7 Levels of Hybrid Human and AI Decision Making

Published 30 June 2021 - ID G00724749 - 19 min read

By Analyst(s): Gareth Herschel, James Richardson

Initiatives: Artificial Intelligence; Data and Analytics Strategies

We trust artificial intelligence to manage retirement accounts and diagnose diseases. But how do we ensure we can trust it to make business decisions? This research discusses the seven ways data and analytics leaders should blend human and AI decision making.

Additional Perspectives

- Summary Translation: 7 Levels of Hybrid Human and AI Decision Making (21 July 2021)

Overview

Key Findings

- In the 2020 Gartner Reengineering the Decision Survey of executives, only 20% agreed with the idea that: "With sufficient data and a well-trained AI, any decision can be automated."
- Many business decisions are made by people, but most could be improved or accelerated by increasing the role of analysis as part of the process. This means reducing or at least changing the role of the human in the process; however, this is often perceived as a threat to employees' credibility, authority or job security.
- Errors in the adoption and use of artificial intelligence (AI) are often driven by the deployment of AI systems with a lack of oversight that is not justified, given their relative immaturity. AI systems can be deployed more rapidly and with lower risk if the right blend of human and machine is maintained.

Recommendations

Data and analytics leaders responsible for artificial intelligence efforts should:

- Begin by using data and analysis to *support* human decision makers.
- *Augment* decisions by presenting recommendations to human decision makers.

- Maintain control of complex systems by using humans to review *automated* decisions after they have been made.

Introduction

Stereotypes about the relationship between analysis or AI and human decision makers are plentiful. Three of the most common are:

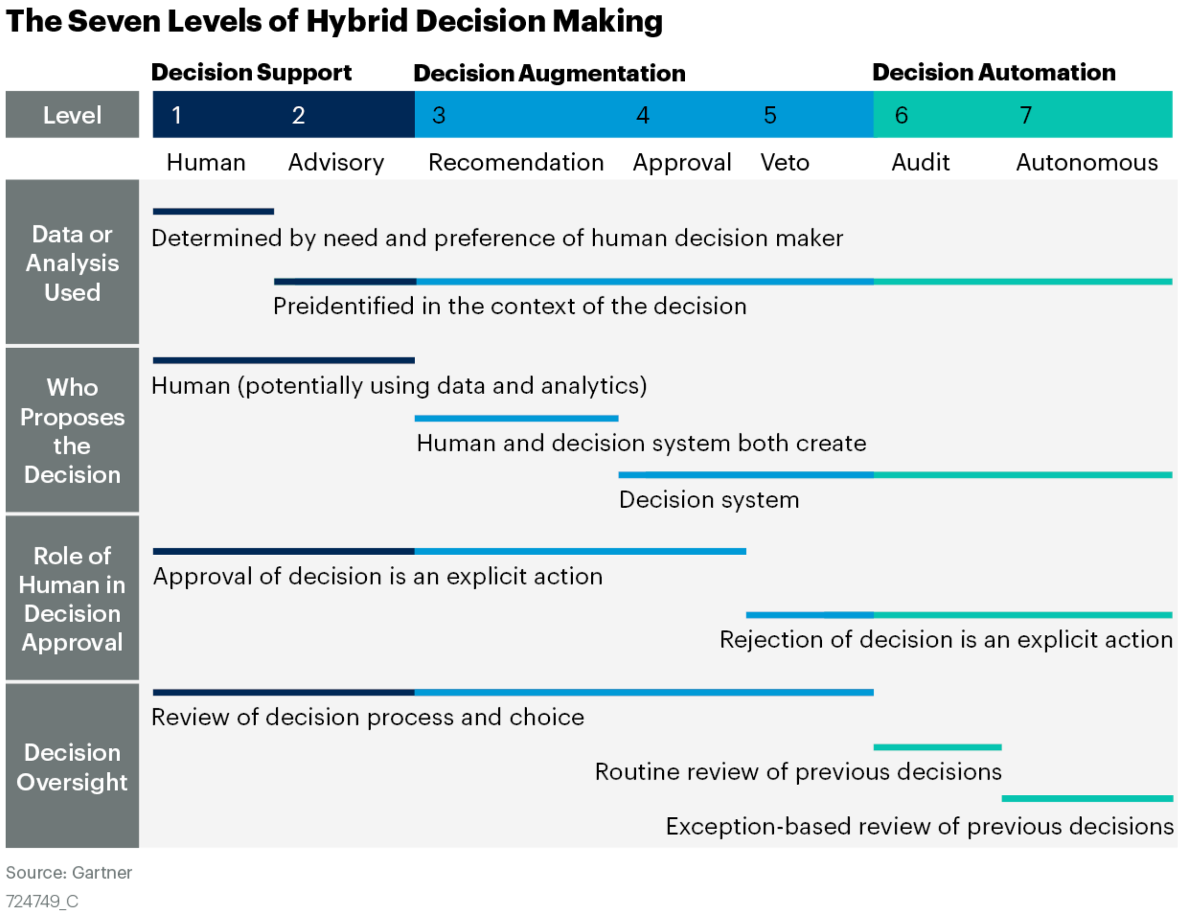
1. The machine does the analysis; the human makes the decision (“decision support”).

