Conviction Analysis
for
Swarm AI

BACKGROUND:

Swarm AI technology enables networked human groups to answer questions by thinking together as a system moderated by AI algorithms. The process involves complex internal deliberations as the system weighs the available options and converges on the best solution it can across many competing constraints. Like human deliberations, some solutions are reached with high confidence, while others are conflicted. This means it's not just the answers produced by a Swarm AI system that are informative, but how each answer was reached – the internal behaviors that drove the system to converge the way it did.

The following sections describe a powerful method for analyzing the internal behaviors of Swarm AI systems called Conviction Analysis. It uses a dense Neural Network to process the deliberations and produce rigorous metrics that accurately quantify the "conviction" expressed for each solution chosen.

SEEING BELOW THE SURACE:

When a Swarm AI system answers a question, a rich dataset is generated representing the behaviors that unfold as participants act, react, and interact in real-time. On the surface, we see hints of these behaviors by watching replays of the user interface during the swarming process. For example, below are three snapshots of a Swarm AI system correctly predicting the 2016 Republican Primary:

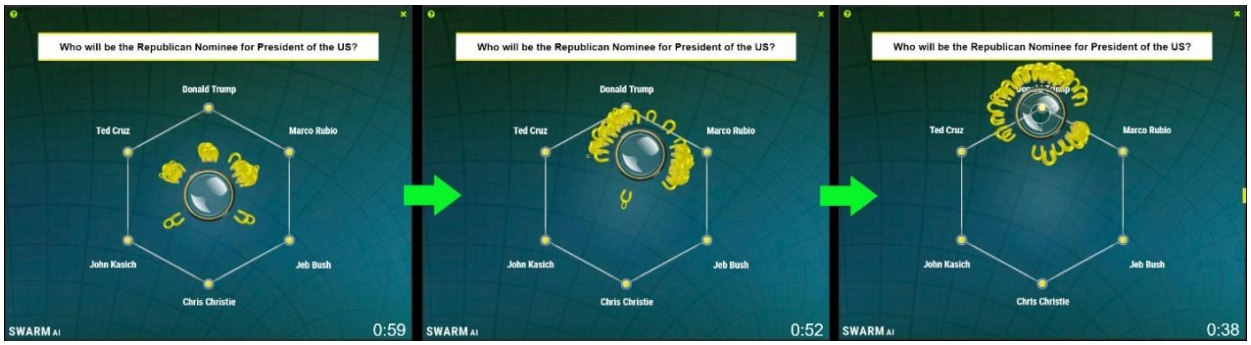

Fig 1. Snapshots of a Swarm AI system Converging upon a Political Forecast

While expressive, the above representation only shows surface features. To reveal what's going on below the surface, we use a unique visualization called a "Brainscan." It represents the subtle internal behaviors of the swarming system as it wrestles with the options and converges on a solution.

EXAMPLE: There are two ways to visualize the Swarm AI system that correctly predicted the 2016 Republican Presidential Primary. The image on the left shows the frontend client, as seen by users in real-time. The image on the right is a "Brainscan" of the backend data processed by the AI engine.

