How Phone Data Can Help With Suicide Prevention

First Findings from the Black Box Project® on Digital Behavioral Patterns Preceding Suicide

Stop Soldier Suicide
Austin R. Grimes, Chief Product Officer
Dr. John Richardson, VP Research & Evaluation
Sarah DePalma, Director of Data Strategy & Solutions
Kim Burditt Bartlett, Sr. Manager of Black Box Project Family Engagement
Dustin Millado, Forensics Lab Manager

Pariveda Solutions Contributing Analysts
Mayur Venkateswaran, Sr. Managing Principal
Daniel Clements, Sr. Solutions Architect (former)
Thomas Krill, Manager, Engineering
Brandon Tang, Associate, Software Engineering
Matt Thompson, Associate, Cloud & AI/ML Engineering (former)

In Partnership With:
Amazon Web Services (AWS)
Cellebrite
Magnet Forensics
Pariveda

August, 20th, 2025

This publication is dedicated to those whose lives ended too soon, whose digital echoes now speak for countless others who still struggle in silence. And, to their families, who transformed profound loss into an extraordinary gift of hope. Together, your legacies illuminate the path from darkness to prevention.

Contact Information:
Austin R. Grimes
blackboxproject@stopsoldiersuicide.org
Stop Soldier Suicide

If you are a US Veteran or Service Member, or you know a US Veteran or Service Member who has ever had thoughts of suicide and is looking for free confidential therapy, please visit GoROGER.org/get-help or call us at 833-697-6437.

TABLE OF CONTENTS

A Promise to the Black Box Families

To the families who have entrusted us with your loved ones' devices. And with them, the most intimate fragments of their final chapters. We begin with the only words that matter: We see you. We see your loss, your love, and your extraordinary courage.

In the field of suicide prevention, we speak often of data points and algorithms, of patterns and predictive models. We promise to never forget that behind every dataset beats a human heart that once loved and was loved in return. Your willingness to transform the most profound personal loss into collective understanding represents an immense act of generosity. This choice to open the digital remnants of a life cut short requires a particular kind of courage. It is the courage to believe that meaning can emerge from incomprehensible loss, that your loved one's struggle was not merely an ending but could become a beginning for others. In forensic science, we call this "evidence." In human terms, we recognize it as sacred ground.

Each shared device carries within it a universe of moments: late-night searches for help that went unanswered, messages that hinted at pain too deep for words, or digital breadcrumbs marking a path through darkness. By allowing us to study these traces, you've given the scientific community something we've never had before: the unfiltered truth of crisis as it unfolds. But more than that, you've given future families the possibility of a different ending to their story.

We do not take this gift lightly. We understand that in choosing to participate, you've placed your trust not just in our technology or methodology, but in our commitment to honor your loved one's memory through action. This is a covenant we accept with the full weight of its responsibility.

Here is our promise to you: We will not rest. We will relentlessly pursue every lead, analyze every pattern, and develop every model that might identify someone standing where your loved one once stood. We will fight through the complexity of this challenge with the same determination you've shown in choosing to help others - despite your own pain. When obstacles arise, and they will, we will remember that our frustrations pale in comparison to your loss, and we will persist.

We promise to treat your loved one's data with the reverence it deserves. We promise to see beyond the tragedy to the person. Their hopes, their struggles, their humanity. And we promise that their story, understood through the lens of rigorous science and compassionate understanding, will become a beacon for those still finding their way through darkness.

Your participation transforms you from survivors into architects of prevention, from bereaved into builders of hope. You've chosen to ensure that your loved one's legacy extends beyond memory into measurable impact. Through your courage, someone's spouse, parent, child, sibling, or friend will receive the intervention they need at the moment they need it most. Lives will be saved because you chose to act when action seemed impossible.

To those families still considering whether to join this effort: We understand the weight of this decision. There is no timeline for readiness, no prescription for when or how to contribute. We will be here whenever you're ready, with the same commitment to honor your journey and your loved one's memory.

Together, we are writing a new chapter in suicide prevention. A chapter where loss leads to learning, private pain generates public good, and the worst day of your life becomes the foundation for saving others. This is the profound alchemy of the Black Box Project: transforming tragedy into knowledge, knowledge into action, and action into lives preserved.

Your loved ones' stories don't end with their final moments. Through your courage, they continue. Written in algorithms that detect crisis, in interventions that arrive in time, in futures that remain possible because you chose to share your past. This is their enduring gift to the world, made possible by yours.

We carry their memory, and your trust, in everything we do. We will not stop until the patterns hidden in their pain become the pathways to others' survival.

With deepest gratitude and unwavering commitment,

The Black Box Project® Team

ACKNOWLEDGEMENTS

Acknowledging Our Partners

The Black Box Project exists because extraordinary partners believed in an unconventional approach to understanding suicide. Their contributions have transformed a bold vision into a functioning reality that now serves families across the nation.

Cellebrite and **Magnet Forensics** have done something remarkable: they've donated the very foundation upon which this entire initiative stands. Their digital forensics platforms have allowed us to access the phone and device data, make copies of the data, and review the data for relevant content. Without their platforms, we would have no ability to honor families' wishes to transform their loved ones' devices into instruments of prevention. These industry-leading tools typically reserved for law enforcement and major investigations now serve a different kind of justice: the pursuit of answers that could save lives. Their generosity extends beyond software licenses; they've provided training, support, and unwavering belief that technology designed to enable the pursuit of justice could help solve one of public health's most complex challenges.

These partners have given us more than tools and services. They've given us the capability to keep our promise to grieving families: that their devastating loss can help prevent others from experiencing the same pain.

Pariveda has proven that the term "vendor" vastly undersells what a true technology partner can be. Time and again, they've gone beyond contractual obligations to ensure not just that systems function, but that they serve our mission with excellence. Their team approaches each technical challenge with the same reverence we bring to handling families' data, understanding that behind every dataset is a life lost and a family seeking meaning from tragedy.

Amazon Web Services (AWS) provides the secure, scalable infrastructure that makes this work possible. But more than just cloud services, AWS has been a thought partner in architecting solutions that protect sensitive data, while enabling groundbreaking analysis. They've helped us build a platform capable of growing from dozens to potentially thousands of donated devices, ensuring that no family willing to contribute will be turned away due to technical limitations.

The Deloitte Health Institute has provided a financial gift, which made part of the analyses presented in this report possible. Their contribution provided resources to conduct research. We appreciate their help to enable this work and advance knowledge in this area.

Acknowledging Our Board of Directors

The Stop Soldier Suicide Board of Directors, which includes the three founders of the organization, has maintained steadfast belief in the need to find scalable solutions for suicide prevention. That belief is reflected in their support for Black Box Project®. Brian Kinsella, Nick Black, Craig Gridelli, Melissa Baird, and Obele-West Brown: Thank you for your enduring leadership, your commitment to innovation in suicide prevention, and your belief in the transformative potential of Black Box Project.

Acknowledging Our Supporters

We extend our sincere thanks to the donors, sponsors, and foundation partners whose generosity make the Black Box Project® possible. Thank you for standing shoulder-to-shoulder with us, believing in this work, and helping turn possibility into progress. We are deeply grateful for your support.

EXECUTIVE SUMMARY

Breaking the Silence: What Digital Echoes Reveal About Suicide's Hidden Journey

In the quiet hours before dawn, a veteran searches for ways to kill himself. Days later, they text a friend about weekend plans. Within their phone lies a story that traditional suicide prevention has never been able to read until now. The Black Box Project represents a paradigm shift in our understanding of suicide, transforming the digital traces left by those we've lost into a roadmap for saving others.

After analyzing data from over 100 donated devices, we've uncovered truths that challenge fundamental assumptions about suicide prevention. The most profound discovery: those at highest risk often maintain perfect facades in their external communications while their private digital behavior screams for help. This stark divergence between internal and external digital personas explains why so many suicides seem to "come out of nowhere". We've been looking in the wrong places.

The Digital Truth About Suicide's Final Chapter

Our analysis has revealed several key insights that warrant further analysis and an eye toward improving intervention and prevention strategies:

1. **The Invisible Crisis Unfolds in Private.** Harvard's clinically-validated Suicide Risk Lexicon showed escalating risk language in private digital spaces (web searches, notes, web history) while external communications remained deceptively stable. This finding alone could revolutionize screening approaches. Traditional assessments based on what people choose to share may miss those in greatest danger.
2. **Sleepless Nights Signal Silent Struggles.** Late-night device activity nearly doubled in the final months of life, jumping from 30% to 53% with a critical inflection point at 10 weeks before death. This pattern held true across diverse behavioral phenotypes, suggesting that changes in overnight device engagement may serve as a universal early warning signal that transcends individual coping styles and could trigger timely interventions weeks before traditional clinical indicators emerge.
3. **Crisis Leaves Digital Fingerprints.** We successfully identified 106 timestamped moments of elevated risk across 42% of devices, proving that suicidal ideation manifests in detectable patterns within routine technology use. These aren't abstract risk scores, they're real moments when someone researched methods, expressed hopelessness, or reached out for help that never came.
4. **Communication & Behavioral Patterns Change Over Time.** Communication patterns revealed a significant change at multiple inflection points before death, where individuals departed from established day-time routines. Some withdrawing, others reaching out desperately. In some of our other analyses regarding financial distress and risk-related language used in private digital spaces, we found that there were substantial behavioral deviations as early as 6 months prior to death. This discovery provides potential windows and timelines for intervention that could save lives.
5. **Financial Distress: The Sustained Signal.** Unlike many risk factors that fluctuate unpredictably, financial distress showed a persistent, measurable escalation throughout the final year, accelerating in the last month. This finding positions financial institutions and assistance programs as unexpected but crucial partners in suicide prevention.
6. **Stability Masks Subtle Shifts.** While core online behaviors remained remarkably stable, specific changes such as increased visits to financial planning sites, life insurance pages, and veteran services revealed end-of-life planning hidden within normal-appearing routines. Detection requires sophistication beyond simple anomaly alerts.

[1] Low DM, Rankin O, Coppersmith DDL, Bentley KH, Nock MK, Ghosh SS (2024). Using Generative AI to create lexicons for interpretable text models with high content validity. PsyarXiv.

7. **The Cycle of Ambivalence Creates Intervention Windows.** Perhaps most critically, we discovered pronounced cyclical patterns in web browsing behavior during crisis periods. Individuals alternated between researching harmful methods and desperately seeking help through mental health resources, support forums, and religious sites. This digital manifestation of psychological ambivalence reveals something profound: even in their darkest moments, people experiencing suicidal ideation continue fighting to live. These help-seeking phases within the cycle represent crucial intervention opportunities that current prevention systems miss entirely.

From Proof of Concept to Proven Impact

This whitepaper marks the culmination of Phase 1 of the Black Box Project, proving that decedent device analysis is not only possible but essential for advancing suicide prevention. Through partnerships with AWS, Cellebrite, and Magnet Forensics, we've built the technical infrastructure to transform raw device data into actionable insights. Our collaboration with AWS Professional Services established that machine learning can detect behavioral patterns invisible to human observation, including late-night device activity that increases suicide risk by a factor of 21.

But we've always known that devices represent just one critical piece of a complex puzzle. That's why we're now launching Phase 2: partnering with researchers across the country to connect device data with medical records, social services data, and other sources through privacy-preserving record linkage. For the first time in history, we're assembling complete suicidal journeys through data, from first warning signs to final moments.

The Path Forward Requires Collective Courage

Two critical needs will determine whether these discoveries translate into lives saved:

First, we need more families to join this mission. Our current insights, while groundbreaking, emerge from a limited sample. Every device donated exponentially increases our ability to detect patterns, validate findings, and ensure our models work across diverse populations. To the families who've already contributed: your courage has illuminated paths we never knew existed. To those considering joining: your loved one's digital legacy could provide the missing piece that helps us crack the code of suicide prevention.

Second, we need the entire ecosystem to shatter the data silos that have stalled progress for decades. Healthcare systems, technology companies, financial institutions, researchers, and policymakers must embrace a fundamental truth: the status quo of fragmented data and isolated interventions has failed. While we debate privacy protocols and institutional boundaries, 50,000 Americans die by suicide each year. The Black Box Project proves that decedent device data can provide answers and insights about the journey to suicide which have never before been available. Now we need the collective will to scale it.

A Revolution Born from Loss, Powered by Love

The Black Box Project began with a simple question: What if the devices we carry every day could help us understand the journey toward suicide? Today, we have the answer. These digital companions capture the unfiltered truth of human crisis: the 3 AM searches, the drafted-but-deleted calls for help, the gradual withdrawal from digital life that mirrors withdrawal from life itself.

This isn't just about data or algorithms. It's about honoring the 50,000 stories lost each year by ensuring they become the foundation for prevention. It's about families who chose to transform their devastation into data that could spare others their pain. It's about recognizing that in our digital age, the path to suicide leaves traces. And following those traces could lead us to intervention moments we've been missing.

The technical infrastructure exists. The analytical methods work. The patterns are real and detectable. What we need now is scale: more devices to strengthen our findings, more partners to break down silos, more courage to pursue unconventional approaches to an unconscionable problem.

Join us in building a future where digital footprints become lifelines, where patterns in data translate to people saved, where no family has to wonder what signs they missed—because technology will help us see them clearly.

To loan a device or partner with the Black Box Project, visit https://stopsoldiersuicide.org/blackboxproject. Together, we can transform suicide prevention from reactive to proactive, from mysterious to understood, from impossible to inevitable.

UNDERSTANDING BLACK BOX PROJECT

When commercial aviation faced its darkest hours in the mid-20th century, engineers developed a simple yet transformative solution: the flight data recorder. This indestructible "black box" captured the final moments before tragedy, transforming catastrophic loss into actionable insights that would save countless future lives. Today, aviation stands as one of the safest forms of transportation. Not despite these tragedies, but because of our willingness to learn from them.

"In investigating [flight] accidents, anything which provides a record of flight conditions or pilot reactions for the few moments preceding the crash is of inestimable value" - Dr. David Warren in his 1954 address to the Aeronautical Research Laboratories.

The Black Box Project®, initiated by Stop Soldier Suicide, applies this same principle to one of our most pressing public health crises. Just as aviation's black boxes reveal the cascading factors that lead to crashes, this groundbreaking initiative seeks to understand the digital footprints that precede suicide, transforming profound loss into life-saving knowledge.

At its core, the Black Box Project® represents a paradigm shift in suicide prevention science. Traditional approaches have long relied on retrospective interviews, psychological autopsies, and limited clinical data. Methods that, while valuable, cannot capture the lived reality of someone's final days and weeks. The project bridges this critical gap by partnering with families who have experienced loss, inviting them to loan their loved one's digital devices (phones, computers, and tablets) that contain the unfiltered narrative of crisis as it unfolded.

The process honors both scientific rigor and human dignity. Using the same forensic procedures as federal law enforcement, the Stop Soldier Suicide team carefully extracts data while preserving the integrity of both the information and the devices themselves. Devices are usually returned to families within 90 days, with all data and hardware intact. Throughout this journey, families receive continuous support from dedicated engagement managers who understand that behind every dataset lies a story of love, loss, and the courage to prevent future tragedies.

What emerges from this data is unprecedented in suicide prevention insights. By aggregating and analyzing information from multiple sources such as text messages, search histories, social media interactions, and app usage patterns, researchers can identify behavioral patterns and warning signs that were previously invisible to clinicians and loved ones alike. Machine learning algorithms work to develop pre-suicidal behavior models, creating an early warning system that could identify individuals in crisis before they reach the point of no return. The implications extend far beyond the military community where Stop Soldier Suicide began its mission.