2. AI will replace human decision makers (“decision automation”).

3. AI will enable new categories of decision making that are not possible for a human.

But reality is more nuanced, with a series of categories defining a transition from decision support to augmentation to automation. This research discusses seven levels of hybrid human/machine decision making, ranging from “human” to “autonomous” (see Figure 1). By using this framework, organizations can be clear about where “in the loop” the human can and should be involved, and the types of investment required.

Figure 1. Seven Levels to Blend Human and Artificial Intelligence in Hybrid Decision Making



Three considerations will improve success in using this framework:

- Move from one level to the next based on organizational need, not technical ability. Using technology to advance beyond the organization's readiness will result in poor acceptance and usage, or misuse of the technology, with potentially catastrophic outcomes. Think of the scenario in British writer Arthur C. Clarke's science-fiction novel "2001: A Space Odyssey." The AI computer character HAL 9000 was given a poorly defined objective of ensuring the success of the mission without considering the implications if it came to believe humans were jeopardizing the success of the mission.

- Blend different levels, even within one decision, to balance risk and opportunity. For clarity, we present these levels as discrete steps. In reality, they will probably be used in combination; for example, decisions with low financial cost or risk may operate at a lower level of automation than the same decision with higher cost or risk implications. This is both an effective way of testing the system as it matures and a way to gradually build organizational acceptance of the system.
- Do not view Level 7 as an automatic best practice or objective; the optimal level will vary for different decisions and depend on your culture. There is a tendency to view "higher" levels as better, as in: "AI is better than reporting" or "All organizations should aspire to use AI." In reality, "better" is an extremely subjective term. Even if we accept the premise that higher levels are more sophisticated than lower levels, this does not imply that they are intrinsically better or more appropriate. A Ferrari may be "better" than a Ford in some ways, but it doesn't mean everyone should be driving a Ferrari.

Analysis

> **"When two men always agree, one of them is unnecessary."**
>
> *— William Wrigley Jr., American salesman, manufacturer and chewing gum industrialist*

Use Data and Analysis to Support Human Decision Makers

Level 1 — Human: Normal Decision Making

This is the starting position for most decisions. Humans are responsible for making the decision, and they may or may not use data or analysis to provide insight to reach one. The equivalent automotive metaphor would be a human driver looking at vehicle data on the car dashboard — the driver may be influenced by the vehicle's speed, but there is no suggestion the car is being driven by the speedometer.

Decisions can be made based on a number of factors, including business or ethical principles, advice from experts, personal experience, or data and analysis. Most decisions that use data and analysis begin with the identification of informal best practices. That is, somebody starts using a particular report or piece of insight to inform the decision making, and the practice spreads as it is discovered to be helpful. This allows for widespread experimentation with a variety of data used in different ways without significant investment (to start with) in creating a formal decision-support environment.

Strengths

- This approach can leverage a wide variety of analysis in combination with the highest level of human autonomy, allowing more nuanced decisions to be made. Ironically, many organizations self-limit the potential of this level by establishing rigid procedures that a human must follow (see Note 1).
- Since a human or group of humans is likely already making the decision with or without the data, adding data does not change the process. Decision makers usually see this as beneficial since they are still free to interpret or ignore the data as they wish. This allows the easiest integration of data and analytics into a decision-making process and the highest probability of success.

Pitfalls

- Cognitive biases prevent humans from making the right choice, even when presented with appropriate data. These biases may be behavioral, such as anchoring, confirmation or selection bias, or social, like ingroup bias.
- The decision maker is not able to understand or correctly apply the data and analysis they are presented with (data illiteracy).