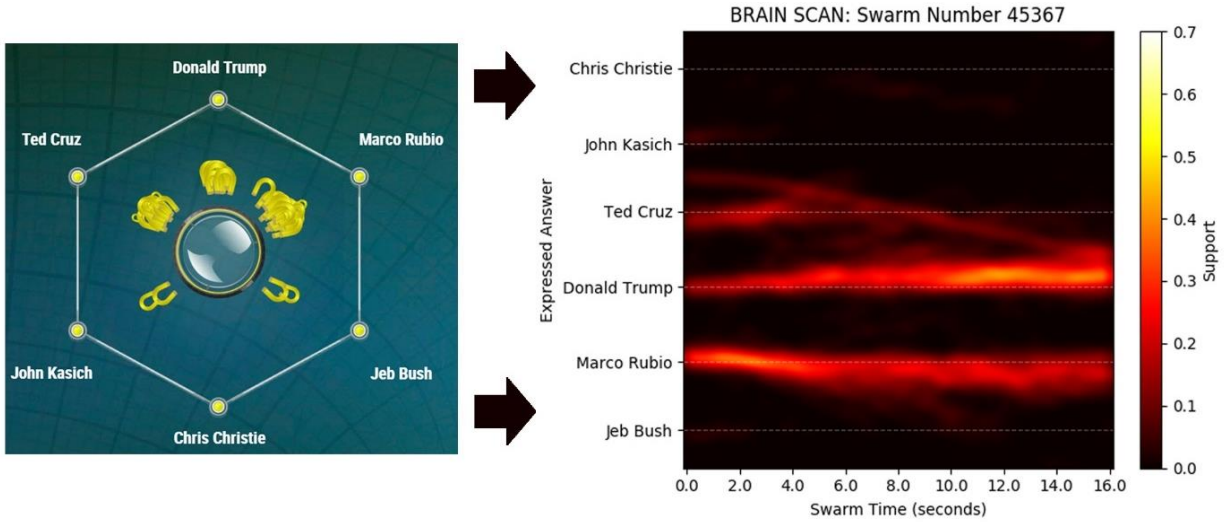

Fig 2. Brainscan represents the inner deliberations of a Swarm AI system

The colored regions on the Brainscan represent the changing sentiments within the Swarm AI system as the group converges on a solution over time, the brightness conveying strength of conviction. As shown, Marco Rubio started with the strongest initial sentiment expressed by participants, but as the deliberation progressed, conviction for Donald Trump surged while support for Ted Cruz waned.

Even before we rigorously process the deliberation data, we get insights by just viewing the scan. For example, we can easily see that support for Marco Rubio was more entrenched than support for Ted Cruz, and that Donald Trump was the beneficiary when Cruz supporters considered other options.

But can we quantify these complex deliberations with statistical rigor? It turns out we can. By training a dense Neural Network on the large database of behavioral data that Unanimous AI has collected over the last three years, we have developed a process called Conviction Analysis. It allows us to compare the strength of the swarm-based sentiments with a high degree of statistical certainty.

BEHAVIORAL NEURAL NETWORK:

Researchers at Unanimous AI have trained a Neural Network on the behavioral data collected across thousands of human swarms. By correlating the training data with known sentiments, we have been able to build an AI system that can quantify the "Conviction" of the output with statistical certainty, based not just on the answer reached, but on the complex behaviors that led to that answer.

Simply put, there's an infinite number of ways a swarming system can converge on an answer. In many cases the swarm is confident, while in other cases it may be somewhat conflicted. We have built a Neural Network that processes the subtle behavioral features and quantifies conviction with high statistical certainty. We call this process Conviction Analysis.

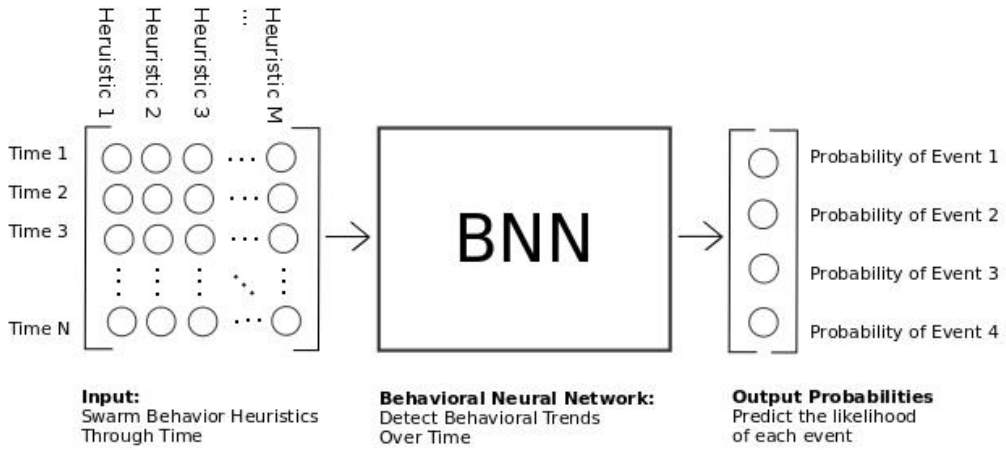

Fig 3. Behavioral Neural Network processes time-varying deliberations within Swarm AI datasets