While suicide rates among veterans and service members are indeed more than 50% higher than the general population, suicide remains the 10th leading cause of death among all U.S. citizens. The models and insights developed through the Black Box Project can scale to serve first responders, youth populations, construction workers, and other groups at elevated risk.

For families who have donated devices, this initiative offers something precious: the knowledge that their loved one's struggle was not in vain. Their contribution transforms personal tragedy into collective understanding, creating a legacy that may prevent other families from experiencing similar loss. Each device donated represents an act of extraordinary generosity, sharing the most intimate details of loss to light the path forward for others.

For researchers and clinicians, the Black Box Project opens entirely new avenues of investigation. The comprehensive digital data provides a counterfactual dataset unlike anything previously available, enabling studies that could revolutionize our understanding of suicidal ideation and behavior. This wealth of information promises to reveal patterns and risk factors that traditional research methods simply cannot capture.

For leaders in suicide prevention and mental health, the project demonstrates that innovative approaches to data collection and analysis can break through decades of stagnation in suicide rates. It challenges us to think beyond conventional interventions and embrace technologies that can identify and support individuals in crisis with unprecedented precision.

The aviation industry didn't eliminate crashes overnight. But by committing to learn from every tragedy through black box technology, they created a culture of continuous improvement that has saved millions of lives. The Black Box Project represents our opportunity to bring that same commitment to suicide prevention, transforming our deepest losses into our most powerful tools for preserving life.

As we stand at this critical juncture in suicide prevention science, the Black Box Project® reminds us that data is never just numbers. It represents real people, real struggles, and real opportunities to intervene before it's too late. By bridging the gap between individual tragedy and collective understanding, we can finally begin to answer the question that haunts every loss: What could we have done differently? And more importantly: What will we do differently tomorrow?

OUR APPROACH

The Black Box Project represents a fundamental shift in how we approach suicide prevention. This shift recognizes that data alone cannot unlock the complexities of human behavior. Our methodology deliberately places data scientists and engineers at the intersection with those who understand suicide's human dimensions: clinical experts, forensic investigators, researchers, and perhaps most importantly, suicide loss survivors themselves.

Figure 1: SSS' approach to contextual interjections.

This collaborative approach forms the foundation of our MLOps (Machine Learning Operations) framework, which we've structured as three interconnected phases: ML Planning, ML Iteration, and ML Delivery. Unlike traditional data science projects that often operate in isolation, we've embedded clinical and research expertise directly into every stage of the process. This ensures that the questions we ask of the data are not merely statistically interesting, but clinically meaningful and actionable for prevention efforts.

Our team structure reflects this philosophy. Rather than separate silos, we operate with overlapping circles of expertise where scientific researchers, forensic and investigative specialists, clinical experts, and engineers work in constant collaboration. This intersection is where insights emerge. Here, a pattern detected by an algorithm gains context from a clinician's experience, or a forensic expert's observation guides the development of new analytical approaches.

The iterative cycle at the heart of our methodology ensures that human expertise shapes every aspect of the analysis. We begin with hypothesis generation, where contextual experts inject their understanding into the engineering process. As we prepare data, train models, and evaluate results, we continuously document our findings and re-prioritize based on what we learn. This isn't a linear process. It's a dynamic conversation between data and understanding, between algorithms and human insight.

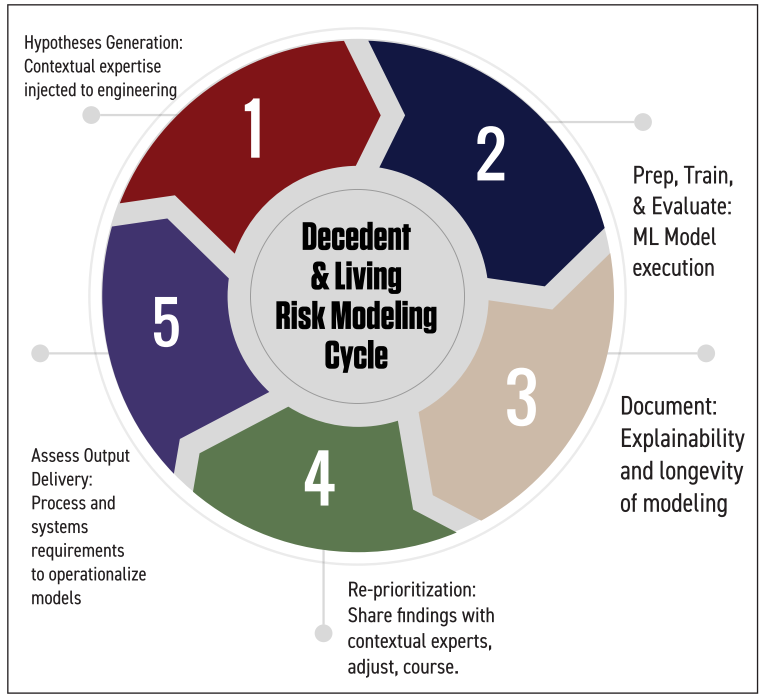

Figure 2: SSS' Machine Learning Ops Cycle

Currently, our work spans both clinical risk modeling for living individuals and decedent risk analysis for those we've lost. These parallel tracks allow us to build comprehensive models that can identify risk factors, predict critical inflection points, and ultimately inform intervention strategies.

This approach has already yielded insights that would have been impossible through traditional methods. By combining the computational power of machine learning with the irreplaceable wisdom of human experience, we're not just analyzing data. We're building a new framework for understanding one of humanity's most complex challenges. The families who have entrusted us with their loved ones' digital legacies deserve nothing less than this level of rigor, compassion, and innovation in our pursuit of understanding.

COMPLEXITY OF DEVICE DATA

Understanding the Challenge

Working with posthumously donated mobile devices presents unique technical challenges. Unlike traditional datasets with standardized formats and controlled collection methods, device data arrives in a bewildering array of structures, timescales, and completeness levels. To illustrate these complexities and their implications, we present a seemingly simple task: determining whether a message was sent or received by the device owner.

This example, while conceptually straightforward, reveals the sophisticated technical infrastructure required to transform raw device data into research-ready insights. The methodological rigor demonstrated here underpins every analysis in the Black Box Project, from social network reconstruction to behavioral pattern detection.

The Deceptive Simplicity of Message Direction

At first glance, classifying messages as incoming or outgoing appears trivial. Surely, one might assume, mobile devices clearly mark who sent and received each message. The reality proves far more complex. Different messaging applications store data in fundamentally different ways. WhatsApp structures its data differently from iMessage, which differs from SMS, which differs from Facebook Messenger. Some applications populate traditional "to" and "from" fields, while others rely entirely on a "participants" array that requires sophisticated parsing to determine directionality.

Our initial approach, utilizing only the standard "to" and "from" fields, successfully classified messages from certain applications but failed entirely for others. This discovery came only after extensive data exploration revealed that some messaging platforms never populate these fields, instead storing all participant information in auxiliary data structures. The seemingly simple task of identifying the device owner within these structures proved particularly challenging, as ownership tags appeared inconsistently and unreliably across different backup methods and operating system versions.

The solution required developing a modal contact identification system that infers device ownership through communication patterns rather than relying on explicit metadata. This approach, while more computationally intensive, provides robust ownership identification across heterogeneous data sources. Furthermore, the discovery of group messaging patterns necessitated expanding our classification schema from a simple binary (incoming/outgoing) to include "indirect incoming" messages such as those sent to groups where the device owner was a participant but not necessarily the primary recipient.

Quality Assurance in a Heterogeneous Landscape

The challenge of data heterogeneity extends far beyond message classification. Among 55 devices containing chat data, only 33 met the criteria necessary for meaningful longitudinal analysis. This dramatic filtering illustrates a critical reality of device data: not all data is created equal, and treating it as such would compromise scientific validity.

Consider the temporal patterns we encountered. Some devices contained rich messaging histories spanning years but mysteriously ceased all activity months before death, perhaps indicating a switch to a new device. Others showed sporadic bursts of activity with months-long gaps, suggesting either technical issues with data preservation or periods of non-use. Still others provided dense, continuous data but only for the final weeks of life, offering no baseline for comparison.

The implementation of DBSCAN clustering with carefully calibrated parameters (minimum 100 observations, minimum 4-week duration, maximum 5-week gaps) represents our solution to this challenge. This approach ensures that only devices with sustained, continuous data near the critical endpoint contribute to analyses. Without such filtering, we risk a subtle but pernicious bias: comparing behavioral patterns from entirely different populations at different time points, mistaking data availability for behavioral change.

The Journey from Raw to Purpose-Ready

Transforming filtered data into purpose-ready analytical datasets requires addressing fundamental questions about comparability and standardization. Death dates in our dataset span multiple years, yet we need to compare behavioral trajectories across all devices. Simply plotting messages by calendar date would conflate seasonal communication patterns with end-of-life behavioral changes.Our temporal alignment strategy anchors all devices to a common endpoint, enabling meaningful cross-device comparisons. Combined with normalization procedures that account for individual baseline communication levels, this approach reveals patterns invisible in raw data. The choice between mean and median normalization, between daily and weekly aggregation windows, might seem like technical minutiae, but these decisions fundamentally shape what patterns emerge from the noise.

The technical journey revealed unexpected insights about data quality itself. What initially appeared as a concerning trend of increasing message volumes over time, likely reflects not behavioral change but data preservation artifacts. Older messages, more likely to be archived in cloud services or deleted during device transfers, appear less frequently in our dataset. This realization underscores a critical principle: in device data analysis, absence of evidence is not evidence of absence.

Implications for Suicide Prevention Research

This deep dive into message directionality and heterogeneity illuminates broader principles that should govern all device data analysis in suicide prevention analyses:

1. The heterogeneity of device data demands flexible, adaptive analytical frameworks that can accommodate platform diversity without sacrificing standardization.
2. Temporal data quality issues require sophisticated statistical approaches to distinguish genuine behavioral patterns from data availability artifacts.
3. Even simple-seeming analyses require extensive technical infrastructure to ensure validity.

These technical complexities directly impact our ability to detect meaningful patterns in suicide risk. The directionality analysis feeds into broader investigations of social isolation, where distinguishing between reduced outgoing communication (potential withdrawal) and reduced incoming communication (potential social network changes) could reveal different risk pathways. Similarly, our social network reconstruction efforts build upon this foundational work, using validated directionality classifications to map relationship dynamics over time.

The methodological rigor demonstrated in this example extends to every analysis within the Black Box Project. Whether examining sleep patterns through device usage, detecting mood changes through communication sentiment, or identifying crisis periods through behavioral anomalies, each investigation requires similar attention to data quality, temporal alignment, and platform heterogeneity. This technical foundation, while invisible in final results, determines whether insights derived from device data can meaningfully contribute to suicide prevention efforts.

As we expand access to this dataset for the broader research community, understanding these technical complexities becomes crucial. Researchers must appreciate that working with device data differs fundamentally from traditional survey or clinical data. The richness of behavioral insights available in these digital traces comes at the cost of substantial technical complexity, a cost that, when properly managed, yields unprecedented opportunities to understand and ultimately prevent suicide.

EARLY PROJECT VALIDATION

During the initial phase of the Black Box Project in 2023, Stop Soldier Suicide partnered with AWS Professional Services to establish the foundational infrastructure and analytical capabilities necessary for this pioneering initiative. This proof of concept engagement encompassed the development of a secure cloud computing environment, implementation of critical data governance protocols, and execution of preliminary machine learning analyses on data from 15 donated devices. The work completed during this period laid the groundwork for scalable data processing, established key analytical methodologies, and generated early insights that would guide future analysis directions. This section provides an overview of the engagement structure, workstream planning, and initial findings that emerged from this foundational effort.

General Overview of the Workstream Planning

The engagement was organized into three primary workstreams to ensure systematic progress.

1. The Account Setup workstream established new AWS commercial and GovCloud accounts within a secure organizational structure, enabling proper data governance and security protocols.
2. The Data Lake Design workstream focused on architecting a scalable repository for storing and processing the sensitive device data.
3. Additionally, the Tool Preparation workstream developed essential analytical pipelines, including natural language processing capabilities for text analysis, search functionality for message content, and standardized templates for exploratory data analysis across different types of device data.

Summary of Findings for Social Isolation

The social isolation analysis revealed compelling patterns in communication behavior leading up to suicide events. By examining text message and chat frequencies over 12-month periods, the team identified statistically significant trends in both message volume and the number of daily conversations. Most notably, the analysis detected critical change points occurring approximately three months and two months before the suicide event, suggesting that changes in social connectivity patterns may serve as early warning indicators. These findings support the established connection between social isolation and suicide risk, providing quantitative evidence that could inform future intervention strategies.

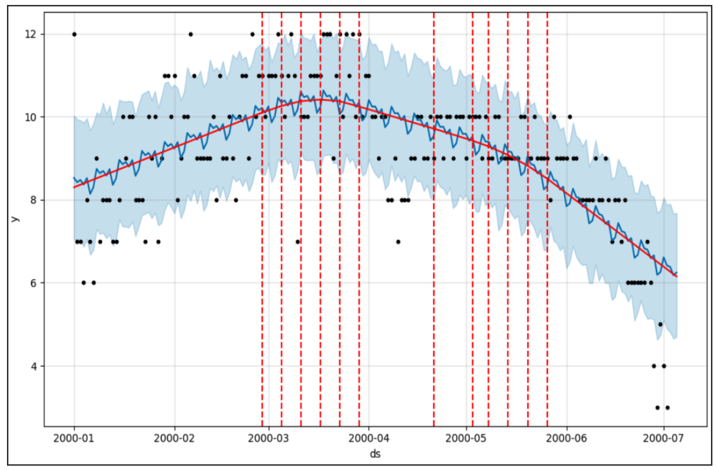

Figure 3: 6-month time range from death showing population level average number of conversations per day.

Summary of Findings for Survival Modeling

The survival modeling analysis yielded two significant discoveries about factors associated with increased suicide risk. First, the time-invariant modeling revealed that changes in sadness levels during the final weeks were strongly predictive, with each 1% increase in sadness corresponding to a 5% increase in suicide risk.

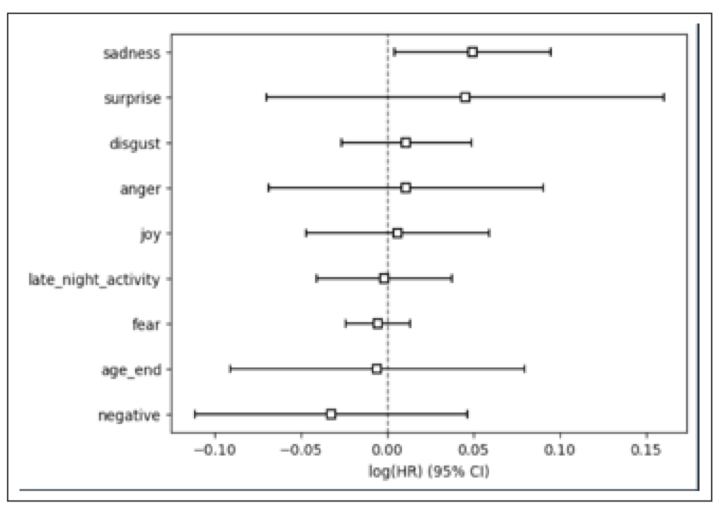

Figure 4: Forest plot using Time Invariant Modeling

Second, the time-varying analysis identified that late- night device activity (between midnight and 4 AM) dramatically increased risk by a factor of 21, while increases in anger-related content also elevated risk. These findings provide quantitative validation of behavioral indicators that could be incorporated into future risk assessment tools.

