clearer during deliberations.

The Database Platform

This platform maintains and manages two databases. One database identifies and maintains the society's enfranchised populace, referred to as *populace panelists* to distinguish them from other panelists. To fill the panels requested by the Petition Platform, the Database Platform draws from a standby pool of candidates who have received several hours of

[7] Examples of criteria: having a sufficient number of legitimate endorsers, being non-duplicative of other petitions, being economically feasible, being clearly stated, and passing other gross screening measures so as to avoid wasting a cyberpanel's time.

generic panelist training on what will be expected of them and the rudiments of the cyberconferencing discourse language that will be used. This process also screens out panelists whose behavior makes them unsuitable for cyberconferencing duty.

The standing Corps of Experts ("CORE"). The Database Platform maintains a second database from which it summons experts who help populace panelists understand the issues before them. The Database Platform checks and registers the credentials and qualifications of every CORE member.[8] CORE members participate actively on the panels but do not have decision-making power. Their input is crucial to the viability of the model, as populace panelists typically lack the background knowledge to understand the full scope of the matters and issues that may arise.

The CHAIR Platform

CHAIR (acronym for Cyberconference Hosting and Algorithmic Issue Resolution) manages and moderates cyberconferences with the goal of resolving petitions. Once a panel has been constituted, CHAIR verifies the presence of the specified number of panelists, panelists from the advisory agencies, and potential impactees. While petitioners appoint their own advocates, CHAIR appoints the advocates for potential impactees who might not be aware of the cyberconference that affects them.

CHAIR then verifies that every panelist has access to the requisite cyberconferencing software and hardware and agrees to be flexibly available for the duration of the panel. Every message that CHAIR directs to a panelist includes a proposed deadline for their response. If the timeframe is insufficient, the panelist can request an extension. Since all

[8] The U.S. has c. 23 million individuals who might be considered qualified experts or specialists in an area. Based on various sources, there are c. 7 million scientists and engineers, 5 million registered nurses, 1 million MDs, 1 million lawyers, 1 million experts in accounting and finance, 2 million real estate specialists, 2 million farming experts, and another *c.* 4 million experts who do not fall into any of the above categories, for a total of *c.* 23 million experts certified by their circumscribed communities. The CORE registry might include 10-20 million of these. REFERENCES.

communications are in writing (keyed or dictated), there are no meeting times, only *response times* (deadlines).

CHAIR customizes itself for each new panel. Upon taking over a cyberconference and panel, and having completed the administrative mechanics described above, CHAIR studies and "learns" the information contained in the petition, including the Action-related information and the new concepts and vocabulary. This primes CHAIR to recognize and manage the interactions and behavioral dynamics that could occur in deliberations among populace panelists, CORE experts, impactees, advisory agencies, and the various advocates. As part of this "learning" process, CHAIR incorporates new concepts and vocabulary that include:

- Information regarding the petition
- information regarding the potential impactees
- information pertaining to the potential related advisory agency(ies)
- the coded identities and specialties of relevant CORE members
- the coded identities of the advocates
- relevant current and historic context for each of the above items.

As deliberations proceed, CHAIR increases its repertoire of what is specific to the current petition and how it differs from prior ones. It learns from the content of the forthcoming proceedings and from the contributions of experts.[9] CHAIR's main learning tool might be a set of generic frameworks with virtual blanks into which novel information can be inserted. *A vocabulary example*: Communications are called missives; *an interaction protocol example*: CHAIR acknowledges and provides feedback for missives it receives. Every missive sent by panelists is tagged with its addressee(s), subject, and the nature of its relevance (e.g., Another approach to . . . ; Disagreement/Agreement with . . . ; Question about . . .; etc.). Interactions involving CORE members, advisory agencies, impactees, and advocates are subject to their own sets of generic protocols.

[9] Rudimentary versions of these capabilities already exist for language translation, search engines, dictation software on phones, editing functions of word processors, and virtual assistants like Siri and Alexa. Emerging AI techniques like ChatGPT are moving these capabilities to new levels.