CONVICTION ANALYSIS:

Imagine we use a Swarm AI system to predict public reaction to a new soft drink. And let's say our goal is to compare the new drink to three existing drinks. We could ask the system a series of questions to compare "*New Drink to Drink A*", "*New Drink to Drink B*" and "*New Drink to Drink C*".

The system might reveal a strong preference for New Drink over each of the existing products. That's valuable insight, *but can we compare across questions?* For example, is the preference for New Drink over Drink A stronger or weaker than the preference for New Drink over Drink B or Drink C?

This is where Conviction Analysis comes in. It allows us to assign Conviction Index to each output from the Swarm AI system, quantified by our Behavioral Neural Network with high statistical certainty. An example is shown below, quantifying New Drink over Drink A with a Conviction Index = **509**/1000.

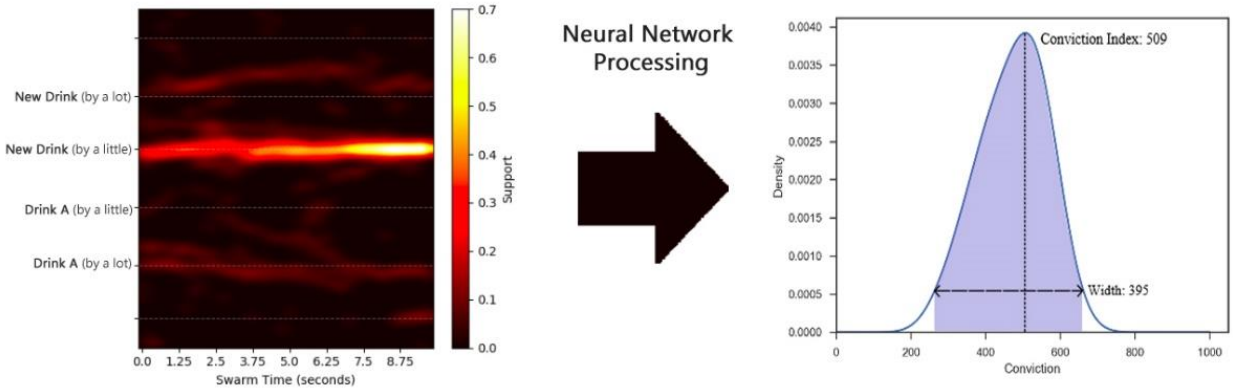

Fig 4. Behavioral Neural Network generates a Conviction Profile from Behavioral Data

We can then generate a Conviction Index for each of the example drinks vs New Drink, along with a confidence interval for 95% statistical certainty. We can plot these on the same axis and immediately compare the strength of conviction across the three soft drinks A, B, and C as follows:

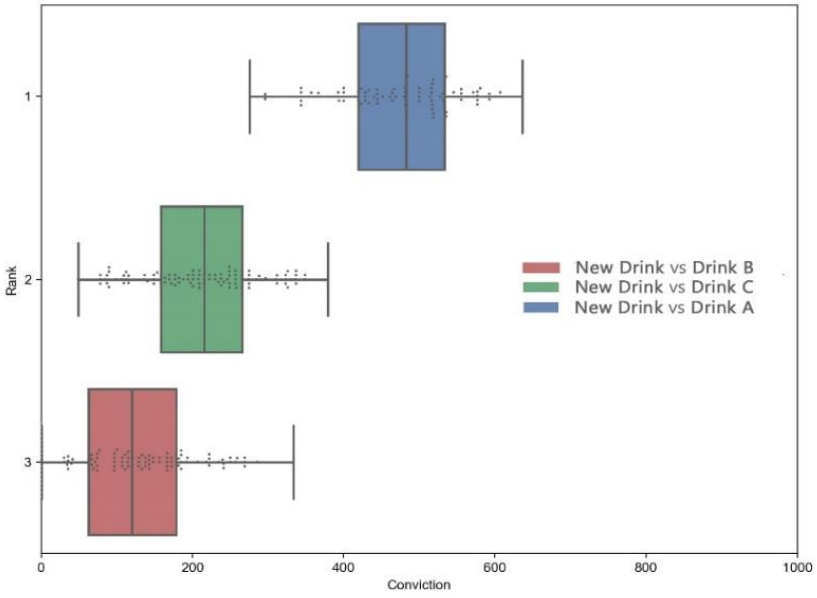

Fig 5. Statistical Comparison of Conviction across Questions

As shown, **we can quickly compare the level of Conviction across questions.** In this case, we see that New Drink was preferred to Drink A with significantly higher conviction than New Drink was preferred to Drink C, which had higher conviction than New Drink was preferred to Drink B. This is a rigorous metric, with confidence intervals, indicating that each of these differences is statistically significant ($p < 0.01$).

In addition to comparing Conviction Index across questions, we can also compare Conviction Index across populations. For example, we might want to generate soft-drink sentiments for multiple demographic groups from millennials to baby-boomers. This would allow us to compare, for example, if men prefer Drink A over Drink B with more or less conviction than women prefer Drink A over Drink B. And we can rigorously assess if the gender differences in conviction are statistically significant.

REAL-WORLD EXAMPLE:

Unanimous was asked to assess the perceived trustworthiness of major news sources in the US. To address this, we assembled a Swarm AI system that connected approximately 50 voting age Americans, controlling for political affiliation to ensure that the number of Democrats, Republicans, and Independents approximately matched the national average. When connected in real-time by swarming algorithms, these participants formed an "artificial expert" that we could ask a series of questions to. Our goal was to compare a set of twelve media sources such that statistically significant comparisons could be made.

When comparing a large set of items, it's often useful to pick one item as the index to compare against. In this case we chose the New York Times as the reference element. We then asked the Swarm AI system a series of questions, each comparing a media item to the New York Times in terms of perceived trustworthiness. An example question was formatted as follows, with the Swarm AI system asked to compare the New York Times to CNN:

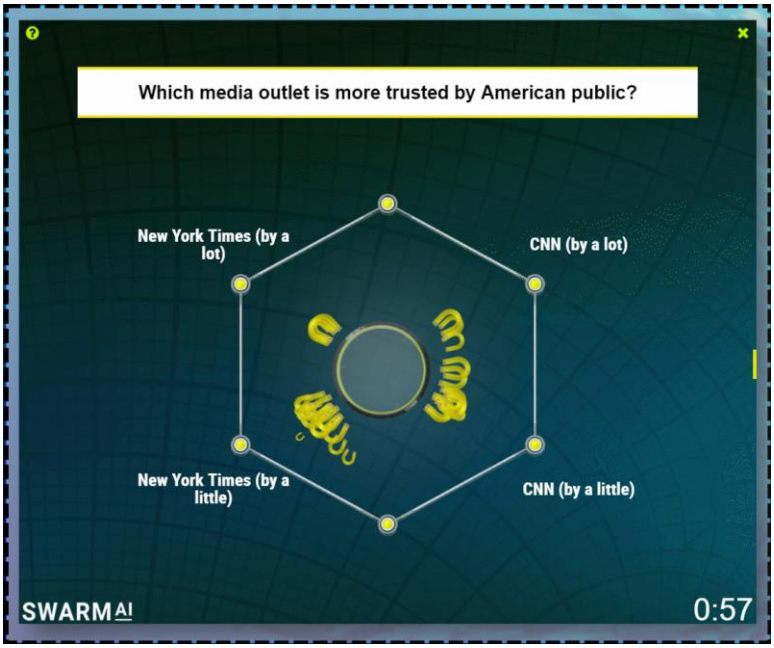

Fig 6. Swarm AI system in the process of comparing CNN to the New York Times

While the above question is asking a relatively simple question that is easy for participants to assess, providing two options for comparison, each with two levels of conviction – the behavioral data captured is very expressive. This can be represented as follows:

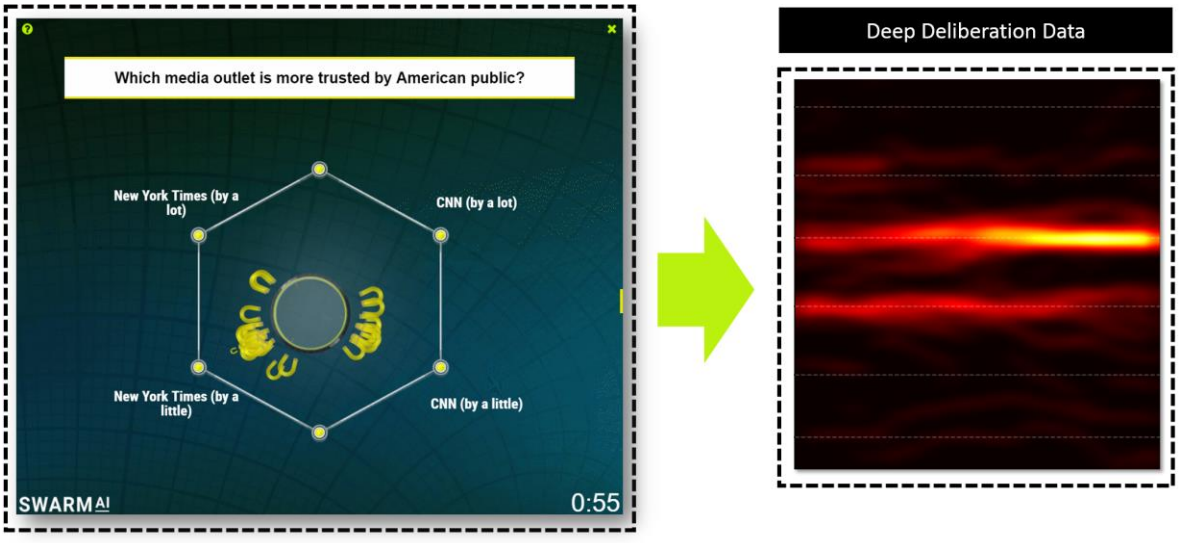

Fig 7. Swarm AI system generates rich behavioral data during every question

This behavioral data can then be processed by our Behavioral Neural Network which has been trained on thousands of prior behavioral profiles, enabling a precise assessment of conviction for the answer converged upon. Similar behavioral data is then captured for each of the other media items under assessment, enabling accurate comparisons across media items.