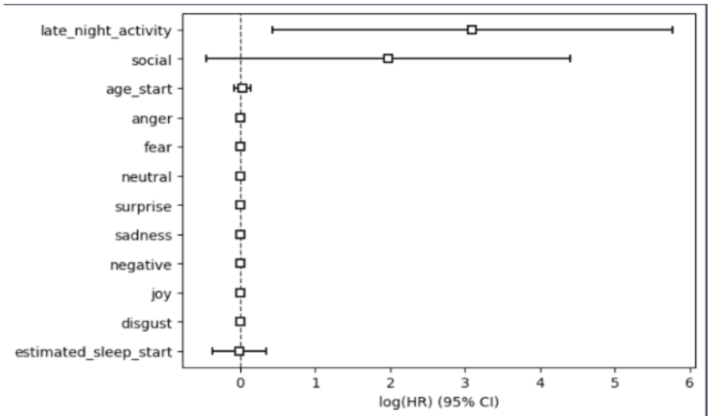

Figure 5: Forest plot using Time Varying Modeling

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
anger	0.34	1.40	0.16	0.02	0.65	1.02	1.92	0.00	2.11	0.04	4.83
disgust	-0.15	0.86	0.13	-0.40	0.10	0.67	1.10	0.00	-1.20	0.23	2.13
fear	0.00	1.00	0.00	-0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00
joy	-0.07	0.93	0.07	-0.21	0.07	0.81	1.07	0.00	-1.00	0.32	1.67
neutral	0.01	1.01	0.04	-0.07	0.09	0.93	1.10	0.00	0.29	0.77	0.38
sadness	-0.02	0.98	0.05	-0.12	0.08	0.89	1.08	0.00	-0.42	0.67	0.57
surprise	0.00	1.00	0.00	-0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00
negative	-0.01	0.99	0.05	-0.10	0.09	0.90	1.09	0.00	-0.14	0.89	0.17
late_night_activity	3.06	21.42	1.35	0.42	5.70	1.53	300.30	0.00	2.27	0.02	5.45
estimated_sleep_start	-0.04	0.96	0.18	-0.39	0.31	0.68	1.36	0.00	-0.23	0.82	0.29
age_start	0.01	1.01	0.06	-0.10	0.13	0.90	1.14	0.00	0.22	0.83	0.27

Figure 6: Results from Time Varying Modeling showing Anger and Late-Night Activity as having statistically signficant relationships with suicide risk (p<0.05)

Summary of Findings for Normalized Time Range for Negative Sentiment

The analysis of negative sentiment patterns across standardized time periods examined data over three, six, and twelve-month windows leading up to the suicide event. While the analysis revealed a general upward trend in negative sentiment across all time ranges, these changes did not reach statistical significance. This null finding is valuable in itself, suggesting that while negative sentiment may increase over time, it may not serve as a reliable standalone predictor without considering other contextual factors or more nuanced emotional indicators. This finding is something we will continue to explore and validate as our sample size increases.

Summary of Findings for Normalized Time Range for Anger Emotion

Similar to the negative sentiment analysis, the examination of anger patterns over three, six, and twelve- month periods showed a slight upward trend but without statistical significance. This finding indicates that while anger may incrementally increase leading up to a suicide event, the changes may be too gradual or variable across individuals to serve as a reliable early warning signal on their own. As we gather more devices, we will continue exploring the role of expressed anger in written messages; however, these results also underscore the importance of multivariate approaches that consider multiple emotional and behavioral indicators rather than relying on single emotional markers.

Building on the Foundation

This proof of concept engagement with AWS Professional Services established the critical infrastructure and analytical capabilities that have enabled the Black Box Project to scale significantly beyond the initial 15 devices. The robust data architecture, processing pipelines, and analytical frameworks developed during this phase provided the foundation necessary to handle larger volumes of data while maintaining security and analytical rigor. With these systems in place, the project has successfully expanded its sample size and pursued more sophisticated analyses, uncovering deeper insights into the complex patterns preceding suicide. The sections that follow present our most recent findings, which build upon these early discoveries to provide a more comprehensive understanding of digital behavioral markers and their potential for advancing suicide prevention efforts.

RECENT WORK AND FINDINGS

This brief section presents concise summaries of the Black Box Project's most significant discoveries from analyzing decedent device data. Each summary distills months of rigorous analysis into key insights that demonstrate how digital behavioral patterns can illuminate the hidden dynamics of suicide risk.

For readers seeking rapid understanding of our findings, these summaries provide the essential takeaways from each of our main analyses. We have deliberately structured this section to highlight both the practical implications of our discoveries and their potential to transform suicide prevention approaches. Each finding represents a unique window into how individuals experiencing suicidal ideation interact with technology in their final months, weeks, and days.

The analyses span from validated clinical tools applied to digital data to novel behavioral analyses that reveal patterns invisible to traditional assessment methods. Some findings confirm long-held clinical beliefs with empirical evidence, while others challenge fundamental assumptions about how suicide risk manifests in daily life. Importantly, we also present null findings where our hypotheses were not supported, as these negative results provide equally valuable guidance for future analysis priorities.

For those requiring deeper technical understanding, comprehensive appendices follow this section. Each appendix details the methodology employed, discusses findings in their full complexity, acknowledges limitations, and explores implications for clinical practice, technology development, and future analyses.

Suicide Risk Lexicon Applications in Digital Phenotyping

Harvard's clinically-validated Suicide Risk Lexicon revealed a stark divergence between internal and external digital personas when applied to posthumous device data. While private digital expressions such as personal notes, searches, and web history showed statistically significant increases in risk language during the year prior to death ($\tau=0.3512$, $p=0.0002$), external communications with others remained remarkably stable, effectively masking escalating distress. This empirical validation demonstrates that traditional risk assessment methods, which rely primarily on what individuals choose to share with others, may systematically miss those at highest risk. These findings necessitate novel analytical approaches that map correlations between external communication patterns and internal distress markers, potentially revealing subtle but detectable changes in otherwise normal-appearing interactions that could serve as early indicators for clinical intervention without requiring access to private data. This work also underscores the urgent need to fundamentally reimagine clinical screening protocols, as we now have definitive evidence that individuals experiencing escalating suicide risk often maintain complete facades during the very period when intervention is most critical, suggesting that effective screening must create psychological safety for authentic disclosure rather than relying on individuals to spontaneously volunteer their distress.

Late Night Activity Analysis

Analysis of late-night device activity (12-4 AM) across 54 devices revealed a striking temporal pattern: overnight device engagement increased from a baseline of 30% to 53.16% in the final week of life, with the most pronounced shift beginning approximately 10 weeks before death. This 77% relative increase in late-night activity suggests fundamental changes in daily rhythms that extend beyond simple sleep disturbance. Device-level analysis uncovered three distinct behavioral phenotypes: persistent users (33%) who maintained high overnight activity throughout, late-increase users (28%) who showed escalating activity primarily in final weeks, and minimal users (39%) who exhibited limited late-night engagement despite approaching end of life. The technical innovation of location-based time zone reconstruction ensured temporal accuracy across 6.25 million activity records, enabling confident identification of true overnight behavior patterns. These findings suggest that digital behavioral markers, particularly the 10-week inflection point, could serve as early warning signals for proactive intervention, offering a practical window for support delivery that current healthcare systems might miss.

Detecting Known Moments of Risk Through Digital Footprints

Advanced computational methods successfully identified 106 timestamped moments of elevated suicide risk across 42% of examined devices, establishing a validated ground truth dataset of authentic distress expressions as they naturally occur in digital life. Using semantic search and zero-shot classification approaches that capture meaning beyond literal text, the analysis revealed 20 distinct risk event types ranging from suicide methods research to help-seeking behaviors, demonstrating that individuals experiencing suicidal ideation leave detectable signals within routine technology use. This breakthrough validates the vision of technology-enabled early intervention and provides concrete evidence that digital phenotyping can yield actionable insights for protecting vulnerable populations.

Cross Referencing: Known Moments of Risk and Web Usage History

Analysis of web browsing behavior surrounding validated risk moments uncovered pronounced cyclical patterns where individuals alternated between engaging with distressing content and subsequent help-seeking behaviors, suggesting internal conflict and creating identifiable intervention opportunities. Most critically, the temporal analysis demonstrated rapid behavioral escalation, with explicit searches for suicide methods often occurring within days of death among devices showing planning-related moments. These findings support the development of continuous monitoring systems that could detect risk escalation in real-time and enable timely interventions during help-seeking phases of the cycle, fundamentally shifting suicide prevention from reactive to proactive approaches.

Social Network Analysis

Comprehensive analysis of communication patterns from call logs and chat data revealed several inflection points before death where individuals departed from their established communication routines, with some showing dramatic shifts in time-of-day patterns such as increases in late-night activity. While communication volume changes varied bidirectionally, meaning that some individuals increased outreach while others withdrew, the consistent timing of behavioral disruption suggests a potential early warning threshold for intervention. The subsequent reversion to baseline patterns in the final week adds complexity but reinforces that algorithmic risk detection built on personalized communication patterns could identify at-risk individuals weeks before crisis peaks.

Financial Distress Patterns in the Final Year of Life

Natural language processing analysis of 53 devices revealed that financial distress communications followed a distinct trajectory, with semantic distress scores rising from 0.32 to 0.35 in the final month, a statistically significant elevation that persisted even as overall communication patterns remained stable. Unlike many risk factors that fluctuate dramatically or appear only immediately before a crisis, financial concerns provided a sustained signal throughout the final year with marked acceleration in the last weeks, suggesting that economic stressors serve as a more reliable indicator than previously recognized. This finding indicates that financial institutions and assistance programs could play a pivotal role in suicide prevention by serving as engagement pathways for at-risk individuals who might otherwise withdraw from traditional mental health outreach.

Digital Behavioral Analysis in the Final Period of Life

Analysis of online behavior revealed that fundamental browsing habits remained remarkably stable even as individuals approached death, with eight of ten most-visited websites unchanged in the final month, challenging assumptions about dramatic behavioral shifts preceding suicide. However, beneath this surface consistency, financial and insurance websites showed marked increases in activity, with some sites like USAA displaying notable rises while others like Intuit and CreditKarma appeared in top rankings for the first time, potentially indicating end-of-life planning behaviors. These patterns suggest that effective digital suicide prevention strategies must evolve beyond simple anomaly detection to incorporate subtle behavioral indicators within sustained patterns, and that partnerships with frequently-visited platforms, particularly financial institutions serving military communities, could create novel intervention pathways.

Digital Interest Profiling: Email Channel Analysis

Despite successfully implementing natural language processing to categorize email content into twelve consumer interest categories across available devices, the analysis found no meaningful patterns linking digital interest profiles to suicidal outcomes, with observed temporal changes falling within expected variance for typical consumer behavior. While the technical framework demonstrated that large-scale interest profiling from unstructured email data is achievable, the absence of suicide-specific signals suggests that email- based consumer interests may not serve as reliable risk indicators. This null finding provides valuable guidance for resource allocation, indicating that suicide prevention efforts should prioritize analyzing communication patterns, social connections, and help-seeking behaviors rather than consumer interest profiles.

Movement and Physical Activity Analysis

Analysis of Apple Health step count and walking distance data using anomaly detection algorithms revealed no consistent relationship between physical activity patterns and documented risk periods across the examined population, with some individuals showing increased activity during risk events while others showed decreases or no change at all. The high variability in both data quality and individual behavioral patterns, compounded by military-specific confounding factors such as mandated training schedules and deployment requirements, suggests that movement data from consumer devices is too inconsistent and context-dependent to serve as a reliable standalone indicator of suicide risk. However, the observation that risk events often clustered around periods of behavioral transition regardless of direction indicates that movement data could contribute value as part of a multimodal assessment approach when combined with other digital biomarkers.

CONCLUSION: FROM DIGITAL ECHOES TO COLLECTIVE ACTION

The findings presented throughout this whitepaper demonstrate how digital behavioral analysis can advance our understanding of suicide risk. Through the generosity of families who loaned their loved ones' devices, we have identified patterns previously invisible to researchers and clinicians: the divergence between private digital distress and public communication, the 10-week threshold for increased late-night activity, cyclical help-seeking behaviors during crisis periods, and sustained financial distress signals throughout the final year of life.

These discoveries inform new approaches to suicide prevention. The individuals whose devices we studied maintained typical communication patterns while their private digital behaviors revealed escalating risk. They continued visiting familiar websites even as they researched methods and composed unsent messages. They alternated between distress-related searches and help-seeking behaviors, indicating that intervention opportunities exist even during acute crisis periods. Each pattern represents a potential moment for support, provided we develop appropriate detection and response systems.

We recognize that device data represents one component of a comprehensive approach to suicide prevention. Complete understanding requires integrating information across multiple domains: medical records documenting health trajectories, financial data revealing economic stressors, military service records capturing unique veteran experiences, and social service interactions marking crisis points. Current data silos prevent researchers from examining these interconnected factors. Aviation safety improved dramatically when the industry integrated mechanical, environmental, and human factors data. Suicide prevention requires similar cross-domain integration.

The technical infrastructure developed with our partners demonstrates that privacy-preserving analysis of sensitive data is achievable. Our analytical methods show that decedent device analysis yields insights applicable to living populations. The patterns we discovered suggest specific opportunities for earlier intervention. However, our current sample size and data sources limit the generalizability and comprehensiveness of our findings.

Our Call to Action

First, we need additional devices from families. Our current findings derive from approximately 100 devices. Larger samples would strengthen pattern detection, improve model validation, and ensure findings apply across diverse populations. Families considering donation should know that their loved one's digital information could reveal previously unrecognized warning signs, potentially enabling earlier intervention for others at risk.

Second, we need institutional data partnership. Healthcare systems, financial institutions, technology companies, government agencies, and social service organizations each maintain data relevant to suicide prevention. We encourage these institutions to explore privacy-preserving data sharing approaches that could contribute to comprehensive risk understanding. Federated computing and differential privacy techniques enable collaborative analysis without compromising individual privacy. Technical solutions exist; implementation requires institutional commitment to suicide prevention as a shared priority.

Third, we need expanded research collaboration. The discoveries in this whitepaper emerged from focused analysis by a small team. Broader collaboration with data scientists, clinicians, ethicists, and domain experts would accelerate discovery and translation to practice. The combination of novel data sources and advanced analytical techniques presents opportunities for breakthrough insights. Collaborative efforts across disciplines and institutions will be essential for developing effective prevention strategies.

The aviation industry's transformation through systematic accident analysis provides a relevant model. Their commitment to learning from every incident through flight data recorder analysis contributed to substantial safety improvements. Digital behavioral data offers similar potential for suicide prevention. We have established that digital traces reveal risk patterns. Families have demonstrated willingness to contribute data for prevention purposes. Technology enables privacy-preserving analysis at scale.

The challenge now is implementation. With approximately 50,000 annual suicide deaths in the United States, the need for improved prevention approaches remains urgent. The data sources exist. Analytical methods continue advancing. Early findings show promise. Progress requires coordinated action across families, institutions, and researchers.

We invite participation from all stakeholders committed to reducing suicide rates. Families who have experienced loss can contribute through device lending. Institutions can explore data sharing partnerships. Researchers can apply their expertise to this emerging dataset. Together, we can develop prevention systems that identify and support individuals before crises become tragedies.

