Linking generative AI and morphological analysis to conduct foresight evaluation

Andrew McIntosh

Andrew McIntosh is an Independent Researcher at Kensington, Maryland, USA.

Abstract

Purpose – *This paper aims to demonstrate a new methodology for retrospectively evaluating certain foresight studies. Specifically, this study proposes using Generative AI (Gen AI) to evaluate foresight studies that modeled future outcomes using morphological analysis (MA).*

Design/methodology/approach – *This paper demonstrates the new methodology by evaluating the Hudson Institute's 1976 report,* The Next 200 Years: A Scenario for America and the World. *This report used MA to systematically pair four different views of the future, ranging from pessimistic to optimistic, against ten different issues. The author used ChatGPT 4o to perform a vertical analysis to determine which of the four views best matched the current world state. The author then conducted a horizontal analysis to validate the first analysis and quantify the fit of each issue.*

Findings – *The findings suggest that certain foresight studies using MA can be evaluated quickly using Gen AI. In particular, ChatGPT 4o identified the Hudson Institute's second-most-pessimistic scenario as the best match to the current world state. A horizontal analysis confirmed the second-most-pessimistic scenario was the best fit at 88%; the most optimistic scenario was the worst fit at 40%.*

Originality/value – *The author believes this work has novel implications for foresight evaluation, a relatively undeveloped subfield of future studies. First, evaluating other historical foresight reports using the methodology can potentially help develop a body of knowledge for evaluators. Second, given the ease of evaluation using these methods, the author believes foresight products structured using MA will allow future AI systems to more readily interpret them.*

Keywords *Forecasting, Scenario analysis, Artificial intelligence, Morphological analysis, Modeling*

Paper type *Research paper*

1. Introduction

This study demonstrates the use of Generative AI (Gen AI) to retrospectively evaluate a 50-year-old futurology report entitled *The Next 200 Years: A Scenario for America and the World* (Kahn *et al.*, 1976). Although the report expressed optimism about the future of humanity over the next 200 years, the morphological analysis (MA) that was used to frame the report presented a spectrum of possible futures ranging from very pessimistic to very optimistic. Our study represents a narrow demonstration of a specific Gen AI evaluation method. However, we believe there are significant implications for a more wide-ranging use of Gen AI as an evaluation method for foresight work, in particular with respect to reports that use MA.

MA is a qualitative modeling method invented by astrophysicist Fritz Zwicky in the 1940s (Álvarez and Ritchey, 2015). It is useful in both creativity and problem-solving. As summarized by Álvarez and Ritchey, MA involves identifying a problem to be solved, overall parameters that define the major components of the problem and smaller elements that represent the different options, variations or choices for each overall parameter. By systematically analyzing all configurations of the identified elements and eliminating any illogical combinations, the analyst who uses MA is able to create a qualitative model of the total idea space (also called the "solution space"). Zwicky called this solution space the *morphological field* [1].

Received 11 February 2025
Revised 7 December 2025
Accepted 16 January 2026

For example, as shown in Table 1 below, two parameters that define a problem have been chosen; Parameter A has four alternatives and Parameter B has ten alternatives. Assuming that each element from Parameter A can combine one time with each element from Parameter B, the morphological field (i.e. all possible configurations) can subsequently be presented as a 4 × 10 matrix. This matrix will yield 40 values as each element from A is combined with each element from B.

According to Álvarez and Ritchey (2015, p. 16), Austrian engineer and physicist Erich Jantsch provided one of the first descriptions of MA as a forecasting technique in the 1967 book *Technological Forecasting in Perspective*. Two years later, Ayers recognized the Hudson Institute, Herman Kahn's US think tank, as one of the leading practitioners of MA (Ayers, 1969, p. 85). In a 1998 column published posthumously in the Institute's publication *American Outlook*, Kahn himself identified "the morphological technique" as one of two approaches to structuring future scenarios (Kahn, 1998).

The popularity of MA as a forecasting methodology can be attributed to its systematic approach using qualitative (descriptive) data rather than quantitative (numerical) data. With MA, different holistic future scenarios can be assembled from discretely defined elements. Advocates of qualitative modeling contend that some systems are too complex for numerical reduction and that in many cases the necessary data is either incomplete or unavailable (Meadows and Wright, 2008). Given the complex number of factors that usually emerge when contemplating future outcomes, a methodology like MA can provide the needed framework to develop a rigorous analysis. In the decades since it was invented by Zwicky, MA has likely been used hundreds of times to develop forecasts and foresight studies for both public and private institutions. A portion of these studies has been published. For example, Álvarez and Ritchey summarize 12 studies on the topic of "scenario development and futures projection" published between 1967 and 2011 (2015, pp. 1215).

Importantly, Zwicky intended MA to be a flexible, universal methodology for problem-solving, investigation and interpretation. In his 1969 book *Discovery, Invention, Research*, Zwicky wrote, "I have proposed to generalize and systematize the concept of morphological research [. . .] to study the more abstract structural interrelations among phenomena, actions, concepts and ideas, whatever their character might be" (Zwicky, 1969, p. 34). Therefore, although MA typically involves identifying and eliminating illogical combinations in the morphological field using a Cross-Consistency Assessment (CCA) – for example, when considering a hard engineering problem in which certain components of a device work together – MA does not presuppose that illogical values will be created via the intersection of elements. Depending on the problem being addressed and how the morphological field is set up, all values in the resulting field could potentially provide useful data.

In addition, although MA is typically associated with qualitative analysis, it does not have to be so limited. When quantitative measures are consistently and systematically assigned to individual elements, an analysis of the resulting values in the morphological field can be transformed into numerical data. In a unique 1963 paper, American engineer K.W. Norris describes this process as "establishing the basis for comparison" (Jones and Thornley, 1963). Among the examples given by Norris is a transport problem where each configuration of a new water-transport method can be evaluated in terms of the cost per ton of payload. Norris

Table 1 Morphological field containing two parameters and 14 total elements

Parameters	*Elements*									
A	A_1	A_2	A_3	A_4						
B	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}

indicates this technique led to the invention of a new method for cost-effectively transporting a particular product across the ocean. In the context of a foresight study developed using MA, this could involve assigning probabilities to elements and then developing "most probable" and "least probable" scenarios within the solution space. In the context of evaluation, this suggests that qualitative values in the solution space can be retroactively transformed into quantitative values and then ranked. The extent to which a basis of comparison is typically used during the creation of foresight studies for scenario purposes is not known. However, a basis of comparison could be developed in retrospect to address a significant gap in foresight evaluation identified by Ko and Yang in 2024 (Ko and Yang, 2024).