Relevant Technology

- Analysis that transforms data into more meaningful data is common in this decision support model. For example, reports are often just simple aggregates or visualizations of data; predictive scores are more complicated transformations of data. Credit scores, for example, are a combination of various data points about a person, but an organization must look at the score and decide whether that person is a good or bad prospect for a loan — the score is not a decision. The analysis is not prescriptive, and the analysis may not have been created for this specific purpose, but it is useful.

- Support for normal decision making also can be as simple as building alerts to decision makers that they need to think about making the decision.
- Smart analytic systems are increasingly starting to learn by example how people make decisions. Simple examples include:
  - Offering personalized suggestions of combinations of reports to look at. (“People who looked at this report also looked at these reports.”)
  - More sophisticated versions at higher levels of automation apply this same thinking by looking at the current conditions, the decision that is made and then “dumbly” replicate the behavior. (“When the situation looks like this, we take this action, even if I don’t really understand why.”)

Level 2 — Advisory: Avoid Catastrophic Human Mistakes

The human still makes the decision, but relevant analysis is presented in the context of the specific decision. Which analysis is relevant and how it should be used are increasingly a defined part of the decision-making and business processes. By presenting an analytically driven “best guess,” the difference between human intuition and the data officially considered to be the most relevant can be easily identified.

Strengths

- Having an analytic system provide a second opinion can reduce the likelihood of a catastrophic mistake because the human considered a key insight with very little gravitas. Indeed, this can be helpful where humans are good at using certain inputs, such as deep insights from a focus group, but the machine is better at broad but shallow analysis of social media sentiment. By adding analytic diversity to the decision process, risk is reduced.

Pitfalls

- Differences of opinion can be seen as a “failure” of the dual approach (i.e., one is “wrong”) or result in analysis paralysis. In reality, they are an opportunity to learn what drove the opinions to diverge.

Incremental Technology

- Technology here is often the same as in Level 1. However, the emphasis shifts to pushing relevant information to the decision maker in the context of the decision (i.e., “you need to see this”) rather than relying on the user to pull data and put it in context. Using the previous example of a credit score, in a Level 2 system, a high-risk score may be color-coded to draw attention to the fact that it represents a heightened risk factor in the decision.

- Decision constraint systems may be added to restrict the ability of a human decision maker to make a mistake. For example, they may force a pause before an email containing emotive language to a manager is sent or warn stock traders they are selling or buying too many shares in a short period of time.

Augment Decisions by Presenting Recommendations to Human Decision Makers

Level 3 — Recommendation: Limited-Context Machine Decisions

The decision system is now replicating a person. It may not make the best decision, but it has moved beyond simply providing more data (Levels 1 and 2) to inferring what action to take. The burden of the decision is still with humans; they may accept or ignore the recommendation.

For example, contact center agents could be given a recommended product to cross-sell, but they may ignore the recommendation because the customer (who has called in to report a problem) is extremely upset. This is an intuitive decision for a human, but one that an automated system would struggle to duplicate.

Strengths

- The decision system is starting to remove the intellectual burden from humans. Humans do not need to evaluate every possible option or interpret and compare all the information that is available — although they can if they want to. This level is suitable for complex or time-pressured decisions where taking the time to identify a decision every time is not feasible.

- This approach also can be tied to other factors such as corporate policies to ensure the decision maker is always aware of the “official” option, even if specific circumstances do not make it the best option.

Pitfalls

- This is the stage where the nature of the human-machine relationship starts to play a role in success. The pressure to conform versus the desire to rebel can be a factor in adoption.
- By presenting a default option, the human decision maker ceases to be expected to look for alternatives, so “blind spots” may occur if the recommendations become too narrowly focused.
- At least initially, recommendations can be quite simplistic: offer this product or that product. The human adds the extra context that a product cross-sell is not the right action.

Incremental Technology

- The decision system is now delivering actual guided recommendations so needs to become prescriptive with technologies such as rule engines or optimization algorithms. The recommendations do not need to be integrated with operational systems (the human can see the recommendation in one system and apply it elsewhere), although there may be efficiency gains by integrating the recommendation into the operational system.
- Depending on the context, some form of explainability also may be desirable to help the decision maker understand how and why the system came to that recommendation.

Level 4 — Approval: Broad-Context Machine Decisions

The decision system analyzes the situation and identifies the best course of action, but the human is still needed to explicitly approve the action. At this level, the decision system has moved from presenting <u>a</u> recommendation to presenting <u>the</u> recommendation. The recommendation is likely to be the correct one based on a wide variety of considerations and possibly more complex (such as an optimized delivery manufacturing and delivery schedule that needs to be accepted or rejected as a whole).