The digital traces of those lost to suicide contain information that could protect others. By analyzing these patterns responsibly and collaboratively, we work toward a future where warning signs are recognized and acted upon. This represents both a technical challenge and a moral imperative. The path forward requires courage, collaboration, and commitment to the belief that suicide prevention can be transformed through data-driven understanding and timely intervention.

STOP SOLDIER SUICIDE | BLACK BOX PROJECT®

The path forward requires collective courage.

As we advance this project, we welcome surviving families, academic researchers, technologists, mental health providers, and veteran service orgnanizations to join us in this continued fight. If you'd like to contribute to this work, please reach out to us at:

BlackBoxProject@StopSoldierSuicide.org

FINDINGS APPENDIX

The following appendices provide comprehensive technical documentation for each finding summarized in the previous section. These detailed analyses represent the methodological foundation of the Black Box Project's discoveries and offer the scientific depth necessary for researchers, clinicians, and technologists to understand, evaluate, and build upon our work. For technical questions about these findings, please email blackboxproject@stopsoldiersuicide.org.

Each appendix follows a standardized structure designed to facilitate both thorough understanding and practical application. The **Key Observations** present the core discoveries without technical jargon, establishing what we found and why it matters. The **Methodology** section documents our analytical approach with sufficient detail for replication, including specific tools, algorithms, parameters, and data processing decisions. This transparency ensures that our findings can be validated and our methods adapted for other populations or contexts.

The **Discussion** sections interpret our findings within the broader context of suicide prevention science, exploring how digital behavioral patterns align with or challenge existing theoretical frameworks. Here we examine not just what the data shows, but what it means for our understanding of suicide risk. The **Limitations** sections acknowledge the constraints that frame our interpretations, from technical challenges in device data extraction to fundamental questions about generalizability across populations. This honest assessment of boundaries strengthens rather than weakens our conclusions by clearly delineating what we can and cannot claim from the current evidence.

The **Implications** sections translate technical findings into actionable insights for different stakeholder communities. Whether addressing clinical practitioners who need practical risk assessment tools, technology developers working on privacy-preserving detection systems, or researchers planning follow-up studies, these sections bridge the gap between data science and real-world application. Finally, the Future Directions outline specific next steps for advancing each line of inquiry, from immediate technical improvements to long-term visions for population-scale implementation.

Together, these appendices demonstrate that the Black Box Project represents more than an exercise in data mining. Each analysis required careful navigation of technical complexity, ethical considerations, and human sensitivity. The digital traces we analyze are not merely data points but the final expressions of individuals who struggled and ultimately lost their battles with suicidal ideation. By documenting our methods with rigor and transparency, we honor both their memory and the trust placed in us by their families, while creating a foundation for prevention efforts that could spare others from similar tragedy.

SUICIDE RISK LEXICON APPLICATIONS IN DIGITAL PHENOTYPING

Key Observations

Our application of Harvard's clinically-validated Suicide Risk Lexicon to posthumous digital data has yielded profound insights into the hidden language of suicide risk. The most striking discovery is the stark divergence between internal and external digital personas: while private digital expressions (notes, searches, web history) showed significant increases in risk language approaching death (Kendall's $\tau=0.3512$, $p=0.0002$), external communications remained remarkably stable. This empirical validation demonstrates that individuals in crisis may maintain facades of normalcy in interpersonal communications while their private digital behavior reveals escalating distress.

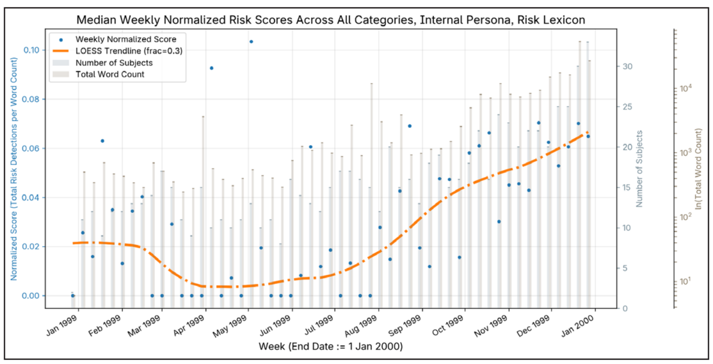

Figure 7: Internal Persona Normalized Risk Scores

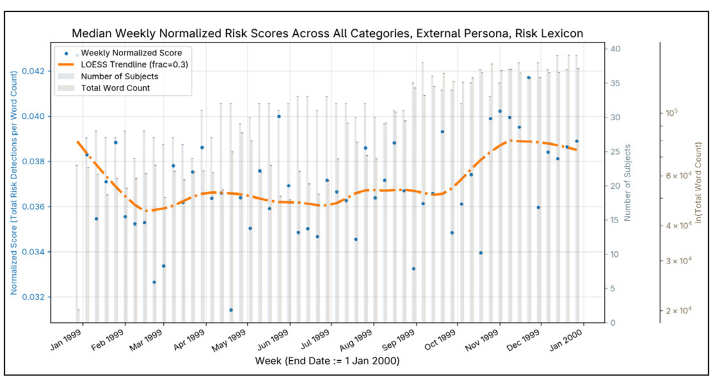

Figure 8: External Persona Normalized Risk Scores

The temporal patterns observed, particularly the "waves" where multiple risk factors activate simultaneously, suggest that suicide risk manifests in discrete episodes rather than linear progression. These waves, clearly visible in individual device heatmaps, create identifiable risk clusters that may represent critical intervention windows. Among the eight risk categories analyzed, anxious symptoms ($\tau=0.3258$, $p<0.0001$) and externalizing behaviors ($\tau=0.3244$, $p=0.0002$) emerged as having the strongest persistent increase over time, indicating that digital traces of agitation and behavioral dysregulation may serve as particularly sensitive early warning signals.

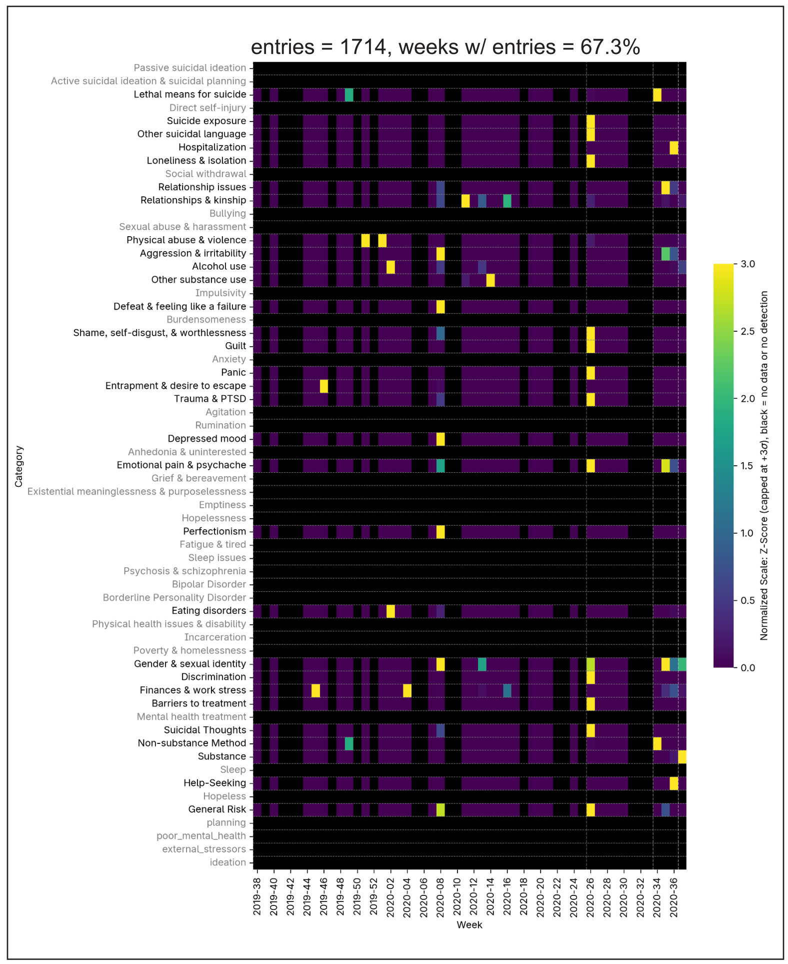

Figure 9: Internal Persona (Singular Device View) – Notice the vertically oriented "waves" of yellow which represent an increase across multiple risk categories during the same week.

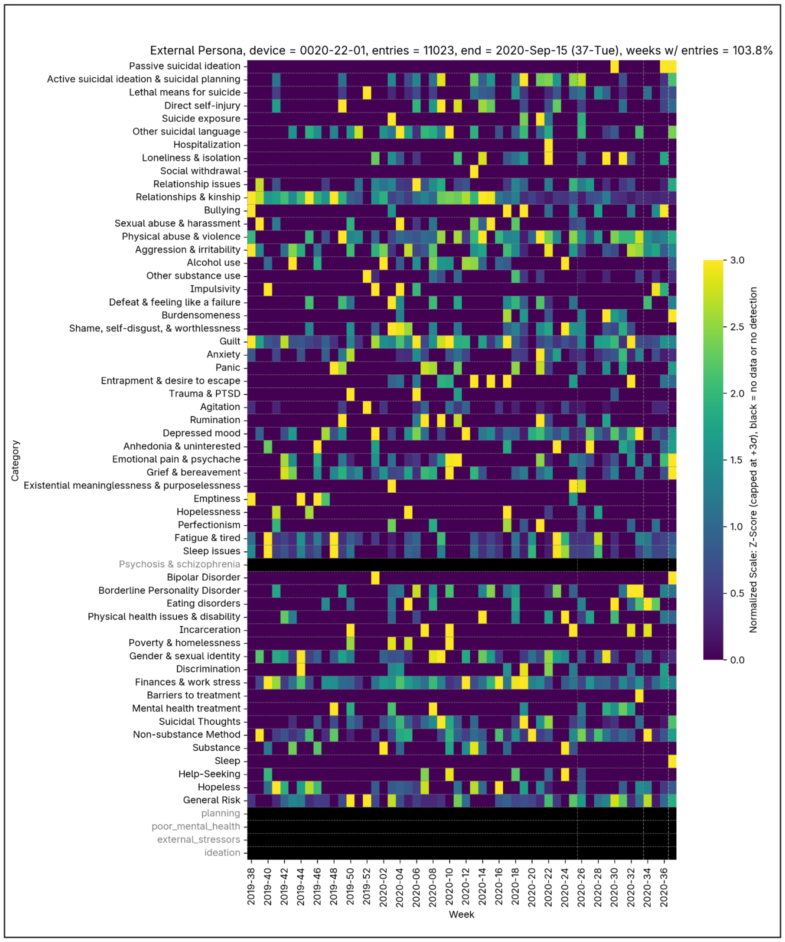

Figure 10: External Persona (Singular Device View).

Methodology

The analysis employed Harvard's open-source Suicide Risk Lexicon, comprising 49 clinically-validated risk factors organized into eight theoretical categories. We applied this lexicon to digital data extracted from 55 available devices, utilizing 54 devices for internal persona analysis and 53 for external persona analysis. The analysis encompassed both device-level heatmap visualization and population-level statistical assessment.

Technical Approach

Heatmap Analysis: We constructed visualizations at the device level by filtering data to the final year of life and normalizing with z-scores (capped at 3σ to handle outliers) per risk lexicon category. Known risk moments were integrated as comparison points. To account for varying message volumes, we divided detection counts for each risk lexicon category by word count at the week-device level before normalization.

Population Analysis: We aggregated weekly risk scores normalized by word count, employing LOESS (Locally Estimated Scatterplot Smoothing) trend analysis, a non-parametric method that makes no assumptions about data shape. Statistical significance was assessed using Mann-Kendall tests for temporal trends. The resulting Kendall's tau statistic is an indicator of the strength and direction of a monotonic relationship (i.e., a consistent increase or decrease of a variable over the entire time frame) and can range from -1 to 1. For category-specific analyses, we utilized Q3 (third quartile) aggregation rather than median values to better capture and highlight emerging trends in the data.

Data Inclusion: Critically, we applied no pre-filtering to the dataset, ensuring representation of devices with both high and low entry volumes. This decision preserves the full spectrum of digital behavior patterns, avoiding selection bias toward more active digital users.

The lexicon's interpretability allowed precise tracking of specific risk factors while maintaining computational efficiency, a critical consideration for potential on-device implementation to protect privacy.

Discussion

The successful translation of Harvard's academic innovation to forensic digital data represents a methodological breakthrough. Unlike traditional studies limited to self-reported data or counselor assessments, our analysis benefits from ground truth outcomes, actual deaths by suicide, transforming probabilistic modeling into definitive pattern recognition.

Internal vs. External Persona Analysis

The data reveals two distinct digital phenotypes. Internal persona markers showed positive trends across all eight risk categories, with seven achieving statistical significance. The most pronounced changes appeared in anxiety-related constructs and externalizing behaviors, aligning with contemporary suicide theory that emphasizes acute anxiety and behavioral dysregulation in the ideation-to-action transition.

Conversely, external persona data demonstrated overall non-significant monotonic trends ($\tau=0.1611$, $p=0.09$), with only depressive symptoms achieving significance ($\tau=0.2017$, $p=0.0336$). This singular exception among external communications suggests that while most interpersonal digital interactions actively mask escalating risk, depressive symptoms may occasionally breach the social facade.

Temporal Dynamics and Wave Patterns

The heatmap analyses revealed striking visual evidence of risk clustering where multiple risk categories often activated simultaneously, creating identifiable "waves" that appear as bright vertical bands spanning multiple risk factors. These episodic patterns challenge linear progression models and suggest that suicide risk may concentrate into acute periods rather than manifesting as a constant state.

Category-Specific Performance

All eight lexicon categories demonstrated positive temporal correlations with approaching death in internal persona data:

- **Anxious symptoms** ($\tau=0.3258$, $p<0.0001$): Strongest consistent increase over time, suggesting changes in digital anxiety markers as a primary indicator of risk
- **Externalizing** ($\tau=0.3244$, $p=0.0002$): Nearly equivalent strength, indicating behavioral dysregulation manifests clearly in private digital behavior
- **Social determinants** ($\tau=0.2932$, $p=0.0019$): Life stressors and contextual factors show clear digital footprints
  Interpersonal ($\tau=0.2903$, $p=0.0008$): Relationship concerns intensify in private expressions
- **Suicidal concepts** ($\tau=0.2511$, $p=0.0004$): Direct risk expressions validate the lexicon's core purpose
- **Depressive symptoms** ($\tau=0.2104$, $p=0.0032$): Traditional risk markers remain relevant but are not primary
  Other disorders ($\tau=0.1909$, $p=0.0112$): Comorbid conditions contribute to overall risk picture
- **Negative perception of self** ($\tau=0.1190$, $p=0.0551$): Trending toward significance, suggesting self-concept deterioration

Limitations

Several constraints frame our interpretation. The analysis represents 53.8% of available devices (54 of 55 for internal, 53 of 55 for external), potentially introducing selection bias toward devices with richer data. The lexicon approach, while interpretable, cannot capture contextual nuances that deep learning models might detect, though this limitation is offset by the guarantee of capturing exact risk markers when present.

