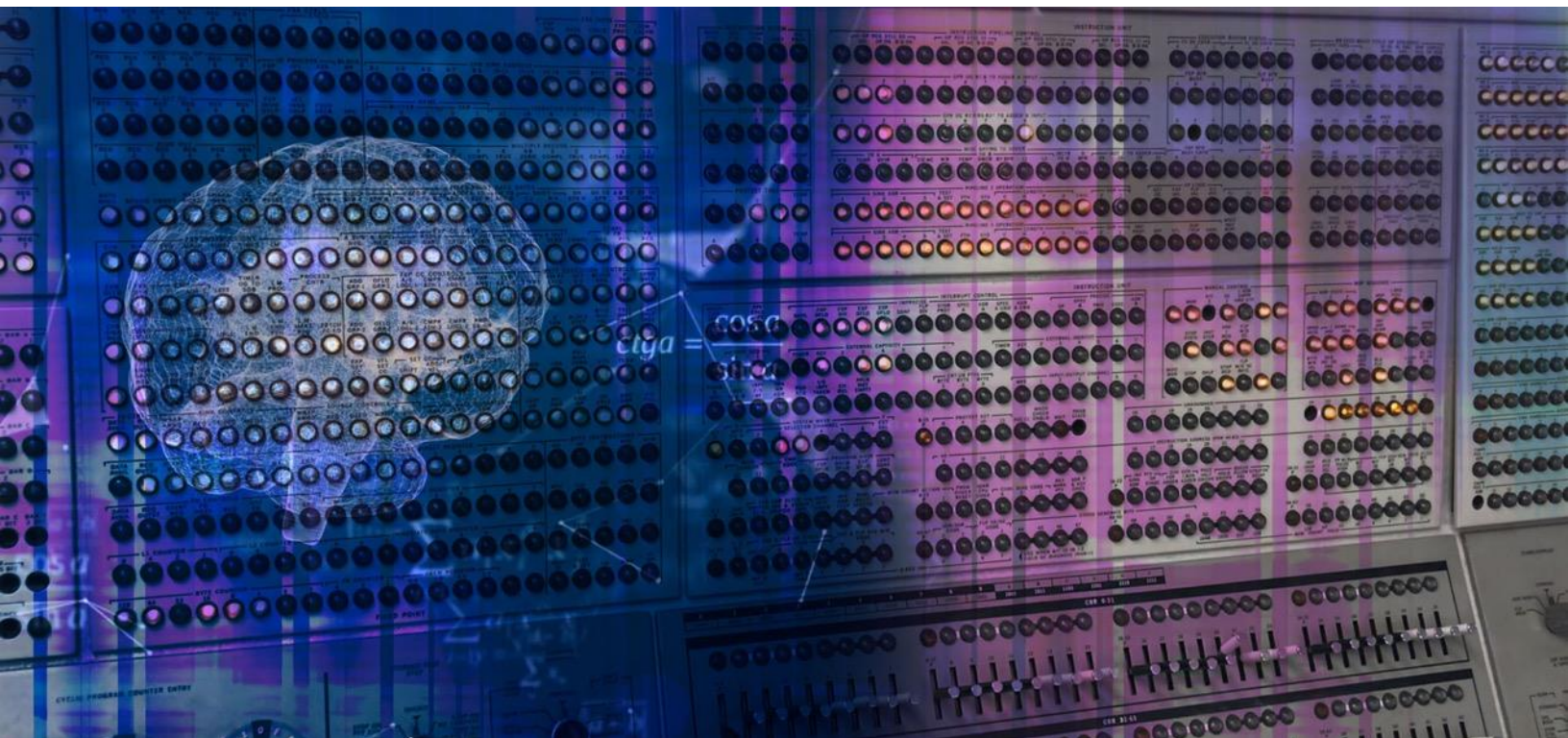




Intellyx™



White Paper

You Don't Trust AI? How to Overcome Your Fears

Jason Bloomberg

President, Intellyx

January 2022



In a recent episode of *Star Trek: Discovery*, the crew struggled with the question of how to trust their newly sentient ship's computer Zora.

The issue of trust came to a head when Zora made a unilateral decision the crew didn't like. In the face of such insubordination, is there any way the crew could trust Zora to follow the chain of command?

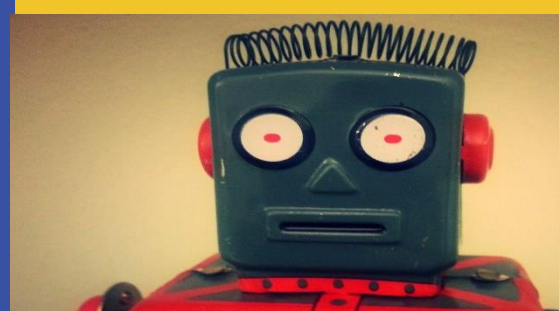
Today's AI is many years away from suddenly waking up sentient, but the question of trust is front and center in every professional's mind.

If there's a chance that some AI-driven software might get an answer wrong – either clearly incorrect or perhaps more perniciously, subtly biased – then how can we ever trust it?

A Different Kind of Software

The reason that people struggle with trusting AI is because AI works differently from other software.