The Cyberconferencing Discourse Language (CDL). CHAIR's functionalities require a vast repertoire of generic *discourse units* and the contexts in which those units occur. CHAIR builds that repertoire over time by continuous learning, often with the use of AI technology. Here are two examples of rudimentary Contextualized Units of Discourse (CUD), a large repertoire of which enables CHAIR to carry on quasi-human context-appropriate discourse:

> ***A*** asks a question, ***B*** answers it, and ***A*** asks a follow-up question regarding ***B***'s answer. *Possible partial context*: ***A*** has asked the same or a similar question prior to that interchange.
>
> ***A*** comments on an assertion or tact by ***B*** and states a reason for the comment. ***B*** responds to ***A***'s comment. *Possible partial context*: ***A*** and ***B*** have or have never interacted prior to that interchange.

Like any language, CDL has its own syntax and vocabulary. The above examples are verb-type CUDs. The language also includes noun-type CUDs such as: question, response, missive, agreement, disagreement, consensus, rating, assistance, expert, advisory agency, names of the platforms, proposal, petition, petitioner, panelist, impactee, advocate, and so forth, and combinations of these. There are also recurring contextualized intraverbals and autoclitic prompts similar to those of present day email language: from:, to:, re:, and various available action steps. Equally important in the definition of CDL is the exclusion of most foreign vocabularies and syntaxes, including English.

To adapt its vast CUD repertoire to each new panel, CHAIR identifies, recognizes, and categorizes CUDs that are novel and specific to the new panel, and adds them to its existing compendium of concepts and vocabulary. Examples of contexts are recent circumstances in which the CUDs occurred and pending issues.[10]

[10] CHAIR's moderator functionalities include voice recognition; tracking, organizing, and archiving verbal discourses; routing communications to their destinations; assisting panelists with language issues; maintaining panelists' active participation and goal awareness; making sure expert assistance is continuously available to the populace panelists; identifying points of disagreement; ushering panels toward consensus and agreement. Many functionalities would

CHAIR's moderation of behavioral interactions. CHAIR's overarching function is to usher populace panelists toward a possible consensus. It asks each panelist to state any qualms they may have after every set of discourses. For each stated qualm, CHAIR provides relevant data or arguments that have previously been made, and then asks the panelist for a new rating. CHAIR recognizes and manages personality-driven behaviors, such as if a panelist digs in their heels or becomes irrational.

The design of CHAIR's specialized CDL and its protocols in moderating provide an exciting challenge to linguists and behavioral scientists. A language of this importance may become the gateway to a general science of verbal discourse given the inevitable spread of cyberconferencing.

Disposal of petitions*.* The processing of petitions by CHAIR includes their final disposal. When a petitioned issue has been finally resolved as an implementation-ready Action, CHAIR forwards it to the Implementation Platform, which selects the most appropriate effectors and executors of the decision or Action, much as a corporation or traditional governmental agency would. Enforcement of Actions remains the province of local court officers and militias when required.

When the implementation process involves procurement, the process could require contracting with public or private entities with the desired expertise. To prevent this process from offering any entry ports for corruption with respect to choice of contractor, quality control, or pricing, the populace panel, aided by its CORE experts, selects the contractor and supervises the other details.

The Economy Platform

The Economy Platform serves to maximize employment, GDP, and price stability—goals similar to those of the U.S. Federal Reserve. Its functions include budgeting, taxation, and monetary policy. Its main tool for performing its functions is an AI-based Modeled Simulation of the

require the extension of the science of verbal behavior into the territory of contextualized verbal discourses and conversational units.

Economy (MSE). MSE is also a primary resource for petitioners, panelists, CORE experts, advisory agencies, impactees, and advocates. They regularly request MSE to conduct simulation tests of potential Actions so as to determine their possible effects.[11] The MSE also alerts the Economy Platform to likely consequences of existing Actions or policies, which it reveals in the course of its continuous research and testing. This iterative process improves MSE's capabilities for predicting the impacts of contemplated Actions. It identifies possibly consequential interactions among Actions and proposes alternative ways to address them.

The LIVE Encyclopedia Platform

The LIVE Platform combines the functionalities of Google, Wikipedia, National Archives, Library of Congress, scientific and technical journals, and other digitized sources of information, with search engines that make them accessible. It includes national and international databases in myriad fields of human endeavor and archives. These resources are updated and reorganized by AI methodologies, assisted by epistemologists and behavioral scientists who study the architecture of human knowledge.

