



Dialectical AI for Bias & Bullshit Identification / Mitigation



Evaldas Taroza, CTO
evaldas@dialexity.com

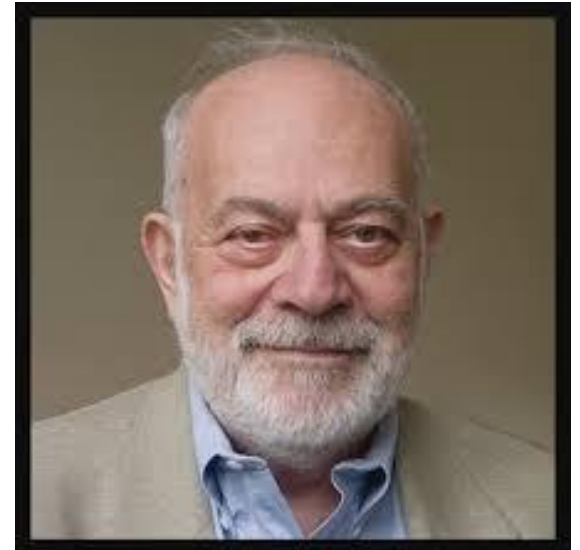


Alanas Petrauskas, CSO
alanas@dialexity.com



Worse Than A Lie

Bullshit is worse
than a lie, as it
totally disregards
the Truth



[Prof Harry Frankfurt on Bullshit](#)



Timing: Why Now?

- ▶ AI makes Dialectics practical
- ▶ We just patented the “Dialectical AI” –
USPTO issued a Notice of Allowance
- ▶ Dialectics helps to mitigate AI’s bullshit:
Dialectical AI = Artificial Honesty
- ▶ Anyone thinking about AGI will need this
technology



Problems we address:

- ▶ **Information / Bullshit overload**
- ▶ Subtle Manipulation
- ▶ Distraction from genuine priorities
- ▶ Impaired decision-making
- ▶ Biased judgment & conflicts
- ▶ Cognitive deterioration ("zombification")

Typical solutions check Facts, Credibility, and Patterns, but interpretations vary and evolve

- ▶ **Fact-Checking Tools and Platforms:**

Google Fact Check Tools, Snopes, PolitiFact, FactCheck.org, Full Fact (UK), AFP Fact Check, Reuters Fact Check

- ▶ **Bias Detection Tools:**

Ad Fontes Media's Media Bias Chart, AllSides, Ground News, The Factual, Media Bias/Fact Check

- ▶ **AI-powered Fact-Checking and Analysis Tools:**

ClaimBuster, Factmata, Logically.ai, Truly Media

No universally accepted criterion for fairness and truth exists



We Apply Dialectical Criteria

- ▶ Identifying obligations and constructivity

Truth is what obliges us to adopt the positive sides of our antitheses/obstacles

- ▶ Identifying missing information

All solutions are within us

- ▶ Identifying causes of dissonances

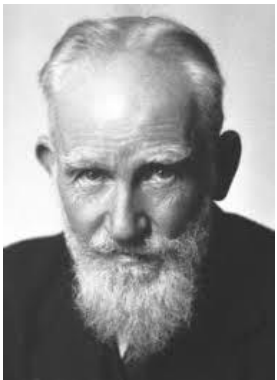
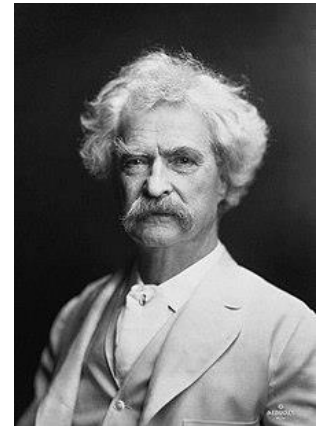
We all follow universal dialectical patterns

Global Impact Potential

Dialectic Unveils (Self-)Deception

“It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.”

Mark Twain



“Beware of false knowledge; it is more dangerous than ignorance.”