An AI model is a program (or simply an algorithm) that relies on a data set to recognize patterns and make predictions or decisions. The behavior of this model depends upon the data you feed it. Good data yield good results, whereas biased data yield biased results.



an algorithm) that relies on a data set to recognize patterns and make predictions or decisions.

The behavior of this model, therefore, depends upon the data you feed it. Good data yield good results, whereas biased data yield biased results.



The data sets we feed our AI models, however, tend to be quite large and opaque. We typically have no idea what problems such data have, bias or otherwise. You might even say that one of the primary uses of AI is to uncover such issues.

Having a program tell us what's wrong with our data, however, is a far cry from providing useful insights for our business or recommendations as to what decisions we should make.

How can we trust the results from our AI, therefore, if we can't trust our data – and we have no way of uncovering their underlying issues other than the AI models themselves?

Putting Humans in the Loop

Trust, of course, must be earned – even when the party in question is AI.

The best way to build trust in an AI routine is gradually over time. Run the routine, have people evaluate the results, and repeat as needed.

Sometimes the results will be off. In such situations, either adjust the model or the data sets to better represent the goals of the initiative. Then rinse and repeat.

Over time, the AI's results will improve, as AI is able to learn over time. The people using the routine will see this improvement as the AI returns gradually improving answers.

Eventually, those answers will be good enough, where 'good enough' depends upon the business goals the organization is looking to achieve with its AI. But not only are the results sufficient, the people using the AI will *know* that the answers are good enough, because they have seen the AI improve with use.

In other words, this iterative approach to improving AI results builds trust – in both the models and the data sets feeding them.



Examples of Building Trust

Optical character recognition (OCR) is an ideal application of AI (machine learning in particular). When a computer 'reads' a scan of a document (an invoice, for example), it will attempt to interpret each character in the document.

Today's OCR programs are quite proficient, but they still confuse 'S' with '5,' 'O' with '0,' and '1' with 'l.' For documents like invoices, even a single incorrect character can become a million-dollar mistake.

Using the iterative approach to building trust, humans evaluate the results of AI-empowered OCR, making corrections if necessary. Continue this process with different documents, and over time, the AI will improve. Eventually it will be good enough to meet the business need – and people will have the confidence to trust that it has met that threshold.

Based upon data sets that include information about previous management decisions, AI can make guesses as to what a manager will decide in a new situation. As long as a manager reviews such guesses and makes corrections as needed, the AI will improve to the point that such a manager will become confident in the AI routine's decision-making skills.



Decision making can also take advantage of this iterative approach to building trust. Say, for example, that managers must approve discounts that call center representatives provide disgruntled customers when they call in with problems.

Based upon data sets that include information about previous management decisions, AI can make guesses as to what a manager will decide in a new situation. As long as a manager reviews such guesses and makes corrections as needed, the



AI will improve to the point that such a manager will become confident in the AI routine's decision-making skills.

The Intellyx Take

Krista Software has built this iterative approach to building trust in AI into its conversational approach to automation.

Krista automatically trains its AI directly from people's words, phrases, and outputs as they use the intelligent automation software.

The various data and requests people ask the AI become the input to training the AI over time, strengthening the business context of the interactions with the software, and, in turn, the eventual goal of having Krista answer a question instead of a person.

Over time, Krista learns in order to tune its output by receiving human feedback regarding desired outcomes. This collaborative process helps humans build trust in the ML and, as confidence grows, can allow Krista to take over these decisions completely.



This training input which improves the automations is conversational, as are the outputs of the software – thus empowering people to trust the AI via the familiar context of human conversation.

Over time, Krista learns in order to tune its output by receiving human feedback regarding desired outcomes. This collaborative process helps humans build trust in the ML and, as confidence grows, can allow Krista to take over these decisions completely.

We haven't made it to Star Trek's Zora quite yet, but conversational interactions with AI are an important step on the path to the future.



About the Author: Jason Bloomberg



Jason Bloomberg is a leading IT industry analyst, author, keynote speaker, and globally recognized expert on multiple disruptive trends in enterprise technology and digital transformation.

He is founder and president of Digital Transformation analyst firm Intellyx. He is ranked #5 on [Thinkers360's Top 50 Global Thought Leaders and Influencers on Cloud Computing](#) for 2020, among the top low-code analysts on the [Influencer50 Low-Code50 Study](#) for 2019, #5 on Onalytica's [list of top Digital Transformation influencers](#) for 2018, and #15 on Jax's [list of top DevOps influencers](#) for 2017.

Mr. Bloomberg is the author or coauthor of five books, including [Low-Code for Dummies](#), published in October 2019.

About Krista Software

Krista Software is in an unrelenting pursuit to help businesses find the right answers. Krista Software produces Krista, a modern Intelligent Automation platform. Krista empowers businesses to leverage existing IT assets by building low-cost automation applications. For more information visit kristasoft.com.

Copyright © Intellyx LLC. Krista Software is an Intellyx customer. None of the other organizations mentioned in this paper is an Intellyx customer. Intellyx retains final editorial control of this paper. Image credits: [Janus Sandsgaard](#), [Luis Pérez](#), and [Freddy Alberto Suárez Guerrero](#).