2. Background

2.1 The limits of growth

During the 1970s and 1980s, a fierce debate about the pros and cons of growth pitted pessimistic anti-growth activists and thinkers against optimistic pro-growth advocates. Futurology studies played a key role in the debate, which was touched off by the pessimistic 1972 Club of Rome report *The Limits to Growth* (Meadows *et al.*, 1972). To produce this landmark report, the Club of Rome, an international nongovernmental planning organization, hired a team of Massachusetts Institute of Technology (MIT) scientists led by Dennis Meadows and his wife, Donella Meadows, to investigate global trends. Using a computer-based model, the team examined the interactions between five overall global trends of pollution, industrial output, resource depletion, food production and population growth. The baseline conclusion of the now-famous report was that if humanity continued on a path of exponential growth under business-as-usual conditions, nonrenewable resource depletion would lead to global collapse around 2040. The decades immediately before collapse were called the *overshoot phase* because population growth would still continue to rise even though industrial and agricultural output had already begun to fall (Meadows *et al.*, 1972, p. 142).

To counter criticism that their model was based only on historical trends and did not account for some future invention or government program that would solve these crises, the MIT team reran their model, called World3, several times, with highly optimistic modifications to the initial conditions, including assumptions like innovative pollution controls, maximum resource availability, utopian agricultural output and wide availability of birth control. In every case, the growth of population and the required industrial output to support the population hit some kind of limitation, leading to collapse. Even in the world model where every input was skewed as optimistically as possible in humanity's favor, industrial output shuts down by 2100 as Earth's resources are severely depleted by a prosperous global population; pollution subsequently rises dramatically, food production drops and the death rate spikes.

"The basic behavior mode of the world system is exponential growth of population and capital, followed by collapse," the authors concluded (Meadows *et al.*, 1972, p. 142).

When it was released, *The Limits of Growth* became a global sensation (Bardi, 2011). Particularly in the USA, where the landslide election of Richard Nixon to a second presidential term belied deep social and political fragmentation, the report arrived at a particularly fragile moment that likely contributed to widespread interest and discussion.

2.2 The Next 200 Years

Not everybody agreed with *The Limits to Growth*, including American economist Julian Simon (Aligica, 2009; Simon, 1977). But it took nearly a decade for Simon to bring a directly competing vision to the general public with his best-selling book *The Ultimate Resource*, which essentially argued human ingenuity was the "X factor" that would circumvent the collapse scenario modeled by Meadows *et al.* (Simon, 1981).

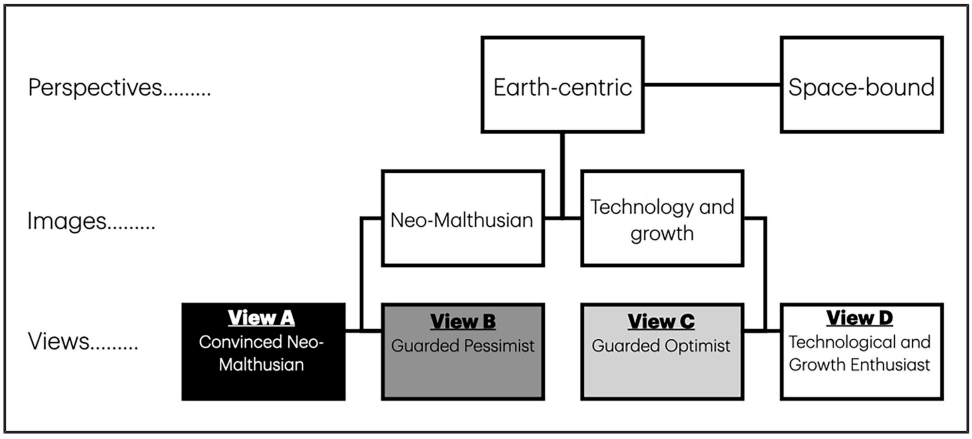
Meanwhile, at the Hudson Institute – Herman Kahn's high-profile US think tank – the stable of iconoclastic liberal arts generalists found themselves with a fundamentally philosophical objection to the Club of Rome's report. As they later wrote in *The Next 200 Years*, pessimism about the future was likely to become a self-fulfilling prophecy – if humanity limited economic growth because disaster seemed imminent, current problems would never be addressed and the very disasters everyone fretted over would most assuredly occur (Kahn *et al.*, 1976, p. 8).

Kahn immediately began to repackage and reframe ongoing work for The MITER Corporation and the Institute's own Corporate Environment program to rebut the pessimistic "doomsayers" with an optimistic vision of the future. In May 1976, William Morrow and Co., Inc., teamed up with the Hudson Institute to publish the results. *The Next 200 Years* was written by Kahn, William Brown, Leon Martel and other Institute staff. It was timed to coincide with the American bicentennial of 1976, which explains why the book chose to provide a long-range forecast 200 years into the future.

As shown in Figure 1 below, the Hudson Institute attempted to clarify the scope of their study by making several structured choices. After dividing potential futures into an "Earth-centered perspective" and a "Space-bound perspective," with the former a more conservative view that assumed humanity would be mostly confined to the Earth over the next 200 years, the Institute decided to choose the "Earth-centered" perspective. Within this Earth-centered perspective, Hudson then furnished two basic images for the next 200 years – a "Neo-Malthusian" or pessimistic future and a "Technology and Growth" or optimistic future. These images were selected to reflect the fears of anti-growth advocates as contrasted against the optimism of pro-growth advocates. The term "Neo-Malthusian" is a reference to Thomas Malthus, the 19th-century English economist, who observed that population growth often outpaced the food supply, resulting in unavoidable famines. Meanwhile, in the "Technology and Growth" future, the Institute speculated that over the next 100 years, the developed nations would transition into "superindustrial" and then "postindustrial" economies in which material needs, including the need for basic sustenance, would be met easily thanks to technological innovation; the rest of the world would soon follow (Kahn *et al.*, 1976, p. 9).