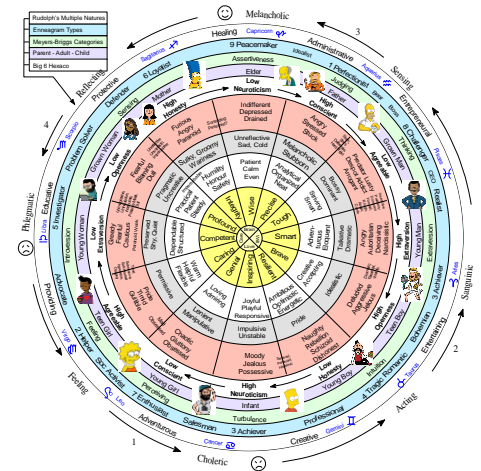
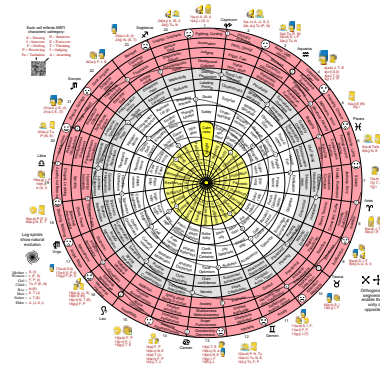
George Bernard Shaw



Our Unique Advantages

- ▶ **Unmatched Dedication to Dialectics:**

We uniquely apply dialectics across diverse fields (see 'My Writings' section [here](#))

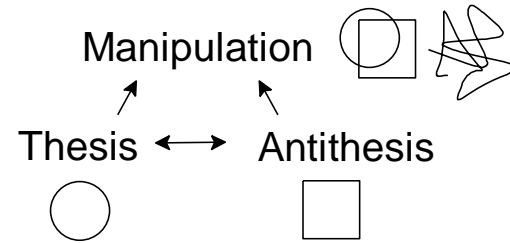


- ▶ **Patented Approach:**

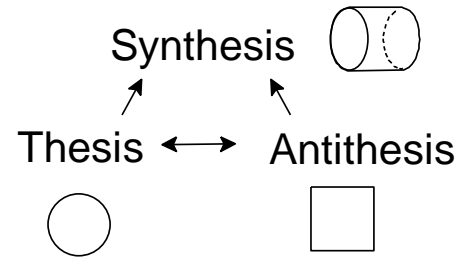
our method was approved by the USPTO (“Integration of machine learning methods with dialectical logic frameworks”)

Dialectics Breakthrough

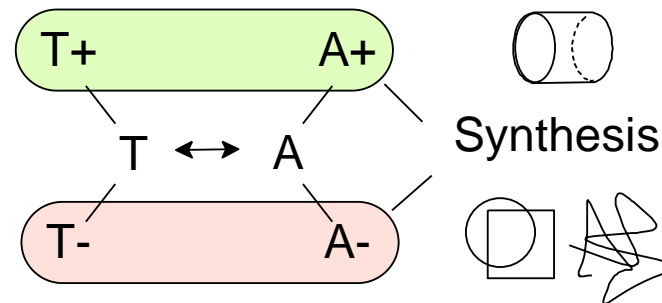
350 BCE: Aristotle systematized argumentation



1812: Hegel conceptualized dialectics



2024: We formalized synthesis prediction



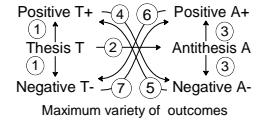
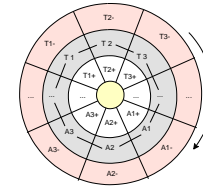


Our patent enabled solutions

- ▶ Mental Health: Identifying self-deceptions and deeper dissonances
- ▶ Cybersecurity: Detecting deceptions in hybrid warfare
- ▶ Ethics: Uncovering subtle AI biases and hallucinations
- ▶ Decision-Making: Combating "quick fix" mentality and entitlement
- ▶ Personal compatibility assessment for dating and team-building