Technical limitations include the Mann-Kendall test's constraint of considering only one value per time point, potentially underestimating significance in periods with multiple data points. It also is limited to measuring the strength of monotonic trends over time and does not give us an indicator of moments or waves of statistical divergence in behavior that could indicate an opportunity for intervention. Additionally, the normalization by word count, while standard practice, may obscure risk signals in individuals who communicate less as crisis deepens. We observed technical artifacts such as week-of-year offset issues that occasionally resulted in coverage exceeding 100% for certain time periods.

The analysis is currently limited to specific file types (.xlsx, .json, .csv) and data sources. While comprehensive for available data, this may not capture all forms of digital expression, particularly newer communication platforms or encrypted services.

Implications

This work fundamentally transforms our understanding of digital suicide risk assessment. The validated divergence between internal and external personas suggests that traditional risk assessments which are reliant on what individuals choose to share, may systematically miss those at highest risk. This finding alone could revolutionize screening and intervention approaches.

For Clinical Practice

The lexicon's performance, comparable to deep learning while maintaining interpretability, enables immediate clinical applications. Mental health providers could implement lexicon-based screening tools that respect privacy through on-device processing while providing actionable risk assessments. The identification of anxiety and externalizing behaviors as primary digital markers should refocus clinical attention on these often-overlooked precursors to suicide.

For Technology Development

These findings provide a blueprint for ethical AI implementation in mental health. The lexicon's lightweight computational requirements enable on-device processing, addressing privacy concerns that have limited digital phenotyping adoption. Technology platforms could implement these tools as protective features, identifying risk patterns while maintaining user autonomy and privacy.

For Research Community

The successful validation against ground truth outcomes opens unprecedented opportunities for suicide prevention research. The clear distinction between internal and external digital personas provides a new framework for understanding help-seeking behavior and the barriers to disclosure that prevent timely intervention.

Future Directions

Immediate Next Steps

- Expand data sources by adding outbound emails to individuals (excluding organizations) to the analysis pipeline
- Evaluate instant messages and social media data for potential inclusion in external persona analysis
- Update population visualizations to show median word count per person per week rather than aggregated totals
  Calculate Mann-Kendall significance for additional time periods (6-month and 9-month windows)
- Use statistical tests that can help us detect the timing of significant inflection points in the population trend lines.
- Develop statistical tests for risk wave clustering to quantify the visual patterns observed in heatmaps
- Generate population-level descriptive statistics including: percentage of individuals showing at least one risk wave, average number of waves per person, typical timing between waves, and temporal distribution of first and last waves

Long-term Vision

The successful validation of clinical tools against real-world outcomes creates transformative analytic opportunities:

Personalized Risk Trajectories: Develop machine learning models that learn individual-specific risk patterns while maintaining the lexicon's interpretability. These hybrid approaches could combine the explicability of rule-based systems with the pattern recognition capabilities of neural networks.

Cultural and Demographic Adaptation: Expand lexicon variants for different populations and linguistic contexts, ensuring equitable risk detection across diverse communities. This includes developing culturally-sensitive risk markers and validating performance across demographic groups.

Intervention Efficacy Testing: Investigate whether lexicon-identified risk periods represent optimal intervention moments. By correlating risk waves with intervention timing, we can develop just-in-time adaptive interventions triggered by digital phenotype changes.

Federated Learning Applications: Create privacy-preserving systems where multiple organizations contribute to risk understanding without sharing raw data. This approach could enable population-scale insights while maintaining individual privacy, a critical balance for ethical deployment.

Privacy-Preserving Implementation Protocols: Design comprehensive frameworks for on-device risk detection that process data locally, transmitting only aggregated risk scores or intervention triggers rather than raw content.

Most profoundly, this work demonstrates that the digital traces left by those we've lost can become powerful tools for protecting others. Each pattern identified, each risk marker validated, transforms tragedy into possibility, creating a future where the silent signals of crisis need no longer go unheard.

Key Observations

Analysis of late-night device activity (12-4 AM) across 54 devices revealed distinct temporal patterns in the months preceding death. At the population level, late-night activity increased from a baseline of approximately 30% to 53.16% in the final week of life. The most pronounced shift occurred at approximately 10 weeks before death, where the percentage of individuals with late-night activity began a marked upward trajectory that persisted through the final days.

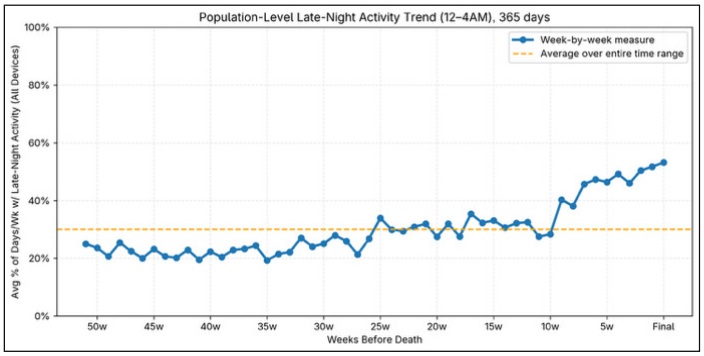

Figure 11: Population level trend for late night activity

Device-level analysis uncovered significant heterogeneity within the population. Three distinct behavioral phenotypes emerged: 33% of devices showed persistent late-night activity throughout the observation period (>70% of nights), 28% demonstrated increased activity primarily in the final weeks, and 39% exhibited minimal late-night engagement (<20% of nights) despite approaching end of life. Sixteen devices maintained complete data coverage across all temporal windows, enabling robust individual trajectory analysis.

Methodology

We implemented a sophisticated approach to ensure temporal accuracy across diverse data sources. The cornerstone innovation involved developing a time zone localization framework that addressed a critical challenge in digital behavioral analysis: accurately determining when activities occurred in the participant's local time.

The methodology began with 6.25 million activity records across 62 devices, utilizing 27 distinct activity tables ranging from communications to web searches. We discovered that device-reported timestamps often reflected UTC offsets that were only accurate if the device's final timezone matched the location where activities occurred—a problematic assumption for individuals who traveled or whose devices defaulted to UTC+00:00.

To address this limitation, we developed a location-based time zone reconstruction approach. By cross-referencing device location data (median device contained 4,050 location records) with the 'timezonefinder' Python package, we created dynamic lookup tables that updated as devices moved across geographic regions. This method required at least three consistent location observations to establish a new time zone, filtering out transient searches while capturing genuine relocations.

The validation results demonstrated the approach's

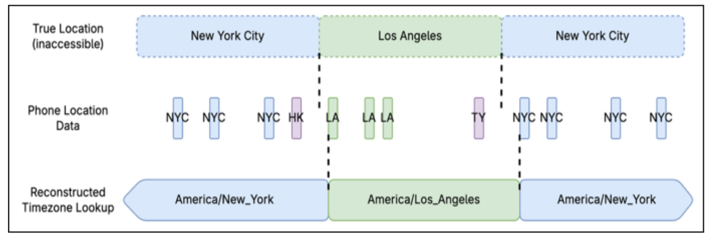

Figure 12: Illustration representing time zone reconstruction

effectiveness: 66% of non-UTC records showed alignment between location-derived and device-provided time zones, while 98% of all records could be accurately localized through either method. This precision enabled confident identification of true late-night activity periods, distinguishing overnight device use from evening activity misclassified due to time zone errors.

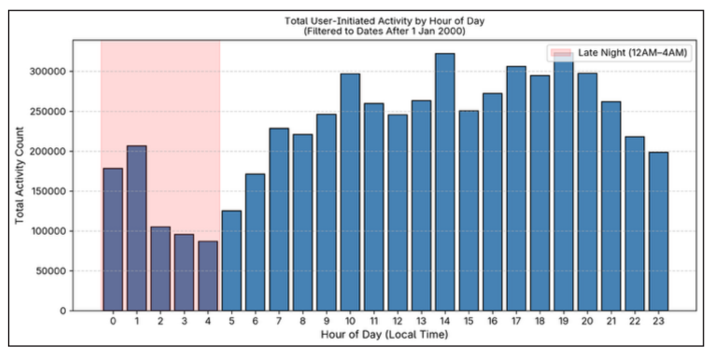

Figure 13: Total user generated activity by hour of day

Discussion

The 10-week inflection point represents a critical finding that aligns with clinical trajectories observed in end-of-life care. This temporal marker suggests a shift in daily rhythms that transcends simple sleep disturbance. The consistency of this pattern across the population, despite significant individual variation, points to underlying physiological or psychological processes that manifest broadly as death approaches. It is also important to note that sophisticated techniques were applied to ensure our analysis only included user-generated activity, excluding device-level activity that could have been driven solely by the device itself such as software updates, wireless and Bluetooth connections, etc.

The heterogeneity observed at the device level provides crucial context for interpreting population-level trends. The identification of three distinct phenotypes suggests that late-night activity serves different functions for different individuals. For persistent users, overnight device engagement may represent an established coping mechanism for chronic conditions. The late-increase group potentially captures acute deterioration or escalating distress, while minimal users may have stronger support systems or different coping strategies that don't involve technology. Further analysis should be conducted in this area.

The technical achievement of accurate time zone localization extends beyond methodological rigor. This innovation enables detection of subtle behavioral shifts that would otherwise be lost to data quality issues. For instance, a person traveling across time zones while experiencing distress might show apparent "late-night" activity that represents normal evening behavior. The ability to distinguish true overnight engagement from artifacts ensures that identified patterns reflect genuine behavioral changes rather than technical limitations.

The progression from 30% to 53% late-night activity represents a 77% relative increase in overnight device engagement. This magnitude of change, occurring over a relatively short timeframe, suggests that digital behavioral markers can detect meaningful transitions in human experience. The gradual nature of the increase, rather than an abrupt shift, indicates that intervention windows may extend across weeks rather than days, offering practical opportunities for support delivery.

Limitations

Several constraints shape the interpretation of these findings. The analysis excluded multiple potentially relevant data sources due to missing time zone offset values, limiting the comprehensiveness of activity capture. Device data continuity and sparsity issues may have influenced the observed patterns, particularly for the dramatic increase seen in the year-level analysis.

The binary classification of late-night activity (presence/absence during 12-4 AM) provides limited granularity about the nature, duration, or intensity of overnight device engagement. Additionally, the requirement for both late-night and daytime activity to count a day as having late-night activity may have excluded isolated overnight sessions that could represent significant distress events.

Individual heterogeneity, while offering opportunities for personalization, also challenges population-level interpretations. Without clinical context about participants' medical conditions, medications, or life circumstances, distinguishing pathological from benign late-night activity remains difficult.

Implications

This analysis demonstrates the potential for digital behavioral markers to complement traditional clinical assessments in identifying at-risk individuals. The 10-week inflection point could serve as a trigger for enhanced monitoring or proactive intervention, particularly when combined with other digital markers such as Known Moments of Risk.

The successful time zone localization methodology establishes a framework for accurate temporal analysis across diverse, real-world datasets—a critical capability for any digital health monitoring system. This technical innovation enables confident detection of behavioral patterns that might otherwise be obscured by data quality issues.

For families who have contributed devices to this research, these findings validate that their loved ones' digital traces contain meaningful signals that could help prevent future tragedies. The ability to detect behavioral changes months before death suggests opportunities for earlier intervention that current healthcare systems might miss.

Future Directions

Immediate enhancements should explore alternative late-night window definitions (such as 12:30-4 AM or sliding windows) to optimize sensitivity and specificity. Statistical validation through regression analysis could establish the significance of observed trends, moving beyond descriptive patterns to predictive insights.

Integration with complementary analyses represents the most promising avenue for clinical translation. Combining late-night activity patterns with Known Moments of Risk, communication frequency changes, and other digital markers could create multi-dimensional risk profiles with enhanced predictive power. Machine learning approaches could identify optimal combinations of markers for different risk scenarios.

The development of normalization methods to address device-level variations would strengthen population-level analyses while preserving the valuable insights from individual behavioral patterns. This might include adjusting for baseline activity levels, device types, or demographic factors that influence technology use patterns.

Ultimately, this work contributes to a larger vision where privacy-preserving analysis of digital footprints enables proactive mental health support. By demonstrating that meaningful behavioral signals exist within routine device usage, this analysis advances toward a future where technology serves as a compassionate early warning system—identifying those in need of support before crises escalate beyond intervention.

Key Observations

The analysis uncovered 106 timestamped moments of elevated suicide risk across 42% of examined devices, revealing that the digital traces of human suffering manifest in identifiable patterns within routine technology use. This finding represents more than statistical significance. It demonstrates that individuals experiencing suicidal ideation leave detectable signals within their digital interactions, offering unprecedented opportunities for understanding and intervention. Of the 59 devices analyzed, 25 contained clear indicators of risk in their notes, web histories, and searched items, with 16 devices (27%) showing evidence of suicidal ideation or active planning behaviors.

These discoveries challenge traditional approaches to suicide prevention by suggesting that the intimate relationship between individuals and their devices creates a new window into mental health crises.

Our analysis of temporal clustering of risk indicators, visualized through overlays with communication patterns, also revealed that distress follows recognizable rhythms. This is knowledge that could transform how we conceptualize intervention timing and support delivery.

Methodology

The analysis team developed a sophisticated dual-method approach that honors both computational power and human judgment. Rather than relying on simplistic keyword searches that would miss the nuanced ways individuals express distress, the methodology employs semantic search using the all-MiniLM-L12-v2 embedding model alongside zero-shot classification via bart-large-mnli. This combination captures meaning beyond literal text, recognizing that expressions of suicidal ideation are deeply personal and varied.

The semantic search transforms text entries into high-dimensional vector representations, enabling comparison based on meaning rather than exact phrasing. The process begins with computing embeddings for each entry, then calculating cosine similarity scores against a carefully curated list of suicide-related queries. Complementing this, zero-shot classification categorizes entries across multiple risk dimensions simultaneously, acknowledging that mental health crises rarely fit neat categories. The system evaluates text against categories including "suicide planning," "suicidal ideation," "external stressors," and "poor mental health," with multi-label capabilities recognizing the complex nature of distress expressions.

Human reviewers then validated computational findings, ensuring that technology augments rather than replaces clinical judgment, a critical balance when dealing with such sensitive human experiences.

Only the personal notes stored on the phone, web history, and searched items were examined in this analysis.

Discussion

Computational Approach to Human Suffering

The successful identification of risk moments across nearly half the sample devices represents a profound validation of the methodology's sensitivity to human distress. By combining embeddings-based semantic search with multi-label classification, the system navigates the complex landscape of how individuals express suicidal thoughts from explicit statements to subtle indicators buried in search histories and personal notes.

The analysis revealed 20 distinct event types, including methods_research, seeking_help, negative_self_talk, relationship_issues, and trauma_fixation, illustrating the multifaceted nature of suicidal crises. This granularity matters: understanding whether someone is researching methods versus reaching out for support fundamentally changes the intervention approach. The system's ability to distinguish these nuances while maintaining scalability offers hope for population-level prevention efforts that don't sacrifice individual sensitivity.