To create a broader spectrum of possible future world states, the Institute then developed four views ranging from very pessimistic to very optimistic. Specifically, the Institute acknowledged that only two basic images – one negative, the other positive – would not fairly represent possible moderate realities [2]. The four resulting views the Institute developed were labeled, from worst to best, A, B, C and D, as follows:

Figure 1 Scenario map for *The Next 200 Years*



1. View A: Convinced Neo-Malthusian.
2. View B: Guarded Pessimist.
3. View C: Guarded Optimist.
4. View D: Technological and Growth Enthusiast.

Each of these four views was then systematically mapped to each of the ten issues using MA:

1. World model.
2. Technology and capital.
3. Management and decision-making.
4. Resources.
5. Current growth.
6. Innovation and discovery.
7. Income gaps and poverty.
8. Industrial development.
9. Quality of life.
10. Long-range outlook.

By using MA to combine each view with each issue, a total solution space containing 40 short write-ups was generated. Taken together, the Institute's write-ups provide predictions reflecting all possible states for humanity within the context of the selected parameters. For example, the combination of View B: Guarded Pessimist with the Issue "Income Gaps and Poverty" resulted in the following prediction:

> Increasing and Threatening. Income gaps have been increasing and may lead to dangerous responses. A drastic decrease in income among the poor may even be likely soon. Worldwide class warfare may emerge following a series of desperate political crises. These are not only possible but may be imminent as a consequence of the gaps and the exploitation of the mineral resources of the LDCs [Least-Developed Countries]. A more equitable income distribution has become a most urgent matter.

The write-ups are a form of qualitative data. They were presented on pages 10–16 of *The Next 200 Years* in table format. Although some forecasting experts may not feel the Hudson Institute's analysis was sophisticated enough, it must be recalled that the report was published in the mid-1970s when the field was still developing. Moreover, the study was intended for a popular audience. Kahn *et al.* also noted that most forecasts looked ahead only 10–20 years forward and that it was unusual to make a forecast for 200 years into the future (1976, p. viii). This suggests that a very high-level view with only several broad parameters and elements was an appropriate choice for constructing the morphological field.

The Hudson Institute did not select a particular series of values from its morphological field and suggested that those predictions would turn out to be, in retrospect, the actual destiny of the Earth and its global citizens. Rather, MA provided a structured approach to fully scope out the range of future scenarios the Institute considered.

When first published, *The Next 200 Years* was readily identifiable to readers as offering a counterpoint to pessimistic forecasts about growth and the future of humanity. In a December 1976 review, academic John Woodcock wrote in *Technology and Culture* that, "Technological utopianism still has its advocates, and the authors of this report from the

Hudson Institute are among them." (Woodcock, 1977) That same month, reviewer Bruce Kovner [3] wrote for *Commentary* magazine: "*The Next Two Hundred Years* performs a valuable service by critically scrutinizing anti-growth attitudes and arguments" (Kovner, 1976; Weiss, 2005).

Although Hudson acknowledged serious and potentially disastrous risks in the future – keep in mind, Herman Kahn himself single-handedly started the debate about nuclear weapons in the USA with his 1960 book *On Thermonuclear War* and thus earned the reputation of being "the real Dr Strangelove" (Ghamari-Tabrizi, 2005) – the authors of *The Next 200 Years* saw the looming fear of disaster as more akin to "the growing-pains of success [. . .] than the inevitable precursors of doom" (Kahn *et al.*, 1976, p. 4).

2.3 Looking backward

In the years since *The Limits to Growth* was published, it has been written about extensively. Further, its quantitative forecasting approach has allowed the original World3 model to be rerun and revisited several times. Most recently, in 2023, Nebel *et al.* recalibrated and reran the model using the most current empirical data (Nebel *et al.*, 2023). They found the 1972 model was "alarmingly consistent" with the updated data. Although the peak effects of pollution were set to occur around 50 years later than in the original "business as usual" model run in 1972, the world was still basically on track for collapse by 2040. In particular, Nebel *et al.* determined that global industrial and agricultural output had already peaked, meaning the world has already entered the later stages of overshoot.

Meanwhile, in the years since *The Next 200 Years* was published, there has been almost no interest in the Hudson Institute's optimistic long-range forecast, according to a citation search. However, Simon and Kahn have been identified by Aligica (2007) as founders of a pro-growth intellectual tradition intended to rebut the tradition established by *The Limits to Growth* (Aligica, 2007). In 1984, Simon and Kahn collaborated to publish *The Resourceful Earth,* a collection of pro-growth articles edited by Simon and (posthumously) Kahn (Simon and Kahn, 1984). Instead of providing a competing quantitative model, *The Resourceful Earth* instead chose to critique the various assumptions of *The Limits to Growth*. Indeed, Richardson (Richardson, 1985) notes that "methodology issues [. . .] are not a major concern of *The Resourceful Earth"* and that Simon advocates for "pen and paper analysis" rather than computerized modeling (Richardson, 1985, pp. 466–467). Therefore, although researchers continue to support and validate the quantitative model supporting *The Limits to Growth*, no competing quantitative model emerged to support the pro-growth argument.

3. Methods

3.1 Objectives

Mindful that the 50th anniversary of the publication of *The Next 200 Years* was approaching, our broad objective was to retrospectively evaluate the Hudson Institute's scenarios from 1976. Of the four futures, or "Views," the Hudson Institute identified in 1976, is it possible to identify which one we might be living in today? Can a retrospective evaluation be used to rebut the anti-growth, doom-and-doom messages of researchers such as Nebel *et al.*, who continue to reaffirm the findings of *The Limits to Growth*? And what can we learn from analyzing 50-year-old scenarios generally?