Dialectical Analysis Routine

1. Convert text into a dialectical map



2. Identify risks, goals, obligations, constructivities

IMPLIED GOAL
Order, Justice, Liberty

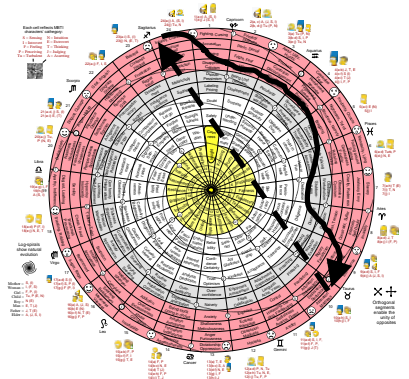
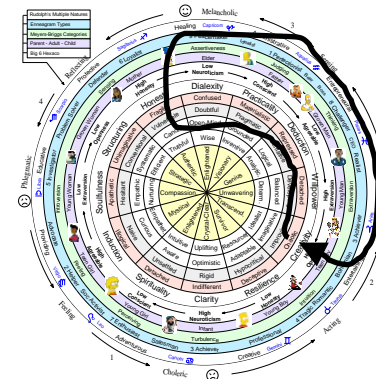
INHERENT RISK
Anarchy, Chaos, Dictatorship

INHERENT OBLIGATION
Order, Stability, Autonomy

SUBSEQUENT RISK
Tyranny, Despotism, Anarchy



3. Identify self-deceptions via "archetype loops" in reference maps

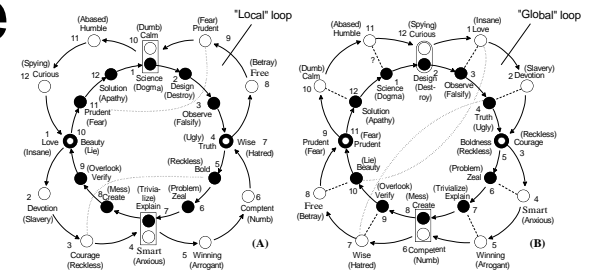


5. Identify compatibilities with others

	Inner	A/A	A/B	A/C	A/D	B/A	B/B	B/C	B/D	C/A	C/B	C/C	C/D	D/A	D/B	D/C	D/D
A-A	1	1.5	2	1.5	2	2	2.5	2	1.5	2	2	2.5	2	2	2.5	2	2
AB	2	1.5	1	1.5	2	2	1.5	2	2.5	2	2.5	2	2	2	1.5	2	2.5
AC	2	1.5	1	1.5	2	2	2.5	2	1.5	2	2.5	2	2	2	1.5	2	2.5
AD	2	1.5	1	1.5	2	2	2.5	2	1.5	2	2.5	2	2	2	1.5	2	2.5
B-A	2	1.5	2	2.5	2	1	1.5	2	1.5	2	2	2.5	2	2	2.5	2	2
B-B	1	2	1.5	2	2.5	2	1	1.5	2	2	2	2.5	2	2	2.5	2	2
B-C	2	2	1.5	2	2	2	1.5	2	2	2	2	2.5	2	2	2.5	2	2
B-D	2	2	1.5	2	2	2	1.5	2	2	2	2	2.5	2	2	2.5	2	2
C-A	2	2	2.5	2	2	2	2.5	2	1	1.5	2	1.5	2	2	1.5	2	2.5
C-B	2	2	2.5	2	2	2	2.5	2	1	1.5	2	1.5	2	2	1.5	2	2.5
C-C	1	2	2.5	2	2	2	2.5	2	2	1.5	2	1.5	2	2	1.5	2	2.5
C-D	2	2	2.5	2	2	2	2.5	2	2	1.5	2	1.5	2	2	1.5	2	2.5
D-A	2	1.5	2	2.5	2	2	2.5	2	1	1.5	2	1.5	2	2	1.5	2	2.5
D-B	2	1.5	2	2.5	2	2	2.5	2	1	1.5	2	1.5	2	2	1.5	2	2.5
D-C	2	2	1.5	2	2	2	2.5	2	2	1.5	2	1.5	2	2	1.5	2	2.5
D-D	1	2	2.5	2	2	2	2.5	2	2	1.5	2	1.5	2	2	1.5	2	2.5

Legend: 1 (white), 1.5 (light green), 2 (medium green), 2.5 (dark green), 3 (black)

4. Explore dissonances



We're in Pre-revenue Stage

- ▶ Marketing in target groups (spiritual growth & mediators)
- ▶ Revenue through subscription-based plugins and pay-per-use APIs
- ▶ So far invested ~\$30K (mostly patents)