LIVE identifies and categorizes trillions of possible relationships among the archived elements—similarities and differences, attributes and functionalities, temporal, logical, and epistemological status, and so forth. LIVE's ever-expanding resources are translated into all languages, and accessible via search queries to its archives.

The self-governance decision-making process

The deliberation process, initiated by a petition and carried out through cyberconferencing, must ultimately generate a decision that reflects the degree of approval or disapproval of the requested Action. To arrive at an

[11] AI methodology will improve this direct democracy model mainly by performing virtual "experiments" that test hypothetical scenarios to determine the effects of potential variables and their interactions. Over time, as empirical economic data are fed into MSE, a further research methodology will be enabled, one in which MSE tests its predictions against actual collected data and learns from those.

Approval Index, CHAIR nets two measures against each other: the long-term economic cost-benefit calculations for the entire society (via the Economic Platform's MSE; see p. 16) and the narrower, subjectively estimated impact on the impactees, including on their quality of life.[12] Since there are usually a number of impactees, the panel needs the monetized value of the total impact on all impactees.

> **Example:** The Petition Platform has received and approved a petition from the Sturdy Bridge Corporation. It proposes to build a bridge over the river that separates two parts of a town. The petition has a long list of co-sponsors and includes engineering and environmental impact studies. To decide whether Sturdy's proposal should go forward, the populace panelists must arrive at an Approval Index for it.
>
> Determination of the Index starts with the long-term cost-benefit calculation for the society as a whole. CHAIR prompts the Infrastructure Agency and the Economics Platform and its MSE to estimate the total economic benefits that could accrue over time from the existence of the bridge.
>
> The economic benefits are balanced against the estimated negative impacts on certain parties, in this case, the residents of the two parts of the town; a ferry company whose business would be depleted by the existence of the bridge; the owners of an apple orchard near the river; relocation costs of homeowners; and so forth. Each impactee's advocate submits a demand for compensation that reflects objective valuations and quality-of-life impacts.[13] The demands are considered and adjusted according to what the panel's consensus considers fair. (The process for achieving consensus and agreement is described in the section below.)
>
> The populace panelists net these two estimates against each other to determine whether the Sturdy Bridge proposal should go forward.

---

[12] Differentiation between these two measures is relevant in governance to addressing distant threats that require near-term actions that a populace might be tempted to defer or minimize because of their cost or aversiveness.

[13] A 1990 documentary by Jordan Mechner relates the case of Chavez Ravine Village, a vibrant California neighborhood that was destroyed by construction of a sports stadium.

Managed cyberconferencing enables decision making and valuation steps to be free of power node influences. It is applicable to any Action with a quantitative economic metric and a qualitative subjectively assessed one. Most governance Actions possess both of these measures.

Quantifying a panel's judgment. Throughout any process in which an Action with identifiable impactees is being considered, CHAIR maintains and ushers the panelists to the goal of securing consensus.[14] CHAIR polls each panelist after every interaction cycle with respect to the weight they would assign to each impactee or impactee group. In the Sturdy Bridge example, CHAIR asks panelists to assign a weight to each impactee group's per-member impact on a scale from +10 (100% for it) to -10 (disregard the impactee), with zero signifying neutrality or lacking information. CHAIR multiplies the average weight assigned by all the panelists by the number of impactees in the group to obtain the consensus index for the panel as a whole. CHAIR terminates the polling routine and considers consensus to have been reached when the consensus index has stopped changing.

There might be instances of Actions in which the entire human species is determined to be an impactee of an Action. In such a case, the impacts would fall into the "entire society" category, distinct from the impacts on individual impactees.

Consensus versus agreement. The existence of a consensus does not mean panelists agree on the substance of a particular issue. It means only that they agree (or are deadlocked) on the status of their process. In an extreme case, if some panelists provide a +10 rating and the others a -10 rating, that would describe perfect consensus on process and maximum disagreement on substance. Consensus and agreement must therefore be measured separately.

[14] In the present context, a useful mathematical definition of consensus is *information* (the opposite of entropy), used here as a kind of Theil index. (REFs: *Entropy Based Approach to Measuring Consensus in Group Decision-Making Problems* by J. M. Tapia, F. Chiclana, M. J. del Moral & E. Herrera–Viedma; Shannon, C. (1950). A Mathematical Theory of Communication; Zadeh, L. (1965). Fuzzy sets and fuzzy logic.)