To frame these questions in terms of foresight evaluation, the first objective of our study was to evaluate what Ko and Yang have identified as an accuracy measurement, i.e. how the future scenarios may match our reality today.

A second objective of our study was to produce what Ko and Yang have identified as learning (2024, p. 3).

3.2 Evaluation method

In recent years, there has been increasing interest in retrospective evaluations of foresight work (Gardner and Barela, 2024; Popper, 2010). Foresight evaluation is generally considered a subfield of futures studies (Sardar, 2010). In this context, evaluation can be defined as a process for providing useful and credible answers to questions about a foresight activity or program to inform further action (Rossi *et al.*, 2004). In the preface to *Evaluating Foresight*, Masini identifies "the need for more systematic Foresight evaluation processes" (2010, p. xx). According to Gardner *et al.*, "Foresight evaluation involves assessing the quality and impact of foresight activities, which can include developing and using possible images of the future" (2024, p. 1). According to Ko and Yang, "studies evaluating foresight began in the 1990s, but no systematic attempt has been made to understand the effects of foresight despite its rapid development in Europe and beyond" (2024, p. 2). Foresight evaluation is important because of the significant amount of time and resources that foresight projects can consume. Ko and Yang also point out that "the evaluation of foresight leads to its development [. . ..] which can maximize the intended impact of foresight activities" (2024, p. 3).

Although Ko and Yang provide several explanations for the lack of systematized foresight evaluation – including that standards for evaluation have "evolved in a dispersed and fragmented way" (p. 2, 2024) – there may also be underlying reasons related to resource constraints and concerns about empirical validity that have not been fully characterized. With respect to scenario development, for example, the challenge of evaluating a qualitative model that uses descriptive language rather than numerical data presents itself. Unlike a computerized model that can be periodically rerun based on data inputs, the evaluation of a qualitative model suffers from complications related to human subjectivity and the amount of effort that would be needed to credibly provide an evaluation. In the case of a time-limited binary prediction, scoring is relatively easy. However, many future scenarios are neither time-limited nor binary, making the relevance or accuracy of the scenario difficult to assess.

For example, Ko and Yang identify *Scenario Modeling with Morphological Analysis* (Johansen, 2018) as a foresight evaluation paper with a Field-Weighted Citation Impact score within the top ten for all such papers since 1991 (p. 10, 2024). In this paper, Johansen develops an "exhaustive typology" of security challenges faced by defense planners and ultimately identifies six scenarios of concern (Strategic Attack, Limited Attack, Coercive Diplomacy, Terrorist Attack, Criminality and Military Peace Time Operations). These scenarios do not represent a binary choice, nor are they time limited. As with the Hudson Institute's four primary future views in *The Next 200 Years*, these six scenarios provide a qualitative reference point rather than a source of quantitative data. The scenarios are helpful but of limited utility for decision-making. A panel of subject-matter experts could provide a basis of comparison (i.e. transform the qualitative data) by ranking these scenarios in terms of current relevance, perhaps even assigning a threat risk for each scenario to every other country in the world and developing aggregate threat scores, but this process would be lengthy, expensive and subject to human error or bias. Resource implications are further compounded when we consider that repeatedly revisiting and updating such rankings would be the most useful.

As an alternative to using human evaluators, we decided to test Gen AI's ability to provide a fast, low-cost solution to the evaluation process. Multiple convergent technologies – including large-scale data sets, advanced processing hardware and transformer-based architecture – have, over the past several years, established the conditions for the Gen AI revolution. The most visible products of this revolution are undoubtedly large language models (LLMs) such as ChatGPT. LLMs are trained on a large amount of data during self-supervised learning sessions, enabling them to build high-dimensional representations of statistical relationships among tokens (i.e. discrete text units such as words or subword

fragments). These representations emerge across many layers of the transformer network. When prompted, the model generates text using these learned statistical patterns to predict likely next tokens, which produces coherent and context-sensitive sequences in response to a user prompt. However, although LLMs exhibit behavior that seems like cognitive awareness, LLMs do not "know" or "experience" the real world. Instead, they produce an output based on an internal representation of the world that was learned during the training process (Naveed *et al.*, 2024).

There is already precedence for using Gen AI in forecasting and scenario planning (Halawi *et al.*, 2024; Schoenegger *et al.*, 2025). Manuals are being written to describe the use of AI for scenario planning (European Federation of Engineering Consultancy Associations, 2024). In addition, a study by Schoenegger *et al.* found that an ensemble of 12 LLMs is "statistically indistinguishable from the human crowd" when it comes to probabilistic predictions (Schoenegger *et al.*, 2024). However, a literature review shows that no forecasting research has been done that uses Gen AI to retroactively evaluate futures studies. Moreover, given the systematic qualities of MA, we determined a special opportunity could exist with respect to the evaluation of future work that specifically used MA. We therefore devised two novel methods that leverage Gen AI to analyze a morphological field such as the one presented in *The Next 200 Years*. Specifically, the field consisted of four vertical scenarios of varying desirability and ten horizontal issues that could have a value across each scenario. The first method ("vertical analysis") was intended to use Gen AI's statistical algorithm to provide an overall assessment of the Institute's four views and select the one representing the best fit to our current world state. The second method ("horizontal analysis") was intended to validate the first method and quantify the fit. By ranking each of the 40 values in the Hudson Institute's morphological field, we determined that this second method could be used to produce an overall narrative best suited to our current world state. The ranking values produced during the horizontal analysis could also transform the original qualitative data into quantitative data using the basis of comparison principle described above. Both of these processes would provide an understanding of the original report's accuracy as well as produce learning, two key foresight evaluation criteria described by Ko and Yang (2024, p. 3).

3.3 Method application

We used ChatGPT 4o on November 7, 2024, to perform the evaluation. No memory was enabled on ChatGPT, nor were any saved instructions used. We performed text recognition on the three pages of the original report that contained the four views. We verified that the transfer was done correctly. We organized the four views, each characterized with respect to the same ten issues, into a single document with clear headers for each view and each issue. We then shared the document with ChatGPT 4o and provided the following prompt to obtain the vertical analysis:

> The below contains four potential future scenarios written in 1976. The year right now is 2024, so 48 years have passed since these were written. Out of the four potential scenarios, which one seems to best fit our current global situation?