Temporal Patterns and Intervention Windows

The overlay of risk moments with communication

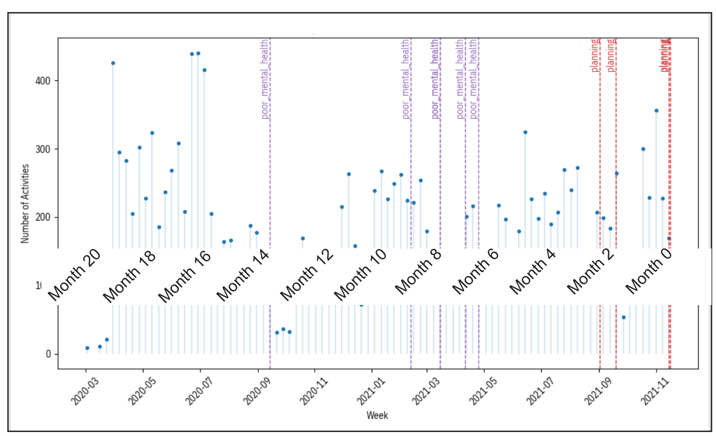

Figure 14: Labeled Known Moments of Risk (KMORs) (Individual Device View)

patterns reveals crucial insights about the temporal dynamics of suicidal crises. These visualizations suggest that risk isn't constant but fluctuates in ways that correlate with digital behavior patterns. Such knowledge transforms our understanding from static risk assessment to dynamic monitoring, opening possibilities for just-in-time interventions that meet individuals in their moments of greatest need.

Ground Truth for Compassionate Technology

Perhaps most significantly, this work establishes a validated ground truth dataset that future predictive models can reference. Unlike theoretical constructs or clinical assessments removed from daily life, these moments represent authentic expressions of distress as they naturally occur. This authenticity is crucial for developing AI systems that can recognize real-world risk patterns rather than idealized clinical presentations.

Limitations

The methodology operates within important constraints that shape its interpretation and application. Privacy-conscious behaviors, particularly incognito browsing for sensitive searches, create blind spots in the data. Ironically, those taking steps to hide their distress may be at highest risk. Additionally, the focus on text-based data misses rich behavioral signals embedded in app usage patterns, location data, and communication metadata.

The current 42% detection rate, while substantial, reminds us that absence of evidence isn't evidence of absence. Many individuals may express distress through channels not captured by this analysis such as physical journals, verbal conversations, or internal struggles that never manifest digitally. Some subjects may use dedicated journaling applications rather than native notes apps, further limiting the scope of detection. These limitations don't diminish the methodology's value but rather contextualize it within the broader challenge of understanding human suffering through technological lenses.

Implications

Transforming Prevention Paradigms

This analysis fundamentally shifts how we conceptualize suicide prevention in the digital age. By demonstrating that meaningful risk indicators exist within routine device usage, it validates the vision of technology-enabled early intervention. The ability to identify 106 genuine risk moments provides concrete evidence that breaking down data silos, the driving mission behind this work, can yield actionable insights for protecting vulnerable populations.

The implications extend beyond technical capabilities to ethical considerations about privacy, consent, and the appropriate use of such intimate data. Success here means not just detecting risk but doing so in ways that preserve dignity and agency, ensuring that surveillance doesn't replace support.

Building Bridges Between Data and Intervention

The validated ground truth dataset creates unprecedented opportunities for developing and testing predictive models against real-world expressions of distress. Future algorithms can now be evaluated not just on statistical metrics but on their ability to recognize authentic human suffering as it manifests in daily digital life. This shifts the conversation from abstract model performance to concrete intervention opportunities where each correctly identified moment represents a potential life saved.

Future Directions

The path forward involves both technical enhancement and thoughtful expansion. The incorporation of advanced embedding models promises improved semantic understanding while maintaining computational efficiency. Implementation of clinician-validated lexicons, particularly the Suicide Risk Lexicon, will enhance domain specificity beyond general categories like "external stressors," bringing clinical expertise directly into the computational framework.

Infrastructure improvements, particularly implementing vector databases like LanceDB for collaborative analysis, will enable researchers to work with this sensitive data at scale while maintaining security. The expansion to additional data streams including social media activity, web bookmarks, cookies, and email communications will create more comprehensive risk profiles. Published suicide risk and crisis lexicons can flag additional scenarios of suicide-specific crises and poor mental health that were not captured in this initial analysis.

Most critically, future work must develop metrics for judging data quality before semantic search processing, better separating devices with insufficient data from those with no detected risk moments. This includes tuning features such as number of inputs and percentage of days with entries in the years prior to death, ensuring that the methodology distinguishes between data absence and risk absence.

Key Observations

Analysis of web browsing behavior surrounding known moments of risk revealed three primary behavioral patterns that warrant further investigation. First, individuals exhibited pronounced cyclical browsing patterns characterized by alternating periods of engagement with distressing content and subsequent help-seeking behaviors. This oscillation manifested as transitions between searches for harmful methods and visits to mental health resources, religious organizations, or support forums, suggesting internal conflict and potential intervention opportunities.

Second, the temporal analysis demonstrated rapid behavioral escalation in critical cases. Among devices where planning-related moments were identified in the final weeks of life, explicit searches for suicide methods often occurred within days of death. These searches exhibited high specificity and intentionality, with query refinement patterns indicating methodical information gathering rather than impulsive behavior.

Third, known moments of risk were frequently preceded by a constellation of digital behaviors including relationship-related searches, increased consumption of adult content, and military or veteran-related stressors. These factors often co-occurred and appeared to contribute to periods of emotional destabilization, particularly when combined with searches related to financial distress or housing instability.

Methodology

The analysis employed a multi-stage approach to identify and analyze digital behaviors surrounding known moments of risk. The Known Moments of Risk utility enabled systematic cross-referencing of risk episodes with corresponding web activity across temporal windows.

Risk Moment Identification

Digital artifacts including notes, search queries, and web history were analyzed using semantic search with predefined suicide-related terminology. Cosine similarity metrics were applied with empirically-determined thresholds to surface potentially relevant content. Results exceeding similarity thresholds were aggregated at the device level and underwent manual review to produce a high-confidence dataset. Each validated moment was categorized into four types: ideation, planning, external stressors, or poor mental health indicators.

Temporal Analysis Framework

Symmetric time windows were constructed around each validated risk moment timestamp, with configurable parameters including ±7, ±14, ±21, and ±30 day windows. This approach enabled examination of behavioral patterns both preceding and following identified risk events. A configurable utility was developed to support temporal extraction, allowing flexible analysis as device volume and data complexity increased.

Topic Modeling Implementation

Latent Dirichlet Allocation (LDA) was employed to extract thematic patterns from search queries during risk windows. The implementation utilized gensim, spacy, and nltk Python libraries with the following parameters:

- Number of topics: 5-10 (dynamically determined based on query volume)
- Alpha: Auto-optimized for topic distribution
- Beta: Auto-optimized for word distribution
- Iterations: 100 passes through the corpus
- Random state: Fixed for reproducibility

Topic coherence scores were calculated using the C_v metric to assess the semantic interpretability of discovered topics, with scores ranging from 0.0257 to 0.7566 across analyzed devices.

Discussion

The analysis revealed nuanced patterns in digital behavior that align with theoretical models of suicide risk. The cyclical nature of browsing behavior supports the ambivalence theory of suicide, where individuals experience competing desires for life and death. This digital manifestation of psychological ambivalence presents unique opportunities for intervention, particularly during help-seeking phases of the cycle.

The rapid escalation observed in planning behaviors challenges assumptions about the timeline of suicide risk. While traditional models often emphasize prolonged ideation periods, the digital footprint suggests that the transition from ideation to active planning can occur within days or even hours. This compressed timeline has significant implications for intervention strategies and risk assessment protocols.

The co-occurrence of relationship stressors, adult content consumption, and military-related searches suggests complex interactions between emotional regulation, coping mechanisms, and environmental stressors. The increased consumption of adult content may represent maladaptive coping strategies during periods of emotional distress, while military-related searches often coincided with searches about benefits, healthcare, or transition challenges.

Limitations

Several limitations constrain the interpretability and generalizability of these findings. Web browsing history demonstrated significant variability across devices, with some showing rich, continuous histories while others exhibited sparse data or substantial gaps. This inconsistency likely resulted from factors including intentional deletion, privacy settings, browser configurations, and device usage patterns.

The analysis was limited to 25 devices with sufficient browsing data from an available pool of 40 devices, potentially introducing selection bias toward individuals with more extensive digital footprints. Additionally, the semantic search approach relied on predefined terminology that may not capture evolving language patterns or culturally specific expressions of distress.

The LDA topic modeling, while revealing thematic patterns, showed varying coherence scores across devices, with some producing highly interpretable topics while others yielded less meaningful results. This variability may reflect differences in search behavior complexity or vocabulary diversity among individuals.

Implications

These findings have profound implications for suicide prevention efforts. The identification of cyclical browsing patterns suggests that continuous monitoring systems could detect risk escalation in real-time, potentially enabling timely interventions during help-seeking phases. The rapid escalation timeline emphasizes the need for immediate response capabilities when planning behaviors are detected.

The strong association between specific digital behaviors and subsequent risk events supports the development of predictive models that could identify individuals at elevated risk before crisis onset. However, such systems must balance sensitivity with specificity to avoid overwhelming intervention resources with false positives.

The prevalence of military-related stressors in the dataset underscores the need for targeted digital interventions for veteran populations, potentially including specialized resources accessible through commonly visited military and veteran websites.

Future Directions

Several avenues for future analyses emerge from this study. Expanding the semantic search vocabulary to include emerging terminology and culturally specific expressions could improve risk moment identification. Integration with other communication modalities, including messaging platforms and email, would provide a more comprehensive view of digital behavior during risk periods.

Advanced machine learning approaches, including transformer-based models for semantic analysis and graph neural networks for behavioral pattern recognition, could enhance both the sensitivity and specificity of risk detection. Real-time analysis capabilities would enable dynamic risk assessment and immediate intervention triggering.

Finally, the development of privacy-preserving federated learning systems could enable analysis across larger populations while maintaining individual privacy, ultimately contributing to more robust and generalizable models of digital behavior during suicide risk periods. This approach aligns with the broader vision of leveraging technology to understand and prevent suicide while respecting the dignity and privacy of those who have been lost.

Key Observations

The Phase 1 Social Network Analysis revealed significant temporal patterns in communication behaviors during the final weeks of life across analyzed devices. Analysis of call log data from 34 devices and chat data from 50 devices demonstrated measurable changes in both the frequency and nature of social interactions as individuals approached death.

Communication volume exhibited bidirectional patterns. While some individuals showed sharp increases in call and chat activity during the weeks preceding death, potentially reflecting help-seeking behaviors or preparatory activities, others demonstrated marked decreases in communication frequency and diversity of contacts. These decreases often manifested as reduced interactions with peripheral contacts while maintaining connections with core relationships.

Time-of-day communication patterns revealed systematic deviations from established baselines at the individual level. At multiple thresholds before death, individuals frequently departed from their typical communication routines. Business hour calls might shift dramatically toward late night activity, or evening communications might transition to early morning exchanges. Remarkably, within the final week of life, these disrupted patterns often reverted to approximate baseline behaviors, suggesting a potential return to familiar routines.

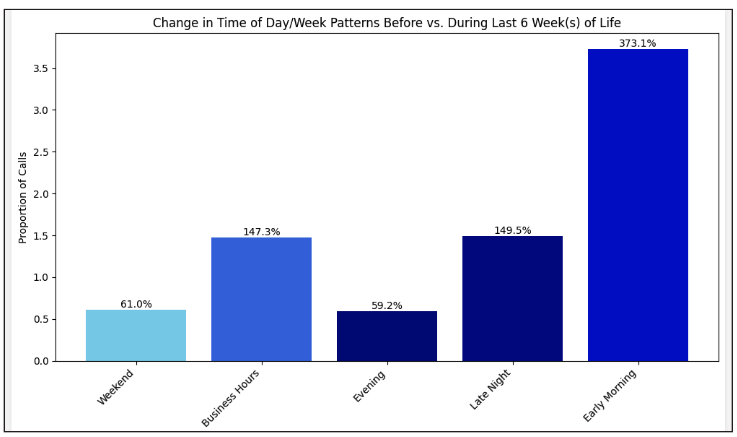

Figure 15: Change in Communication Patterns at the 6-week threshold before death (Individual Device View)

Methodology

The analysis employed a systematic approach to reconstruct egocentric social networks from multiple communication modalities. For each device, we identified the ego (device owner) and their alters (communication partners) through comprehensive data extraction and processing.

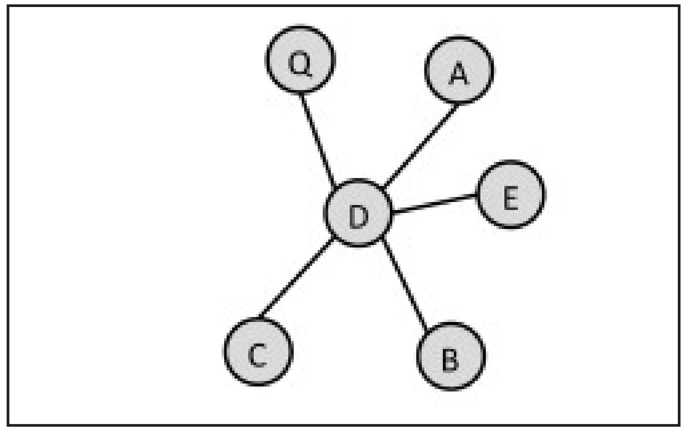

Figure 16: Illustration depicting Ego ("D") at center of Alters

Alter Ranking Algorithm: A weighted composite score determined the relative importance of contacts within call logs. The algorithm incorporated three key metrics: total call volume (weight = 2.0), average call duration (weight = 1.0), and temporal spread of interactions (weight = 0.5). This approach captured both the intensity and persistence of relationships over time.

Exclusion Criteria: Rigorous data quality standards were applied. For call log analysis, devices required a minimum of 100 calls, temporal coverage spanning at least two weeks, activity within the final year of life, and a verified date of death. Chat data underwent similar filtering, with additional criteria addressing sender/recipient identification, timestamp completeness, and group chat attribution.

Feature Engineering: The analysis generated comprehensive metrics at both ego and alter levels. These included communication directionality ratios, time-of-day distributions (categorized as business hours, evenings, weekends, late night, and early morning), burst detection algorithms, and temporal stability indices. Visualization outputs included time-series graphs of communication volume, comparative bar charts for pre/post threshold periods, and call disposition analyses using Wilson score intervals.

Data Processing: Chat analysis required specialized handling due to platform heterogeneity. Missing sender/recipient fields were imputed when possible using participant lists, and a distinction was maintained between direct and indirect incoming messages to prevent inflation of engagement metrics in group contexts.

Discussion

The observed patterns suggest complex psychosocial dynamics in end-of-life communication behaviors. The bidirectional nature of activity changes, both increases and decreases, indicates that no single behavioral trajectory characterizes the final weeks of life. Instead, individual responses appear highly contextualized, potentially reflecting factors such as health status, social support structures, and personal coping mechanisms.

The six-week inflection point for time-of-day pattern disruption represents a particularly intriguing finding. This temporal marker may reflect the onset of significant life changes whether medical, psychological, or social that disrupt established routines. The subsequent reversion to baseline patterns in the final week could indicate either a return to familiar comfort zones or the constraints of declining capacity limiting communication to established patterns.

Social isolation manifested through multiple indicators beyond simple volume metrics. Increased ratios of missed incoming calls, decreased response rates to communications, and narrowing of active contact networks all suggested progressive social withdrawal. Notably, several individual-level analyses revealed spikes in unreturned calls from specific contacts in the final days, potentially indicating network members who sensed risk but were unable to establish contact.