Strengths

- This can be the best approach in situations with an ethical, legal or compliance dimension because human action implies conscious ownership of the decision — no action is taken without human approval.
- Building on the previous level, the decision system is now capable of considering multiple aspects of the decision, reducing the risk of an outside context problem.

Pitfalls

- If the human is expected to approve the system's decision, the human needs to understand how that decision was reached. For this reason, both the analysis and the decision logic should be explainable (no AI black boxes).
- If this approach is overused, there is a risk of decision fatigue setting in where the human starts automatically approving everything without due consideration. This is similar to the need to repeatedly click "OK" in some poorly designed computer systems when completing an action.

Incremental Technology

- This level is based on the assumption that the decision system now considers all the issues that the human decision maker would have done. This broader context probably requires more complete, richer situational awareness and so may include capabilities such as the data fabric or graph technology.
- The analytic and decision system technology is now integrated into the operational system so that approving the decision is required for the process to continue and then seamlessly advances the process after approval has been received.
- Because it is assumed that the human needs only to approve the action, an escalation process needs to track situations where the human does not approve the action so that the analysis or decision logic can be adjusted.

Level 5 — Veto: Avoid the Human Bottleneck

This is the inversion of the previous level. Instead of the human explicitly giving approval, the human's role is to prevent the machine from making a catastrophic mistake. Logically, this is a subtle distinction (in both cases, the machine and the human must agree on the course of action). In reality, there is a significant emotional distinction between the two (as the popular "trolley problem" thought experiment can demonstrate, see Note 2). This is the highest level of automation in which a human can intervene before the action takes place.

Strengths

- The advantage of this level is that it removes the human as a potential bottleneck from the process. In the decision survey we completed in 2020, 89% of respondents said that any decision is better than no decision. There may be a risk that by requiring a human to approve any action before it can occur, opportunities will be lost or problems will escalate. In such cases, it is better to allow the decision by default, particularly if the decision is one that can be reversed if found to be incorrect.

Pitfalls

- There is a logical inconsistency between Levels 4 and 5. Allowing humans the chance to veto the decision assumes they will review it before it executes, in which case they may as well make it a Level 4 decision where the human can approve the decision after the review. Level 5 makes sense only if the human will probably not review the decision, in which case permitting them the chance to veto a decision they have not reviewed could be viewed as pointless. For this reason, Level 5 decisions are best where the opportunity cost of not taking action is higher than the worst-case risk associated with a failure to examine every decision. Therefore, most decisions can be left to occur, but in certain (special) circumstances the human may choose to exert more oversight.

Incremental Technology

- The technology is generally similar to the previous level. The difference is that instead of needing to review and approve the decision (where the emphasis will probably be on explaining and justifying the logic of the decision), the time sensitivity means that fast review of a potentially large number of decisions becomes key. In this environment, visual display techniques that highlight the worst-case potential outcome or augmented reality techniques that allow for intuitive understanding of the decision become more important.

Maintain Control of Complex Systems by Using Humans to Review Automated Decisions

Level 6 — Audit: Trust, but Verify

Level 6 gives humans a "management" role rather than an "operator" one. At this level, the decision system is taking action, with a human reviewing performance and status after the event. An example is an e-commerce system in which automated cross-sell offers are continually generated based on customer spending patterns. As a routine part of the process, the system generates reports that are reviewed to evaluate performance and identify areas of concern or opportunity.

Strengths

- By allowing the automation of decision making, human involvement is focused on assessing outcomes. This allows for highly scalable decision making, but the routine nature of the review process should ensure that any negative outcomes are limited in duration and, therefore, impact.

- Extremely rapid decisions (such as self-driving cars or the control systems in a manufacturing environment) require this level of decision automation, although the human monitoring and evaluation cycle may need to be extremely rapid.

Pitfalls

- Automating decisions requires a clear sense of what the objective is. However, if the objective is poorly defined, high levels of automation can create risks that may not be apparent until later. For example, a marketing campaign that targets customers who are good prospects for a cross-sell promotion can be quickly assessed for its immediate success. But if it targets customers who have both a high propensity to purchase (good) and a high propensity to return purchases (terrible), this problem may not become apparent until much time has passed and losses have been incurred.

Incremental Technology

- Improved reporting is key to success at this level. The reporting may be accelerated (real-time reporting to allow continuous monitoring of outcomes) or multidimensional (looking at a variety of different metrics to ensure there are no unintended consequences).