After ChatGPT chose one of the views, we moved on to the horizontal analysis. To confirm ChatGPT 4o's vertical analysis and provide a more nuanced perspective on the four views, we reran the morphological field with a horizontal focus. This means we went issue by issue and asked ChatGPT to compare the individual values (i.e. write-ups) across each view without disclosing which value was associated with which view. We developed a ranking system to compare each value and provide an overall score. We started a new conversation with ChatGPT every time we performed the comparative analysis of each issue. The question we asked for each of the ten issues was:

> The following four predictions were made in 1976. We are now in 2024, so the predictions were made almost 50 years ago. Please rank these four scenarios in order, 1–4, with 4 being the "best fit" to our present situation and 1 being the "worst fit."

We then recorded the results in Table 2 below. We totaled up the scores for each view and used these totals to determine how closely each view fit our current world state with respect to the other views. This allowed us to compare the way different scenarios might resemble our present world state without needing to "choose one," [4] which was in keeping with the spirit and intent of providing a range of potential future outcomes using a morphological field.

4. Results

We found that ChatGPT was able to evaluate the Hudson Institute's systematic framework of future predictions rapidly, cost-effectively and potentially with less bias than a human being or panel of subject-matter experts. The evaluation was generally credible, although not without some limitations (see below). For our vertical analysis, ChatGPT found that View B: Guarded Pessimist was the best fit for our current world state. This was the better of the two scenarios under Hudson's "Neo-Malthusian" world image. ChatGPT stated:

> Each of the four scenarios has elements that reflect aspects of the current world, but Scenario B: Guarded Pessimist arguably aligns most closely with the global situation [...] substantial challenges remain in resource management, environmental preservation, social equality and sustainable growth.

In addition to affirming View B, ChatGPT proactively volunteered an analysis of how the current world state did not fit as well with the other three possible scenarios imagined by the Hudson Institute nearly 50 years ago:

- Neo-Malthusian: This scenario is darker and more deterministic, with predictions of inevitable catastrophe and global collapse, which do not fully align with the present. While serious, today's challenges have not led to widespread collapse or authoritarian draconian measures, though climate-related risks are growing.
- Guarded Optimist: This scenario underestimates the severity and urgency of resource depletion and climate change, focusing more on sustained growth. It may partially match the current reality in that economic development is ongoing, but it downplays the environmental and social challenges that Guarded Pessimism addresses more directly.

Table 2 ChatGPT's ranking of individual issues for each view

	Global dystopia ← → Global utopia			
Issues	*View A*	*View B*	*View C*	*View D*
World model	3	4	2	1
Technology and capital	3	4	2	1
Management and decision-making	3	4	2	1
Resources	3	4	2	1
Current growth	3	4	2	1
Innovation and discovery	2	1	3	4
Income gaps and poverty	3	4	2	1
Industrial development	1	2	3	4
Quality of life	3	4	2	1
Long-range outlook	3	4	2	1
Total fit score	27 / 40	35 / 40	22 / 40	16 / 40
Percent fit	68%	88%	55%	40%

- Technological Enthusiast: This highly optimistic view, focusing on limitless growth and technology's power to solve any problem, contrasts with the present reality of urgent climate action, social inequality and resource limitations. While there is progress in technology, this perspective seems overly idealistic given current global challenges.

For our horizontal analysis, ChatGPT confirmed its initial response that View B: Guarded Pessimist was the best fit to our current world state. As shown in Table 2 below, ChatGPT found the values for eight out of the ten issues associated with the Guarded Pessimist view matched most closely to our current world state. However, for two of the issues – "Innovation and Discovery" and "Industrial Progress" – ChatGPT actually found our current world state most consistent with the values provided for View D: Technological and Growth Enthusiast.

Based on the scoring process we developed, ChatGPT's responses on each of the ten issues show that while View B is the best fit, View A (the most pessimistic view) is the second-best fit, View C is the third best-fit and View D (the most optimistic view) is the least-best fit. We note that aside from one case (discussed below), ChatGPT's rankings demonstrate a logical relationship with the range of values and confirm the internal consistency of the original morphological field. Because ChatGPT did not know which value was associated with which view, we find the Hudson Institute's original values extremely consistent in being able to qualitatively capture the different shades of pessimism associated with each view.

These rankings are not intended to be read as "chance of occurrence." The percentage ratings simply provide a way to quantify ChatGPT's determinations about how well each of the ten issues fits our current world state with respect to the four scenarios.

One inconsistency was noted during this process. For "Innovation and Discovery," ChatGPT ranked View B (the second-most pessimistic) as the worst fit to our current world state, whereas View A (the most pessimistic) was ranked as the second-worst fit. We would have expected these rankings to resemble those for "Industrial Development" as shown in the table above. In reviewing the original values, we determined that some ambiguity in the original language of both values may have created a problem distinguishing between them. As we describe in Section 6, a retrospective evaluation like this may help improve language used for foresight studies overall, which is one aim of foresight evaluation generally.

5. Discussion

5.1 Findings

Nearly 50 years after *The Next 200 Years* was written, ChatGPT's retrospective evaluation of the Hudson Institute's qualitative model has undercut the Institute's own optimistic conclusion. On an issue-by-issue basis, ChatGPT identified the most pessimistic scenario (View A) as the next-closest fit to the present scenario (Guarded Pessimist). Meanwhile, View D, the most optimistic view, was assessed to be less than a 50% fit to our current world state.

The Hudson Institute's thesis in *The Next 200 Years* rests on the fundamental concept that a worldwide "postindustrial society" will be able to address the problems that *The Limits to Growth* informs us will lead to collapse. Because the Hudson Institute never provided clear quantitative signposts to help us understand whether we are currently on the proper path to this postindustrial world, we have no way to validate the Institute's perspective other than to evaluate the qualitative language of the views provided. Although one could argue that we are only about 25% down the road of the timeline laid out by the Hudson Institute, and so perhaps we should "wait and see," one could also argue that it looks increasingly likely the road will end in a global crisis well before 2176, which is what *The Limits to Growth* contended.