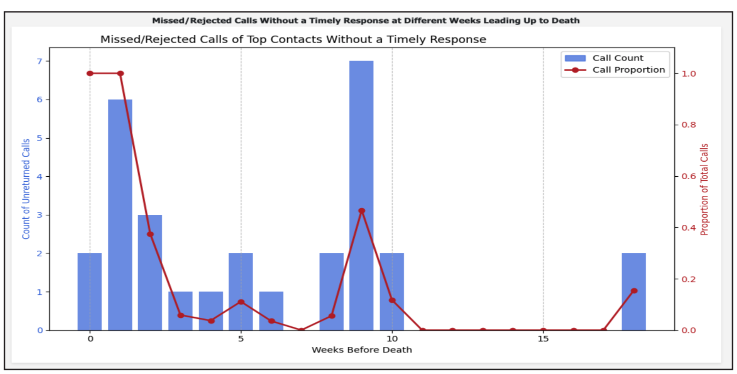

Figure 17: Missed/Rejected calls WITHOUT a timely response (Individual Device View)

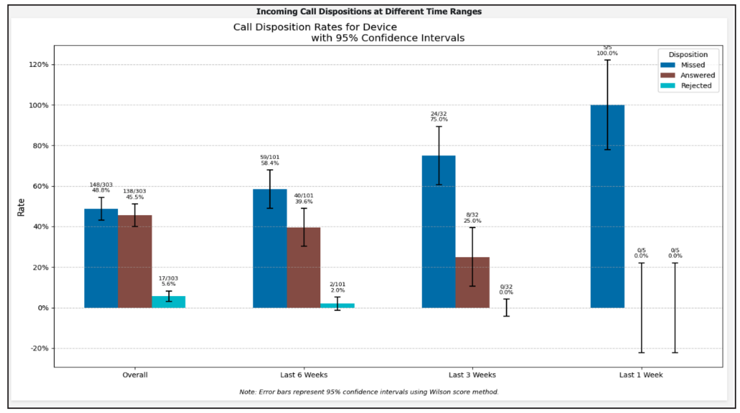

Figure 18: Call Disposition Rates (Individual Device View)

Limitations

Several methodological constraints warrant consideration. The alter ranking algorithm, while capturing multiple relationship dimensions, employed weights that may not optimally represent relationship importance across all contexts. Future iterations could benefit from machine learning approaches to weight optimization.

Data completeness varied significantly across devices and platforms. The exclusion criteria, while necessary for analytical rigor, resulted in the elimination of devices that might have provided valuable edge-case insights. Additionally, the current analysis treated all communication platforms equally, despite potential differences in their social significance and usage patterns.

Group chat attribution remains challenging. The Phase 1 analysis excluded group conversations with more than three participants, potentially overlooking important social dynamics in larger group contexts. Time zone inconsistencies in some datasets prevented complete time-of-day analysis for a subset of devices.

Implications

These findings carry significant implications for suicide prevention efforts. The identification of measurable communication pattern changes weeks before death suggests potential opportunities for algorithmic risk detection. The six-week threshold for behavioral disruption could inform the timing of intervention strategies.

The heterogeneity of observed patterns underscores the importance of personalized approaches to risk assessment. While some individuals may benefit from interventions triggered by increased communication activity, others might require monitoring for social withdrawal indicators. The reversion to baseline patterns across some devices, in the final week particularly, highlights the complexity of end-of-life behaviors and the need for multi-faceted assessment approaches.

Future Directions

Phase 2 will address current limitations through unified contact resolution across communication platforms, enabling true multi-modal social network reconstruction. Advanced attribution algorithms for group chat participation will expand the analysis scope to include larger group dynamics.

Planned enhancements include machine learning approaches for alter importance scoring, incorporation of communication content analysis to complement structural metrics, and development of real-time risk scoring algorithms based on deviation from personal baselines. Integration with external data sources may provide additional context for observed behavioral changes.

The ultimate goal remains the development of privacy-preserving, federated computing approaches that could enable population-scale analysis while maintaining individual privacy. Such systems could transform our understanding of end-of-life social dynamics and inform more effective, timely interventions in suicide prevention efforts.

Key Observations

Our analysis of digital communications revealed distinct patterns of financial distress that intensified during the final year of life. Across the 53 devices analyzed, we observed a marked increase in financial-themed communications, with semantic distress scores rising from approximately 0.32 to 0.35 over the final month, a statistically significant elevation that persisted alongside stable device-level participation. Most notably, this pattern diverged from general communication volume trends, suggesting that financial distress manifests as a distinct behavioral signal rather than merely reflecting overall changes in digital engagement.

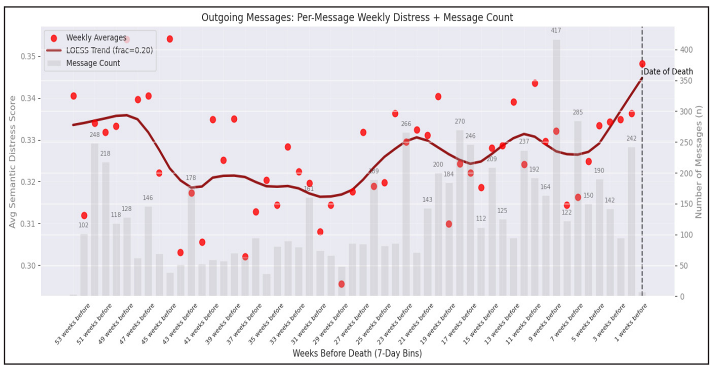

Figure 19: Per Message financial distress score (population level)

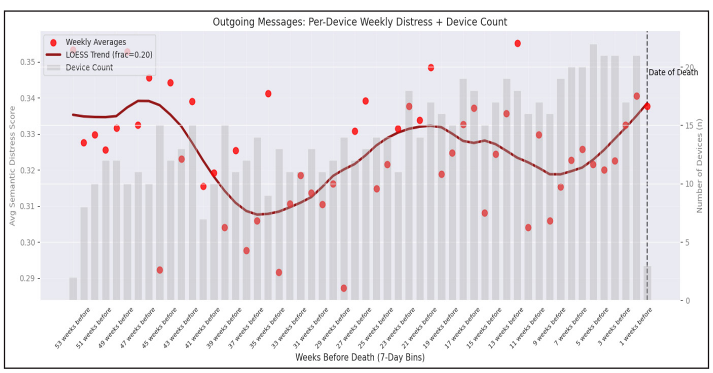

Figure 20: Per Device financial distress score (population level)

The emergence of financial concerns followed a troubling trajectory. LOESS smoothing analysis revealed that while financial distress fluctuated throughout the year, the sharp rise in the final weeks occurred alongside more than 20 active devices continuing to transmit messages. This sustained engagement amid escalating financial anxiety may represent a meaningful behavioral indicator of acute risk, particularly when combined with the abrupt cessation of distress communications that could signal behavioral disengagement or loss of perceived efficacy in outreach.

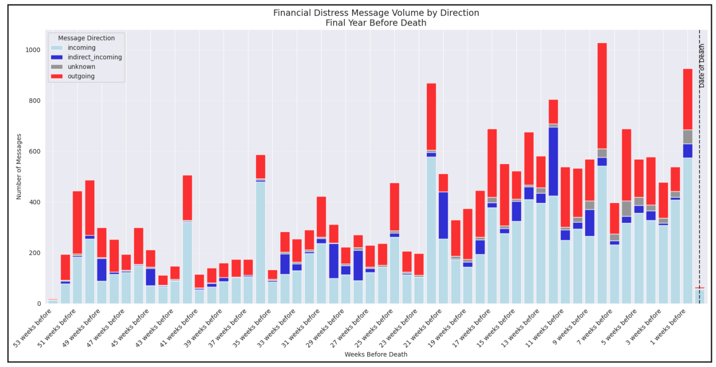

Figure 21: Population level financial distress message volume by direction (incoming, outgoing, indirect incoming, unknown)

Methodology

To identify and quantify financial distress patterns, we employed a multi-layered natural language processing approach on communications from 53 devices selected from an available pool of 62 based on data quality criteria. We first applied Latent Dirichlet Allocation (LDA) topic modeling to identify emergent communication themes across the entire corpus, which revealed eight distinct topics including a prominent cluster of financial concerns encompassing daily money management, long-term financial planning, and acute monetary distress.

For semantic analysis, we developed curated exemplar messages representing financial distress and calculated cosine similarity scores between these exemplars and all messages in the dataset. The similarity threshold of 0.20 was established through manual review of message samples to balance inclusivity with specificity, ensuring capture of genuinely distressed content while excluding tangentially related messages. Each message was embedded using contemporary transformer-based approaches and compared against our exemplar set, with distress severity scores ranging from 0 to 1. To capture temporal patterns, we applied LOESS smoothing with a 0.20 bandwidth parameter to the weekly aggregated semantic scores, providing a smoothed trajectory that reduced noise while preserving meaningful fluctuations in distress levels.

Discussion

The financial distress analysis uncovered a complex interplay between economic concerns and communication behavior in the months preceding death. The Topic 5 cluster, which emerged as the most prevalent financial theme, encompassed a broad spectrum of monetary anxieties ranging from immediate needs to structured financial planning elements. This diversity suggests that financial distress manifests not as a monolithic concern but as a multifaceted stressor affecting various life domains simultaneously.

The temporal evolution of financial distress scores revealed several critical insights. The baseline fluctuation between 0.30 and 0.34 throughout most of the final year suggests a persistent undercurrent of financial concern that may represent normative stress levels. However, the acceleration to 0.35 in the final weeks, while numerically modest, occurred against a backdrop of otherwise stable patterns and maintained device participation. This elevation gains significance when considered alongside the message volume data, which showed institutional communications representing an increasing proportion of total messages over time.

Particularly noteworthy was the divergence between overall message volume and financial distress patterns. While general communication showed expected variability with some decline in certain periods, financial distress communications remained remarkably consistent before their final surge. This persistence suggests that financial concerns may serve as a reliable signal precisely because they resist the communication fatigue or withdrawal patterns often seen in other domains. The sharp cessation of financial distress signals in the final period, despite continued device activity, may indicate a critical transition point where individuals either resolve their financial concerns through alternative means or experience a form of behavioral shutdown regarding financial help-seeking.

The co-occurrence of rising financial distress with stable device counts in the final month provides important context for interpretation. Unlike scenarios where increased distress might simply reflect a selection bias of more distressed individuals remaining active, our data shows consistent participation levels even as distress intensifies. This pattern strengthens the interpretation that the observed changes reflect genuine escalation in financial anxiety rather than artifacts of differential attrition.

Limitations

This analysis faces several important constraints that should guide interpretation of findings. The semantic similarity threshold of 0.20, while carefully selected through manual review, necessarily involves subjective judgment that may not capture all expressions of financial distress, particularly those using indirect language or cultural-specific idioms. Additionally, our focus on English-language communications and specific semantic patterns may underrepresent financial distress expressed through other linguistic or behavioral channels.

The limited sample size of 53 devices, while substantial for this preliminary investigation, constrains our ability to examine subgroup differences or establish robust population-level estimates. The absence of demographic information, maintained to protect privacy, prevents analysis of how financial distress patterns might vary across age, geographic, or socioeconomic contexts. Furthermore, our reliance on digital communications captures only one channel of distress expression, potentially missing financial concerns expressed through voice calls, in-person conversations, or non-verbal behaviors.

Implications

These findings suggest that monitoring financial distress patterns in digital communications could provide valuable early warning signals for suicide risk assessment. The persistent elevation of financial concerns throughout the final year, culminating in the sharp increase during the last month, indicates that financial distress may serve as a more stable and reliable indicator than previously recognized. Unlike many risk factors that fluctuate dramatically or appear only immediately before a crisis, financial concerns appear to provide a sustained signal that could enable earlier intervention.

For clinical practice, these patterns suggest the importance of routine financial distress screening and the potential value of integrating financial counseling or resource connection into comprehensive suicide prevention strategies. The finding that financial communications persist even as other communication types may decline highlights the opportunity for engagement through financial assistance programs or economic support services as a pathway to reach at-risk individuals who might otherwise withdraw from traditional mental health outreach.

The distinct trajectory of financial distress also implies that current risk assessment tools may benefit from more nuanced incorporation of economic stressors. Rather than treating financial problems as a binary risk factor, our findings suggest value in monitoring the trajectory and intensity of financial concerns over time, with particular attention to acceleration patterns that may signal escalating risk.

Future Directions

Immediate next steps should focus on validating these patterns across larger and more diverse samples, including examination of how financial distress manifests across different demographic groups and geographic regions. Development of real-time monitoring systems that can detect concerning trajectories in financial distress communications could enable proactive intervention before crises escalate. Integration with other behavioral indicators, such as sleep patterns, social engagement, and help-seeking behaviors, may reveal multimodal signatures of escalating risk.

Longer-term research should investigate the causal relationships between financial distress and suicide risk, potentially through natural experiments examining the impact of financial interventions on subsequent outcomes. Collaboration with financial institutions, social services, and community organizations could create comprehensive support networks that address both the economic and psychological dimensions of distress. Additionally, development of privacy-preserving technologies that enable broader analysis while maintaining individual confidentiality will be crucial for scaling these insights to population-level prevention efforts. The ultimate goal remains creating predictive models that can identify windows of opportunity for intervention, transforming our understanding of financial distress from a static risk factor into a dynamic signal that guides timely, effective support.

Key Observations

Analysis of digital footprints from Black Box Project participants revealed remarkably stable online behavior patterns, even as individuals approached their final days. The most striking finding was the consistency in core browsing habits: eight out of ten most-visited websites remained unchanged between the period before and during the final month of life. This stability suggests that fundamental online behaviors may persist despite the psychological turmoil often associated with suicidal ideation. However, beneath this surface consistency, we observed critical behavioral shifts: financial and insurance websites showed marked increases in activity during the final month, with USAA displaying notable rises in engagement while Intuit and CreditKarma entered the top 25 sites for the first time. These patterns may indicate end-of-life planning behaviors that warrant further investigation.

Community engagement patterns revealed selective but meaningful participation in support-oriented spaces. Among Reddit activity, r/SuicideBereavement ranked seventh overall with 38 visits, suggesting that some individuals seek connection with others who have experienced similar losses. Facebook group engagement remained notably low compared to other online activities, with the most-visited group receiving only 11 visits, indicating that group-based social media interactions may not be primary channels for those at risk.

Methodology

Our analytical approach employed a comprehensive examination of browser cookies and web history data from 59 available devices, focusing on identifying the most relevant websites, online communities, and user interests. To ensure data quality and meaningful insights, we implemented several preprocessing steps. Known advertising affiliates and tracking domains were filtered out to focus on content sites indicative of user intent rather than passive tracking activity. For temporal analysis, we calculated time deltas between browsing activity and each participant's date of death, segmenting behavior into crisis windows: the last 7 days, last 30 days (imminent risk window), and last 60 days (extended risk window).

We employed a sophisticated ranking system for both cookies and web history analysis. For each device, we extracted the top 100 domains based on cookie count or visit frequency, applying a reverse-rank weighting system where the highest-ranked domain received 100 points, decreasing sequentially to 1 point for the 100th-ranked domain. This approach balanced frequency with relative importance across devices. To address data variability, we deduplicated records at specified time intervals: cookies at 1-minute intervals and web history at 2-minute intervals, and implemented statistical outlier capping based on established conventions for distinguishing programmatic from human behavior.

Discussion

The persistence of core browsing patterns throughout the final month challenges common assumptions about dramatic behavioral changes preceding suicide. The stability in eight of ten top websites suggests that individuals maintain their established digital routines even while potentially experiencing significant psychological distress. This finding has profound implications for detection algorithms that may rely too heavily on sudden behavioral shifts as warning signs. Instead, our analysis points to more subtle indicators embedded within otherwise normal-appearing browsing patterns.