Agreement on substance can be defined mathematically by a measure of central tendency, such as the standard deviation of a distribution: the narrower the standard deviation, the stronger the agreement. If disagreement on substance is high but consensus has been reached, CHAIR convenes a new cyberconference, this one focused on the area of disagreement.

> **Example.** There is a +10/-10 disagreement about which method to use to build levees to prevent rising sea levels from inundating a coastal area. Panelists reach a consensus that one of two proposed methods should be used (re: process), but disagree as to which one (re: substance). To resolve that issue, CHAIR requests a new cyberconference in which panelists (assisted by CORE experts) reach consensus *and* agreement that the real issue (substance) is whether it makes sense in the long run to abandon the affected coastal areas to the seas rather than build levees. This Action, they agree, would have a severe impact on many impactees. To resolve this last issue, CHAIR requests a third cyberconference, this time with the various impactees and their advocates plus CORE experts in earth science, oceanography, climate change, and local real estate.

This example shows how reaching agreement on process (reaching consensus) and separating that from reaching agreement on the substance can resolve complex and refractory problems, especially when they involve many impactees. Traditional voting systems, by comparison, cannot provide as valid an assessment. Voting systems generate internecine factions and do not require reasoned judgment, understanding of the issues, or even participation. The present model, in contrast, helps the populace's decision makers base their judgments on relevant information and reasoned opinions. Technical assistance is provided, when needed, by MSE, CORE, and advisory agencies.

Justice and fairness

The much-sought-after fairness of traditional codified justice is elusive, as parties and situations differ (e.g., adult versus child, motive, and

"extenuating circumstances."[15] Given the diversity of parties and situations, any given law cannot be both uniformly applied and also fair. Thus, a society's administrators of justice—judges, jurors, legislators—*cannot be held accountable*. Combine that with the possible participation and intervention of power nodes in the traditional administration of justice, and the elusiveness of fairness becomes irremediable.

Justice in self-governance and direct democracy. Direct democracy replaces the traditional judge and jury with a randomly selected cyberpanel of 20 to 200 randomly selected populace panelists. Other panelists are the defendant(s) and plaintiff(s) and their advocates, and legal and subject matter experts. The primary focus is on the facts of the case—their completeness, accuracy, and any presence of distortion or disinformation—and, when applicable, the credibility of witnesses.

Since all communication is in writing, emotions and interpersonal factors are largely excluded, except in the rare instances in which they need to be revealed to the panelists and would be provided on demand. The archives consisting of written records are highly and instantly accessible, unlike traditional records.

The concept of justice is replaced by a quantitative *culpability metric*: the defendant's degree of culpability of the alleged tort or crime, and its severity or nature. Populace panels reach consensus or agreement for each of these metrics separately. Since they are quantitative, they can be converted into a verdict according to a legal code accepted by the society. This approach encompasses and supersedes the concepts of reasonable doubt or unfair representation.

[15] Not considered here is the additional issue of how societies can reconcile the need for responsibility and punishment with the scientific fact that socially undesirable behavior is often occasioned by conditions that the society itself created. REFs

Overview of subsequent chapters

Chapter 1 of this work sketches out a platform-managed technology of direct democracy and self-governance for a human society. A more complete bibliography and additional discussion of the topics covered will follow when the document is completed.

Regardless of whether the model would be capable of delivering the desired results to a future society, its component elements might already deliver benefits in the near term, as described in Chapter 2 of this work, an overview of which follows.

Overview of Chapter 2: The prospect of a new industry

This chapter describes how the direct democracy and self-governance project's development process may create a byproduct: a business able to fund the overall model's development costs. As the world shifts away from in-person business, commerce, and interactions in general, and more toward virtual interactions and the movement of information, a cyberconferencing industry is likely to emerge. The immediate economic benefits of managed cyberconferencing, as described in Chapter 1, include the elimination of the need to travel and its associated costs and disruption of lives. It also brings the logistical benefits of being able to assemble large widely dispersed groups of conferees on short notice.

Some of CHAIR's software components overlap with those of applications in contemporary geographically dispersed groups and organizations. Developing CHAIR's components would therefore advance the capabilities of itself and of many diverse applications:[16]

[16] CHAIR and its various functionalities would be sold as services to organizations, international corporations, educational institutions, dispersed professional or scientific associations, town meetings, communities, and so forth. Such near-term applications, especially if they prove valuable to their users, may justify further investment in the development of CHAIR, even if self-governance and direct democracy never materialize.