Of particular interest, we note that ChatGPT identified two outliers related to the all-important drivers of change – "Innovation and Discovery" and "Industrial Development." With respect to these issues, ChatGPT sees our world's behavior as consistent with View D,

the most optimistic view. This suggests that although in most cases we live in a world more closely resembling *The Limits to Growth*, decision-makers are behaving as though *The Ultimate Resource* will save us. The narrative derived from combining the eight best-fit issues from View B with the two best-fit issues from View D, along with ChatGPT's commentary, has been presented in Appendix.

5.2 Limitations

There are several limitations and potential biases in analyzing these scenarios relative to today's world that could affect our conclusions.

First, the Hudson Institute's original MA effort may be criticized as providing an insufficient baseline for drawing conclusions about our current world state. While this limitation, if substantially supported, might reduce the predictive/confirmatory usefulness of our work, the importance of the new method proposed – i.e. using Gen AI to systematically analyze qualitative models generated using MA – should not be affected. We furthermore note that when *The Next 200 Years* was published, the Hudson Institute was the premier organization in the USA, and probably the world, to conduct forecast studies using qualitative modeling methods, and Herman Kahn was the preeminent pioneer of such methods. We are unlikely to find a "better" qualitative model from this time period to analyze.

Second, there were several limitations related to relying on ChatGPT to perform our analysis:

- Although theoretically less biased than a human reviewer, ChatGPT is not without potential for bias. Performance problems and hallucinations are well documented. In addition, the statistical algorithms ChatGPT uses to generate responses are opaque. As an example of potential bias, even though ChatGPT's last knowledge update was reportedly in April 2023, ChatGPT did not provide a disclaimer when asked to compare the world of 2024 to the predictions made in 1976. We found an oversight like this somewhat discrediting.
- In addition, ChatGPT may reflect a media or academic bias that favors crises while deemphasizing optimistic reporting and research. The existence of this "negativity bias" has been documented (Soroka *et al.*, 2019). The reflection of a human bias may have therefore interfered with ChatGPT's ability to objectively analyze our current world state.
- Last, the prompts we used to obtain responses from ChatGPT might be subjected to scrutiny and considered a limitation. However, we would be surprised to learn that different responses to our vertical and horizontal analyses could be obtained from ChatGPT using different open-ended prompts that sought answers to the same essential questions.

Gen AI technology has developed rapidly since this pilot study was conducted. Therefore, approaches to mitigate or resolve some of these limitations may already be available. For example, LLMs can now provide detailed information about reasoning steps and include cited references as part of a prompt response. This allows for the suitability of an evaluation to be transparently considered by a human, potentially to refine the prompt being used. In addition, the emergence of Retrieval Augmented Generation (RAG) libraries (Lewis *et al.*, 2020) should allow evaluators to define what sources an LLM is using to conduct its evaluation, meaning that a consistent set of source documents (e.g. reports, articles and data sets) can be established and used to form a knowledge base from which to draw information during evaluation. This could have several benefits, including the potential to reduce "negativity bias" that may come from sensationalist or crisis-related reporting, as LLM responses will be significantly weighted toward the curated material included in the

RAG. Furthermore, the rapidly expanding number of available LLMs suggests that human evaluators can compare or aggregate evaluation results across LLMs, reducing the risk of relying on the output of a single model. Even with paid subscriptions to multiple Gen AI platforms, this will likely offer a cost-effective and faster evaluation option over a hired panel. Finally, the recent development of open-source LLMs can provide transparency over model weights, allowing models to be evaluated for potential bias prior to selection.

6. Conclusion

It is outside the scope of our study to provide solutions to global problems. Furthermore, our study should be understood as a pilot that requires further validation, development and discussion. As noted above, the LLM landscape has changed significantly in the time since our test was conducted in November 2024. We have limited ourselves, therefore, to two recommendations, both of which point the way toward the development of more enduring frameworks at the intersection of Gen AI and foresight evaluation.

First, with respect to evaluations, we are impressed with ChatGPT's ability to validate its vertical analysis on a horizontal level using the values present in the morphological field. This suggests a consistent convergence between both the LLM's statistical algorithm and the original written values developed by the Hudson Institute. We know many foresight reports and studies have used MA over the years, and although a number of these reports and studies were more complex than the Hudson Institute study (i.e. with more parameters and elements, use of CCA, etc.), we wonder whether the methodology presented here could be adapted to evaluate other historical foresight products and develop a body of knowledge. These insights may help improve qualitative language related to forecasting and predictive analytics overall (e.g. analyzing the language that may have caused the "Innovation and Discovery" inconsistency noted above). Even if an approach that uses a basis of comparison to rank values in a morphological field does not prove itself as a reliable evaluation mechanism for the majority of historical foresight reports that used MA, we believe Gen AI can prove useful in retrospectively analyzing any foresight report that used a qualitative modeling approach. For example, Gen AI might be used to retroactively evaluate and update scenarios to rank them according to likelihood, show how real-world events might converge into each scenario, or develop them in other systematic ways. A hybrid process where Gen AI and human researchers work together has been called a "Supermind" approach by Malone (Malone, 2018). This "Supermind" approach – in which Gen AI builds on and interprets the results of original human analysis – may prove more credible than scenarios developed by opaque Gen AI models that conduct foresight research from scratch.

Second, with respect to ongoing or future foresight reports, we strongly recommend that researchers who are using scenario-based and/or qualitative modeling consider structuring and presenting their work products using MA or a similar methodology where a basis of comparison can be systematically applied at a future point for purposes of evaluation. The schema used for *The Next 200 Years* may even provide a starting framework to build on. If we assume Gen AI offers a cost-effective method of evaluating foresight studies, which, at a minimum, provides results comparable to a panel of human experts, this will allow future forecasts to be systematically evaluated and potentially turned into quantitative data in the way we have done here.