Level 7 — Autonomous: Exception-Based Review

Level 7 is as far toward automated/machine-driven decision making as is reasonable to move. The idea of completely autonomous machine minds making decisions with no oversight is the stuff of dystopian science fiction. In reality, even the most autonomous decision-making systems have a degree of oversight. The difference between Levels 6 and 7 is that at Level 6 there is an assumption that decisions should be reviewed on a regular basis. Level 7 assumes that decisions will only be reviewed as needed, so parameters of acceptable decisions and outcomes are identified, and reviews only occur when these parameters are approached or exceeded.

Strengths

- The routine auditing of decisions requires effort that may not always be justified. By identifying scenarios or levels of performance/outcome that are unacceptable, resources can be focused on areas that will have the greatest impact.

Pitfalls

- Models and decisions are based on expectations of a consistent environment, an assumption that is not always justified. For example, a surge in sales of prescription medication for joint pain may be seen as a good thing, until it is realized that this is to people in their 20s and this might imply prescription abuse. If the system is not preprogrammed to look at changes in the age distribution of customers, this problem may go undetected for long periods of time. This is the classic “black swan” scenario (see Note 3) in which organizations make decisions based on experience, without realizing that their prior experience may not represent reality.

- As more and more systems and applications are automated, they may start to interact in unforeseen ways. One example is an e-commerce system that automatically discounts a slow-moving item. This increases sales, so the independently operating inventory reordering system automatically reorders more of the “increasingly scarce” product. Without human oversight of the entire system, these conflicts become increasingly likely to occur within a single organization. The problem becomes even more likely to occur when the automated actions of multiple organizations start to interact. For example, two sellers were using Amazon’s algorithmic pricing to generate marginally more revenue than each other and ended up pushing up the price of a book to nearly $24 million (see Note 4).

- Near misses can occur undetected. In the medical and aerospace industries, the idea of a near-miss accident represents a situation that could have been catastrophic, but in reality had no (or at least little) adverse outcome. Unless a conscious decision is made to classify near misses as events that must also be monitored, an autonomous system could operate (or even optimize itself to operate) dangerously close to catastrophe without triggering alerts until it was too late. Picture a self-driving car that was "trained" to consistently miss other cars by at least one inch. Technically, the vehicle is achieving its objective of not hitting other vehicles, but the margin of error should prompt reevaluation of the model.

Incremental Technology

- The ability to automatically detect outliers from normal behavior is critical at this level. Preidentifying all possible significant outcomes, and judging the appropriate threshold for a notification is impractical, implausible or both, so the increasing ability of analytic systems to detect outliers is a key technique for autonomous decision making.

- ModelOps or other approaches to monitoring for drift in the underlying data or adaptive models can detect problems earlier than waiting for changes in the environment to manifest in changing performance.

Evidence

The 2020 Gartner Reengineering the Decision Survey was conducted online from 4 August through 12 August 2020 with 147 members of Gartner's Research Circle (a Gartner-managed panel). The survey focused on understanding the context, process and examination of decision making. A team of Gartner analysts developed it, and Gartner's Research Data and Analytics team reviewed, tested and administered it.

Note 1: The Reverse Turing Test

In 1950, Alan Turing proposed an "Imitation Game" (now commonly referred to as "The Turing Test" [except in Hollywood]) to test a machine's ability to exhibit behavior indistinguishable from a human. If an observer cannot tell the machine from the human, the machine has "passed" the test. A Reverse Turing Test is one in which, when speaking with a human, you are unable to tell that you are not speaking to a machine. One example is: "I'm sorry, the system won't let me do that."

Note 2: The Trolley Problem

The trolley problem is a thought experiment in ethics about a fictional scenario in which an onlooker has the choice to save five people in danger of being hit by a trolley. To do this, the person must divert the trolley to kill just one person. The term is often used more loosely with regard to any choice that seemingly has a trade-off between what is good and what sacrifices are "acceptable," if at all (Merriam-Webster).

Note 3: Black Swan Events

The black swan theory is based on an ancient saying that confidently asserted that black swans did not exist — until they were discovered in Australia. It was popularized in a book by Nassim Nicholas Taleb.

Note 4: How a Book About Flies Came to Be Priced at About $24 Million on Amazon

Two sellers using Amazon's algorithmic pricing to ensure they were generating marginally more revenue than their main competitor ended up pushing the price of a book to nearly $24 million (Wired.com).

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

The Future of Data and Analytics: Reengineering the Decision, 2025

How to Manage the Risks of Decision Automation

When to Automate or Augment Decision Making