- Operating open, organization-wide petitioning systems
- reading cyber panelists' written messages and sending them to their intended destinations within the conference.
- Polling groups on controversial issues, with voting replaced by the tracking of consensus.
- Conducting referendums on contentious issues for widely dispersed populations or cultures.
- Cyberconferencing for the deaf.
- Conducting cyberconferences for widely dispersed
    - multilingual panelists.
    - Parties that need to address emergencies or other time-sensitive matters.
    - international scientific, technical, or other types of conferences.
    - classes and lectures for student populations by educational institutions.
    - parties that need to resolve a conflict or settle a dispute.
    - international sales and marketing meetings of international businesses that have worldwide sales forces.
    - villages, townships, and cities for decision making and issue resolution.
    - international environmental conferences.
    - tracking, facilitating, and accelerating consensus convergence in cyberconferencing.

The potential clients for these services are domestic and international organizations, governmental and educational institutions, professional and scientific associations, and political organizations that want to conduct conferences for geographically dispersed or linguistically diverse populations. These potential clients can be approached during the

development phase for possible test sites. Given the breadth of these applications and their clienteles, one can visualize the emergence of a worldwide cyberconferencing industry within the next decade.

Overview of Chapter 3: The ongoing project

This chapter outlines the launching phases of the model and its technology described in Chapters 1 and 2. Development work is being conducted and funded by the Mechner Foundation with the initial goal of building the team that can advance the platform development work to the point where field testing—the project's first major milestone, described in Chapter 2—can be initiated. The initial software modules to be developed will be ones that permit productive field tests to be conducted. These are the envisioned steps:

1. Selection and prioritization of the **functionalities to be programmed.** Conducted by the author in collaboration with 2 to 3 behavioral scientists.
2. Assembly of a team of **behavioral scientists** who will perform **task analyses** for the software development.
3. Assembly of a team of **software developers** who will **develop and manage** the development of the platforms' software programs.
4. Lining up developmental **test sites** for cycles of testing, debugging, and revision.
5. Field testing and demonstration of functional systems.

Overview of Chapter 4: Philosophical implications

This chapter discusses the project's theoretical and philosophical implications, and the breadth and depth of behavioral science's potential for contributing to human affairs.

Is world self-governance a possibility? This question usually, and appropriately, leads to a dead end because of the specter of a malign and

predatory world autocracy. But the possibility of self-governance without a power node at the top might eliminate that specter. If platform-managed cyberconferencing as described in Chapter 1 turns out to enable self-governance of very large groups, such as nation states, might a path to world self-governance and direct democracy come into view? The specific question is whether platform-managed cyberconferencing, AI, and participation of the behavioral sciences can produce self-governing nations and, by extension, breathe new life into conversations about the feasibility of world governance without a power node at the top.

Providing for the worst-case scenario. We don't yet know whether the currently looming threats can be averted. If we were to conclude, one day, that we must anticipate catastrophes so devastating as to leave behind mere enclaves of human survivors, today's scientists would still be duty-bound to create virtual time capsules with potentially helpful roadmaps and technology for use in future go-arounds.

Key features of platform-managed self-governance:

The benefit of self-governance to a society or to an organism is that it bars self-benefitting entities, such as power nodes, from interfering in the society's governance processes. In a self-governed society, governance decisions are made by panels of randomly selected members of the populace who govern with assistance from experts as needed as they decide on actions, laws, and policies directly, without intermediaries, leaders, or chiefs, the feature termed *direct democracy.* The governance system is managed by a network of digital platforms including the Petition Platform, the Database Platform, the CHAIR Platform, the Economic Platform, the Implementation Platform, the LIVE Encyclopedia Platform, and others.

- The society's governance agenda is driven by petitions from citizen groups, advisory agencies, and the society's various special interest groups.
- The Database Platform recruits panelists randomly from a database of the entire populace. Panelists confer remotely from their dispersed locations, thereby eliminating the cost and disruption of travel.
- Petitioners or impacted parties (impactees) have advocates on the governance panels to represent their interests.
- Panelists communicate in writing only, via keyboard entry or voice transcription, without video or voice transmission.