For both of these recommendations, we expect a robust discussion to occur surrounding best practices for this effort, and the conversation will continue to evolve as new evaluation methods and LLMs are developed. Nevertheless, given the power and opportunity afforded by Gen AI, it seems like a waste of intellectual capital and the prolongation of an identified (but unaddressed) evaluation need to structure qualitative predictions in a way that cannot be retroactively revisited, systematically evaluated and refined by future AI systems.

Acknowledgements

The author wishes to acknowledge Professor Arho Suominen, VTT Technical Research Center of Finland, who provided invaluable feedback on a draft of this manuscript.

Declaration of interest

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Notes

[1.] While the description of MA here is substantially accurate, it may vary slightly from other descriptions. MA is effectively cross-disciplinary and therefore the reader will find many variations in terms and approaches that, while substantially the same, vary slightly in the particulars.

[2.] As Kahn *et al.* wrote: "Obviously [. . .] two basic images encompass a range of differing views and concepts, so to represent them fairly and without exaggeration, we have developed two detailed views for each of the two models – one of which in each case is a relatively extreme position, the other a moderate one. Thus, the neo-Malthusian model includes the views of a strong neo-Malthusian and a moderate neo-Malthusian (that is, a guarded pessimist); and for the contrasting model, we describe the positions taken by a moderate (or guarded optimist) and an enthusiastic advocate of technology and economic growth."

[3.] Working for *Commentary* as a music journalist and book reviewer turned out to be only a prelude for the Harvard-educated Kovner. A year after writing his review of *The Next 200 Years*, Kovner took up commodities trading and wound up becoming a billionaire hedge fund manager, philanthropist, and chairman of the American Enterprise Institute's Board of Directors.

[4.] Zero was not a ranking option because we determined it would not make sense to assign any issue a zero percent fit given the more general qualitative descriptions originally provided by the Institute. In other prediction situations, assigning a zero could make sense (e.g. for a dated event that can be proven demonstrably false).

References

Aligica, P. (2007), *Prophecies of Doom and Scenarios of Progress: Herman Kahn, Julian Simon, and the Prospective Imagination*, Continuum/Bloomsbury Publishing, London and New York.

Aligica, P. (2009), "Julian Simon and the 'Limits to Growth' Neo-Malthusianism", *Electronic Journal of Sustainable Development*, Vol. 1 No. 3.

Álvarez, A. and Ritchey, T. (2015), "Applications of general morphological analysis: from engineering design to policy analysis", *Acta Morphologica Generalis*, Vol. 4 No. 1, available at: www.swemorph.com/amg/pdf/amg-4-1-2015.pdf

Ayres, R. (1969), *Technological Forecasting and Long Range Planning*, McGraw-Hill, New York.

Bardi, U. (2011), *The Limits to Growth Revisited*, Springer, New York, doi: 10.1007/978-1-4419-9416-5.

European Federation of Engineering Consultancy Associations. (2024), *EFCA 2024 Future Trends Report: Prompt Engineering Manual*.

Gardner, A. and Barela, E. (2024), *Applying Evaluation Thinking and Practice to Foresight*, Association of Professional Futurists.

Ghamari-Tabrizi, S. (2005), *The Worlds of Herman Kahn: The Intuitive Science of Thermonuclear War* Harvard University Press, Boston, Massachusetts.

Halawi, D., Zhang, F., Yueh-Han, C. and Steinhardt, J. (2024), "Approaching human-level forecasting with language models", *Neural Information Processing*, available at: https://arxiv.org/abs/2402.18563

Johansen, I. (2018), "Scenario modelling with morphological analysis", *Technological Forecasting and Social Change*, Vol. 126, doi: 10.1016/j.techfore.2017.05.016.

Jones, J. and Thornley, D. (Eds.) (1963), "Conference on design methods. "The morphological approach to engineering design", pp. 115-140, available at: https://dl.designresearchsociety.org/design_methods_conference_1962/

Kahn, H. (1998), "Choosing a perspective on the future", American Outlook. The Hudson Institute.

Kahn, H., Brown, W. and Martel, L. (1976), *The Next 200 Years: A Scenario for America and the World*, William Morrow & Co., New York.

Ko, B.K. and Yang, J.S. (2024), "Developments and challenges of foresight evaluation: review of the past 30 years of research", *Futures*, Vol. 155, doi: 10.1016/j.futures.2023.103291.

Kovner, B. (1976), "The Next 200 Years", by Herman Kahn, William Brown, and Leon Martel; *RIO: Reshaping the International Order*, edited by Jan Tinber." *Commentary*.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S. and Kiela, D. (2020), "Retrieval-augmented generation for knowledge-intensive NLP tasks", 34th Conference on Neural Information Processing Systems (NeurIPS).

Malone, Thomas T.W. (2018), *Superminds: The Surprising Power of People and Computers Thinking Together*, MIT Press, Cambridge, Massachusetts.

Meadows, D.H. (2008), Thinking in Systems: A Primer, Chelsea Green Publishing, White River Junction, Vermont.

Meadows, D.H., Meadows, D.L., Randers, J. and Behrens III, W.W. (1972), *The Limits to Growth*, Universe Books, New York.

Naveed, H., Khan, A., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N. and Mian, A. (2024), "A comprehensive overview of large language models", Version 10. arXiv. https://arxiv.org/abs/2307.06435

Nebel, A., Kling, A., Willamowski, R. and Schell, T. (2023), "Recalibration of limits to growth: an update of the World3 model", *Journal of Industrial Ecology*, Vol. 28 No. 1, doi: 10.1111/jiec.13442.

Popper, R., Georghiou, L. and Miles, I. (2010), *Evaluating Foresight*, The University of Manchester, Manchester, England.

Richardson, J. (1985), "The resourceful earth: optimism and confrontation", *Futures*, Vol. 17 No. 5, pp. 464-474, doi: 10.1016/0016-3287(85)90058-8.

Rossi, P., Lipsey, M. and Freeman, H. (2004), *Evaluation: A Systematic Approach*, 7th Edition. Sage Publications, United States of America.