Beta Version of Bias / Argument Inspector App

Progressive web app with voice input, installable on any device or browser: app.dialexity.com

The screenshot displays the Bias Detector web application interface. On the left, a chat history shows messages from 'Alanas Petrauskas' and 'Parent to Teen'. The main chat area shows a message from 'Parent to Teen' discussing the app's goals and revenue. Below the message, the app identifies three biases: Commercial bias (90%), Confirmation bias (70%), and Ingroup bias (60%). To the right, a 'Dialectic Constructivity Score' of 0.9 is shown, indicating a 'Nearly Ideal' score. The score is represented by a horizontal bar with a gradient from red (destructive) to green (constructive). Below the score, a detailed analysis of the text's tone and content is provided.

Bias Detector

biased and primitive results.
That's where our framework steps in— it identifies biases in GPT-4 and suggests missing information to make text more constructive.
Achieving truly ethical AI will require a dialectical approach, and our patent-pending method is at the forefront.
We've launched our MVP, Bias Detector, which is ideal for second opinions, conflict resolution, and negotiations.
Our revenue will come from subscriptions, tokenization, and licensing our IP.
We're seeking funding, partners, and co-founders to join us on this journey.

Bias Type	Percentage
Commercial bias	90%
Confirmation bias	70%
Ingroup bias	60%

Dialectic Constructivity Score

Destructive 0 Constructive 1

0.9 - Nearly Ideal

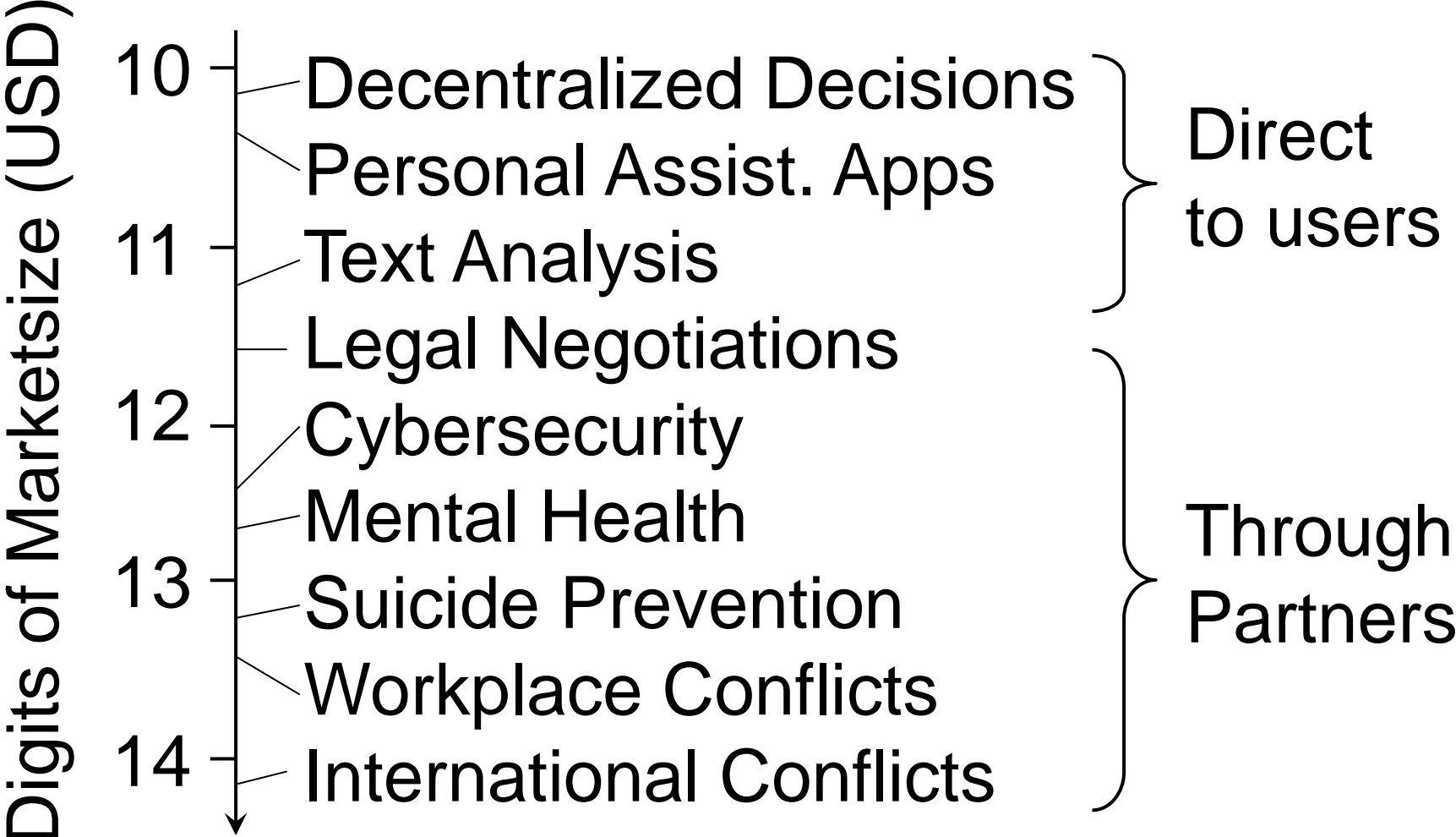
The text powerfully highlights the struggle for racial equality, while maintaining a hopeful and constructive tone. It acknowledges the existing racial disparities and discrimination, but strongly emphasizes the dream of achieving racial justice and unity. The text effectively guards against perpetuating racial discrimination and denial of racial disparities.