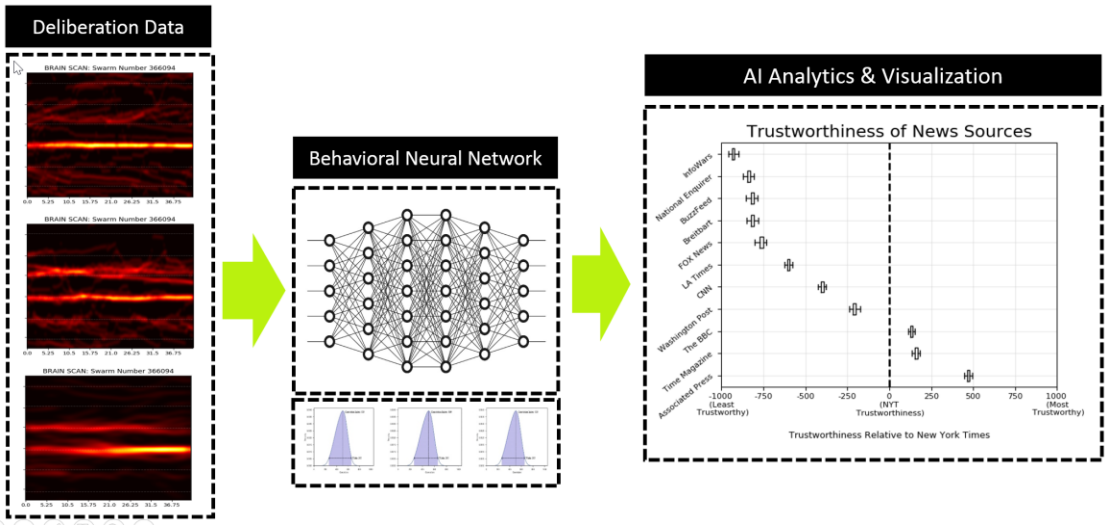

Fig 8. Behavioral Neural Network enables accurate Conviction Comparison

The output of this process is a set of comparisons between the New York Times and each of the other media items. Some sources were assessed as more trustworthy than the Times, while others were assessed as less trustworthy. And for each, a conviction index was generated by the Behavioral Neural Network, with confidence band, enabling this output:

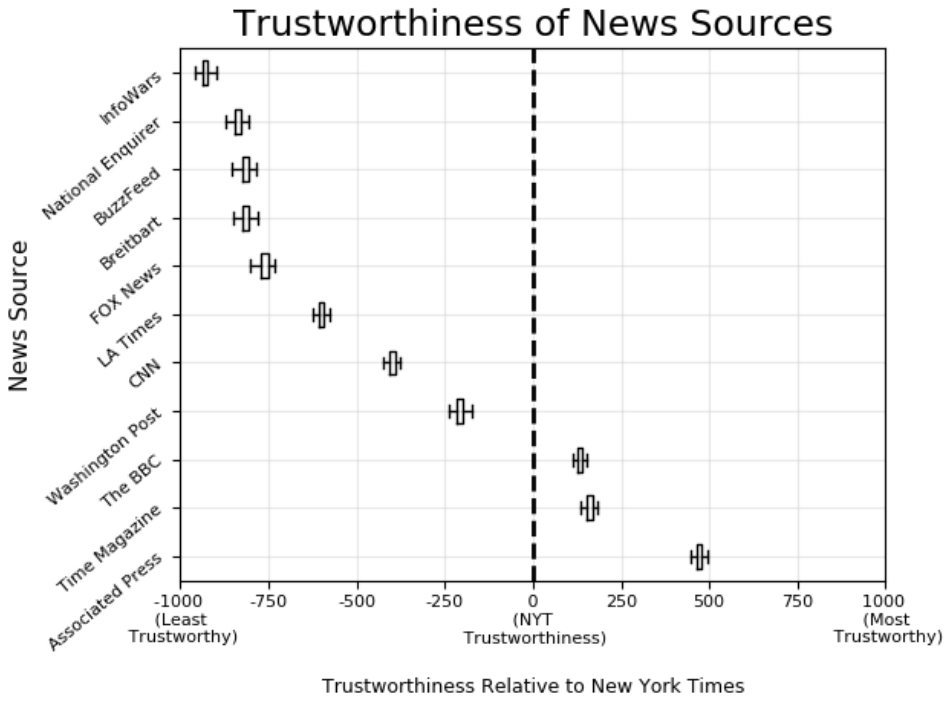

Fig 9. Comparison of News Sources using Swarm AI technology

In this way, we can quickly generate assessments with a rigorous Conviction Index values assigned to each item. While the above example involves the trustworthiness of media items, a similar process can be followed across a wide range of applications, from financial forecasting and sports handicapping to market research and business prioritization.