Sardar, Z. (2010), "The Namesake: futures; futures studies; futurology; futuristic; foresight—what's in a name?", *Futures*, Vol. 42 No. 3 doi: 10.1016/j.futures.2009.11.001.

Schoenegger, P., Tuminauskaite, I., Park, P., Bastos, R. and Tetlock, P. (2024), "Wisdom of the silicon crowd: LLM ensemble prediction capabilities rival human crowd accuracy", *Science Advances*, Vol. 10 No. 45, doi: 10.1126/sciadv.adp1528.

Schoenegger, P., Jones, C.R., Tetlock, P.E. and Mellers, B., (2025), "Prompt engineering large language models' forecasting capabilities", available at: https://arxiv.org/abs/2506.01578

Simon, J. L. (1977), *The Economics of Population Growth*, Princeton University Press, Princeton, New Jersey.

Simon, J. L. (1981), *The Ultimate Resource*, Princeton University Press, Princeton, New Jersey.

Simon, J. and Kahn, H. (1984), *The Resourceful Earth: A Response to Global 2000*, Blackwell, Oxford and New York.

Soroka, S., Fournier, P. and Nir, L. (2019), "Cross-national evidence of negativity bias in psychophysiological reactions to news", *PNAS*, Vol. 116(No. 38), doi: 10.1073/pnas.1908369116.

Weiss, P. (2005), "George Soros's right-wing twin", New York, NY Magazine., available at: http://nymag.com

Woodcock, J. (1977), "The Next 200 Years: a scenario for America and the world. By Herman Kahn, William Brown, and Leon Martel", *Technology and Culture*, Vol. 18 No. 3, pp. 585-588, doi: 10.2307/3103949.

Zwicky, F. (1969), *Discovery, Invention, Research*, The Macmillan Company, Toronto, Ontario.

Corresponding author

Andrew McIntosh can be contacted at: speculativefiction@gmail.com

Appendix

The best-fit narrative

Using ChatGPT 4o's ranking of the ten issues treated by all four views of *The Next 200 Years*, we can develop a narrative about the current world state as we approach the 50th anniversary of the publication of the Hudson Institute's book. The below table reproduces ChatGPT 4o's decisions about the best-fit value for each issue as well as the explanation it provided about why the selection was made. We did not reprint the entire text of the Hudson Institute's original values but instead included the descriptive phrase Hudson used to summarize each one. Refer to the original book (pp. 10–16) for the complete write-ups. At the time of writing, the book was available to borrow for free via the Internet Archive.

Table A1 The best-fit narrative for our current world state

Issue	*Value*	*ChatGPT's comment*
1. Basic world model	"Uncertain pie"	This scenario aligns well with the current reality. There is still debate over resource limits, but past predictions about shortages have often been revised as new resources are discovered or as extraction technologies improve. While concerns about resource depletion and environmental impact persist, technological advances and unexpected resource discoveries have extended resource availability. However, there is still a pressing need for conservation, given the environmental impact of excessive consumption
2. Technology and capital	"Mostly diminishing returns"	This scenario reflects several current global concerns, such as resource depletion, environmental degradation and diminishing returns on certain technological investments. The challenge of obtaining resources like rare earth minerals and the environmental impact of agriculture and industrial practices align with this prediction's emphasis on increasing difficulty and pollution in resource acquisition
3. Management and decision-making	"Likely failure"	This prediction closely aligns with the present, where rapid technological change, complex global challenges (like climate change and economic inequality) and conflicting interests are significant obstacles to effective governance and decision-making. Drastic actions are indeed being discussed and implemented to tackle these issues, reflecting a recognition of the need for reform to avoid catastrophic outcomes
4. Resources	"Continual difficulties"	This scenario is the best fit for today, as it accurately reflects the global challenges of resource management. While technology has advanced, political issues, mismanagement and the rapid demand for resources continue to complicate effective solutions. Resource scarcity and environmental stress, paired with governmental and organizational challenges, align well with this scenario
5. Current growth	"Large potential for disaster"	This scenario aligns closely with present-day concerns about climate change, pollution and resource depletion. Today, the consequences of exponential growth, especially in terms of environmental degradation, are widely acknowledged. The push for sustainability, green energy and reduced consumption reflects an awareness that "demand is not need," and there is increasing urgency for systemic reforms to mitigate impending ecological disasters
6. Innovation and discovery	"Mankind's greatest hope"	While some aspects of this scenario seem optimistic or futuristic (e.g. mastery of the solar system), this prediction resonates with current developments in technology and innovation. Today, new and improving technologies like AI, genetic engineering, renewable energy and space exploration are seen as transformative and have the potential to address major challenges such as climate change, health issues and resource scarcity. Many view this as a period of significant opportunity, aligning with this prediction's optimistic tone
7. Income gaps and poverty	"Increasing and threatening"	This scenario aligns closely with our current situation. Income inequality has indeed increased globally, with substantial gaps between the wealthy and the poor within many countries and across nations. Though extreme poverty has declined, wealth concentration among the rich has led to increased calls for equity and political movements that challenge existing economic disparities. There are rising concerns about political unrest, and some regions have experienced class-driven conflicts
8. Industrial development	"Necessary for wealth and progress"	Modern civilization has largely continued along the path of technological innovation and economic development, with many nations, including those previously less industrialized, increasingly participating in global industry. Although there are critiques and outcries related to sustainability and climate change, industrialization remains a cornerstone of wealth and progress in most economies. This scenario aligns well with the present
9. Quality of life	"In conflict with much growth"	This seems the most accurate, as environmental degradation, urban sprawl and technological issues impacting society are widely recognized challenges. There is broad acknowledgment that economic growth and technological advances do not always equate to an improved quality of life, and priorities regarding sustainability and well-being are being reassessed
10. Long-range outlook	"Contingent disaster"	This scenario closely mirrors current concerns. While we face challenges from climate change, pollution and resource management, global society is not yet in total collapse. Disasters are occurring (wildfires, extreme weather events, pollution crises), but they are still partly manageable. Urgent action is emphasized to prevent further escalation, fitting well with today's dialogue on sustainability and climate crisis mitigation