Second opinion in tough situations, through estimating biases, constructivity, flaws, and more

Prospective Customers

Users	What For	Requirement	Status	Payments
Young Adults	For fun	Real-time analysis	Beta testing starts now	Free
Cautious communicators	Check own text	Gmail plugin?	1-2 month work	Subscription
Social media users	Check integrity of messages	Telegram Whatsapp plugin(s)	As above	
Mediators Negotiators	Resolve conflicts	Multiparty analysis	As above	
Journalists & writers	Improve constructivty	MS Word plugin?	As above	
Business	Decision making	Note-taking plugin	As above	
Fact Checking Services	Enhance existing services	Make API	As above	Pay-per-use
HR, Dating, Mediating services	As above	As above	As above	

Related Markets



Market Size and Opportunity

- ▶ Total Addressable Market (TAM)
- Trillions USD
- ▶ Serviceable Available Market (SAM) – Billions USD
- ▶ Serviceable Obtainable Market (SOM) - Millions USD

Transforming Perspectives

“The problem is not the problem. The problem is your attitude about the problem”



See [Simple Use Case](#)

Competitors

Fact-Checking Tools and Platforms:

Google Fact Check, Snopes, PolitiFact, FactCheck.org, Full Fact (UK), AFP Fact Check, Reuters Fact Check

Bias Detection Tools:

Ad Fontes Media Bias Chart, AllSides, Ground News, The Factual, Media Bias/Fact Check

AI-powered Fact-Checking and Analysis Tools:

ClaimBuster, Factmata, Logically.ai, Truly Media

Browser Extensions:

NewsGuard, SurfSafe, B.S. Detector (no longer maintained but influential)

Social Media Fact-Checking Initiatives:

Facebook's Third-Party Fact-Checking Program, Twitter's Birdwatch

None of them use the dialectical criteria

Traction

- ▶ Provisional patent application with USPTO in November 2023
- ▶ The alpha version [Bias Detector browser plugin](#) released in March 2024 - 50 users
- ▶ Non-provisional patent application and CIP in April and May 2024
- ▶ USPTO Notice of Allowance in August 2024 (“Integration of machine learning methods with dialectical logic frameworks”)
- ▶ The [application's beta version](#) release in September 2024
- ▶ Mobile app Bullshit Detector planned in November 2024
- ▶ First integration partner using our API expected for mid-2025

Team of 2 Co-founders



↑
Alanas Petrauskas:
dialectical algorithms
& prompt developer

↑
Evaldas Taroza:
Systems Architect,
Software Developer

**Seeking cofounders in Coding,
Marketing, and Business**

Pitch

- ▶ Seeking *to be determined* amount under a SAFE Agreement
- ▶ Funding will support user research and product development
- ▶ We seek Impact Investors for strategic guidance and connections

[One Minute Pitch](#)