The emergence of financial planning websites in the final month represents a potentially actionable signal. The increased activity on USAA, coupled with the new appearance of Intuit and CreditKarma in the top 25, may indicate practical preparations that individuals undertake when contemplating end-of-life decisions. This pattern aligns with clinical observations about the period of apparent calm and purposeful activity that sometimes precedes suicide attempts. Notably, a transportation service (Uber) showed dramatic increases in the final month, with 81% of all visits occurring during this period across nine devices, potentially reflecting increased mobility needs or final visits.

Social media and community engagement patterns revealed nuanced behaviors that merit careful interpretation. While overall activity on platforms like Facebook remained consistent, the specific communities visited tell a more complex story. The prominence of r/SuicideBereavement suggests some individuals may be processing grief or seeking understanding about suicide's impact on survivors. The relatively low Facebook group activity, contrasted with sustained Reddit engagement, indicates platform preferences that could inform outreach strategies.

The analysis also revealed interesting patterns in content consumption categories. Politics, gaming, and general interest dominated Reddit activity, suggesting that individuals maintain diverse interests even while potentially struggling with suicidal ideation. This finding reinforces the importance of meeting at-risk individuals within their existing online spaces rather than expecting them to seek out dedicated mental health resources.

Limitations

Several methodological constraints must be acknowledged when interpreting these findings. Browser-based data collection cannot capture app-based activity, potentially underrepresenting mobile-first platforms and their associated behaviors. Some participants may have used private browsing modes or cleared their histories, resulting in incomplete behavioral profiles. Additionally, cookie counts do not reliably indicate visit frequency, as some sites deploy numerous cookies in single sessions while others rely on third-party trackers that inflate counts independently of actual user engagement.

The current analysis lacks a control group of non-suicidal veterans, limiting our ability to definitively attribute observed patterns to suicide risk rather than general veteran online behavior. Time-based comparisons within individuals provide valuable insights but cannot fully substitute for between-group analyses. Furthermore, the relatively small number of devices with complete data for specific time windows may limit the generalizability of temporal patterns, particularly for the final week of life.

Implications

These findings suggest that effective digital suicide prevention strategies must evolve beyond simple anomaly detection to incorporate subtle behavioral indicators within sustained patterns. The stability of core browsing habits indicates that at-risk individuals remain reachable through their regular online channels, challenging approaches that rely solely on crisis-specific platforms. Organizations developing targeted interventions should consider embedding resources within mainstream websites rather than expecting individuals to seek out specialized mental health destinations.

The increased engagement with financial planning sites during the final month presents a critical opportunity for partnership-based interventions. Financial institutions, particularly those serving military communities like USAA and Navy Federal, could play a pivotal role in suicide prevention by training staff to recognize concerning patterns and providing appropriate resources. The temporal clustering of these visits suggests a potential window for intervention if appropriate privacy-preserving detection mechanisms can be developed.

The platform-specific engagement patterns observed high Reddit activity versus low Facebook group participation have immediate implications for resource allocation in digital outreach efforts. The prominence of r/SuicideBereavement and veteran-focused communities indicates that peer support and shared experience remain valuable even in digital spaces. However, the selective nature of this engagement suggests that one-size-fits-all social media strategies may miss significant portions of the at-risk population. Instead, platform-specific approaches that respect the unique culture and engagement patterns of each digital space may prove more effective.

Future Directions

Immediate analytical priorities should include developing behavioral baselines through comparison with control populations of veterans without suicide outcomes. This foundational work would enable more definitive attribution of observed patterns to suicide risk rather than general demographic characteristics. Additionally, investigating the specific content and context of financial site visits could reveal whether these represent practical end-of-life planning, financial distress, or other motivations entirely.

Longer-term research directions should explore the integration of multiple behavioral signals into comprehensive risk assessment models. While individual indicators like increased financial site visits may have limited predictive value in isolation, combinations of subtle behavioral shifts could provide more robust detection capabilities. The development of privacy-preserving federated learning approaches could enable model training across multiple data sources without compromising individual privacy. Furthermore, partnership development with key platforms and services identified in this analysis, particularly financial institutions and community platforms, could create novel intervention pathways that meet individuals within their existing digital routines rather than requiring them to seek help through unfamiliar channels.

INCONCLUSIVE WORK

Scientific integrity demands that we present not only our successes but also the hypotheses that failed to yield meaningful results. This section documents two analytical approaches that, despite sound theoretical foundations and rigorous implementation, revealed no significant associations with suicide risk in our dataset. These null findings represent essential contributions to the field, delineating which digital behavioral patterns warrant continued investigation and which may be less fruitful avenues for suicide prevention efforts.

The analyses presented here, digital interest profiling through email communications and physical activity patterns from health tracking data, consumed considerable resources and employed sophisticated computational methods. Our team developed novel analytical pipelines, implemented state-of-the-art machine learning algorithms, and carefully processed thousands of data points searching for signals that ultimately were not present. Rather than viewing these efforts as failures, we recognize them as critical components of the scientific process that narrow the search space for future researchers and prevent duplication of effort across the field.

Each null finding emerged from a reasonable hypothesis grounded in existing literature. Email interest patterns could theoretically reveal shifts in focus or engagement that precede suicidal crises. Physical activity changes might manifest as behavioral markers of depression or agitation. The absence of clear signals in these domains does not diminish their theoretical plausibility but instead highlights the complexity of suicide risk and the importance of empirical validation for all assumptions.

The documentation that follows maintains the same rigorous structure as our positive findings, demonstrating that null results deserve equal analytical care. By thoroughly describing our methods, we enable other researchers to build upon our work, perhaps with larger samples, different populations, or modified approaches that might reveal patterns invisible in our current analysis. The limitations sections prove particularly valuable for null findings, as they illuminate potential reasons why expected patterns failed to emerge and suggest conditions under which future analyses might succeed.

These inconclusive results underscore a fundamental truth about the Black Box Project's approach: we follow the data wherever it leads, not where we hope it might go. Behind every finding presented in this white paper lie numerous additional analyses, explorations, and hypotheses that helped refine our understanding of which digital traces carry meaningful information about suicide risk. By sharing both our successes and our instructive failures, we contribute to a more complete scientific understanding of digital phenotyping in suicide prevention.

The implications drawn from null findings often prove as valuable as those from positive results. They guide resource allocation, prevent investment in unproductive approaches, and challenge assumptions that might otherwise persist without empirical scrutiny. In presenting these analyses, we demonstrate that the path to breakthrough discoveries requires the courage to acknowledge when promising avenues lead to dead ends, and the wisdom to share these lessons with the broader research community.

Key Observations

Analysis of email communications from decedent devices revealed no meaningful patterns linking digital interest profiles to suicidal outcomes. While the zero-shot classification approach successfully categorized email content into twelve consumer interest categories, the resulting patterns failed to demonstrate predictive value or significant temporal changes approaching death. Shopping & E-commerce dominated as the most prevalent category across all devices (average probability 0.40), followed by Technology & Electronics (0.26), but these distributions appeared consistent with general consumer behavior rather than crisis-specific indicators.

The temporal analysis, examining interest evolution across multiple time periods leading to death, showed only minor fluctuations that fell within expected variance for typical email communication patterns. Although Financial Services demonstrated a 25% increase above baseline in the final week and Technology & Electronics declined to 70% of baseline, these changes lacked the magnitude or consistency necessary to serve as reliable behavioral markers.

Methodology

The analysis employed natural language processing techniques to extract and classify interest patterns from email communications across all available time periods. Device end dates were integrated to enable longitudinal analysis, creating defined observation windows at 6+ months, 3 months, 1 month, and 1 week before death. A multi-stage processing pipeline ensured data quality and classification accuracy.

Initial data preparation involved implementing spam detection using the DistilBERT multilingual model fine-tuned for email classification, removing non-human communications that could skew interest profiles. Financial purchase communications were extracted through a rule-based approach utilizing regular expression patterns across three tiers: primary keywords indicating strong purchase intent, secondary contextual terms, and negative exclusion patterns to minimize false positives. This approach captured transactional emails while filtering newsletters and promotional content.

The core classification leveraged Facebook's BART-large-MNLI model for zero-shot classification, enabling categorization without training data specific to this domain. Eight candidate labels spanning consumer categories (Technology & Electronics, Shopping & E-commerce, Entertainment & Media, Health & Wellness, Financial Services) and specialized categories (Military & Tactical, Firearms & Ammunition, Dating & Relationship Services) were applied to each email. The implementation utilized Ray distributed computing framework to process the large volume of communications efficiently, generating probability scores for each interest category per message.

Discussion

The analysis produced several technical insights despite the absence of meaningful suicide-related patterns. The successful implementation of zero-shot classification demonstrated that consumer interest profiling from email content is technically feasible, with clear differentiation between interest categories and consistent probability distributions across the dataset. The clustering analysis identified four distinct user personas based on engagement patterns: Tech-Entertainment Enthusiasts, Shopping-focused users, Minimal Engagement profiles, and Finance-Focused individuals.

The temporal analysis revealed that while individual interest categories showed some variation over time, these changes aligned more closely with typical consumer behavior cycles than crisis indicators. The U-shaped pattern in Shopping & E-commerce engagement and the gradual increase in Financial Services interactions could reflect seasonal patterns, life events, or random variation rather than deteriorating mental health. The convergence of interest categories in the final months, initially hypothesized as potential focus narrowing during crisis, appeared insufficient to distinguish from normal behavioral variance.

The device-level analysis confirmed substantial heterogeneity in digital engagement patterns across individuals. This diversity, while validating the classification approach's ability to capture individual differences, simultaneously highlighted the challenge of identifying universal markers. Categories expected to potentially correlate with crisis states, such as Firearms & Ammunition or specific health-related content, showed minimal presence in email communications, suggesting either low engagement with these topics via email or successful filtering by email providers.

Limitations

Several constraints affected the scope and interpretability of this analysis. The examination was limited to email communications, potentially missing critical behavioral signals present in other digital channels such as web browsing, search queries, or social media interactions. Email content itself may not adequately capture the full spectrum of personal interests or crisis-related behaviors, as individuals may compartmentalize different aspects of their digital lives across platforms.

The zero-shot classification approach, while enabling analysis without labeled training data, may have introduced classification biases based on the pre-trained model's understanding of consumer categories. The predetermined interest categories might not optimally capture suicide-relevant themes, and the probability thresholds for meaningful engagement remain uncertain without validation data. Additionally, incomplete email records and varying communication patterns across individuals likely affected the representativeness of temporal trends.

Implications

Despite producing null findings for suicide prevention applications, this analysis provides valuable methodological insights for future digital phenotyping efforts. The technical framework successfully demonstrated that large-scale interest profiling from unstructured email data is achievable using modern NLP techniques. The approach could be adapted for other behavioral analysis applications where email content may provide stronger signals.

The absence of meaningful patterns in email-based interest profiling suggests that suicide prevention efforts may need to focus on other digital channels or combine multiple data sources for effective behavioral modeling. The heterogeneity observed across individuals reinforces that personalized approaches, rather than population-level patterns, may be necessary for identifying crisis states. This finding aligns with the complex, multifactorial nature of suicide risk that may not manifest clearly in consumer interest patterns.

For the Black Box Project's broader mission, these results indicate that not all digital traces carry equal weight in understanding pre-suicidal behavior. Resources may be better allocated to analyzing communication patterns, social connections, or help-seeking behaviors rather than consumer interests.

Future Directions

Several avenues could build upon this work's technical foundation while addressing its limitations. Integrating email-derived interest profiles with other digital channels such as web browsing history, search queries, and app usage data may reveal convergent patterns invisible in single-channel analysis. Implementing Named Entity Recognition to extract specific brands, services, and organizations could enable more granular analysis of financial distress indicators or help-seeking behaviors that current category-level classification may obscure.

Advanced temporal modeling using change-point detection or time-series anomaly detection algorithms could identify subtle shifts in individual behavior patterns rather than population-level trends. Developing suicide-specific classification models, trained on communications from individuals who died by suicide versus matched controls, may capture domain-specific language patterns that general-purpose models miss. Finally, exploring alternative feature representations such as emotional tone, linguistic complexity, or communication frequency changes may prove more informative than topical interests for crisis detection.

Key Observations

This analysis examined whether physical activity patterns, as captured through Apple Health step count and walking distance data, demonstrated meaningful deviations around validated high-risk periods for suicide. Using Random Cut Forest anomaly detection algorithms, we found no consistent behavioral signals across the analyzed population. While some individuals showed anomalous activity deviations around known risk events, others did not, and the directionality of changes varied considerably. Some risk periods coincided with increased activity, others with decreased activity, and many showed bidirectional changes over time. Most critically, the analysis revealed substantial gaps and inconsistencies in passive movement data collection across devices, limiting our ability to establish reliable behavioral baselines or detect meaningful changes.

Methodology

We extracted step count and walking distance data from Apple Health records across 15 available devices, ultimately analyzing 8 devices that contained both movement data and documented risk events. Data preprocessing included alignment with device end dates to ensure analysis excluded post-mortem artifacts and removal of extreme outliers through manual review and winsorization at the 95th percentile. To detect behavioral anomalies, we implemented Random Cut Forest (RCF) using AWS's open-source Java libraries with a Python wrapper, configured with a 15% anomaly rate scaled to time series length to balance sensitivity with false positive reduction.

For trend analysis and visualization, we applied LOESS (Locally Estimated Scatterplot Smoothing) regression with manually-tuned parameters to reveal underlying activity patterns while preserving meaningful fluctuations. Each individual's time series was temporally aligned with validated risk events including suicidal ideation, planning episodes, poor mental health days, and external stressors, enabling visual and statistical assessment of activity patterns around these critical periods.

Discussion

The analysis revealed highly individualized activity patterns that defied simple categorization or predictive modeling. Across the eight individuals analyzed, we observed remarkably diverse behavioral trajectories: some maintained relatively stable activity levels before experiencing sharp declines coinciding with risk events, while others demonstrated cyclical patterns spanning months or years with risk events occurring across various phases of these cycles. Several individuals showed extended periods of minimal activity followed by recovery phases, with risk events sometimes preceding, following, or occurring during these transitions.

The temporal relationship between activity changes and risk events proved particularly complex. Risk events clustered around activity transitions rather than consistently occurring during periods of high or low activity. This suggests that behavioral change itself, regardless of direction, may be more informative than absolute activity levels. However, without contextual information about military service requirements, deployment schedules, or mandated physical training, we cannot distinguish between internally-driven behavioral changes potentially related to psychological state and externally-imposed activity modifications.

The Random Cut Forest algorithm successfully identified statistical anomalies in movement patterns, but these anomalies showed no consistent relationship with documented risk periods. Some individuals exhibited anomalous activity spikes or drops near risk events, while others showed their most anomalous behavior during apparently stable psychological periods. This inconsistency undermines the utility of movement data as a standalone risk indicator in this population.

Limitations

The analysis faced several critical limitations that constrained our ability to draw meaningful conclusions. Passive data collection through consumer devices resulted in substantial gaps in the movement record, with periods of non-wear, delayed synchronization, and device switching creating an incomplete behavioral picture. The military-affiliated population introduced unique confounding factors, as occupational requirements for physical fitness, deployment-related activity changes, and mandated training schedules could drive activity patterns independently of psychological state.

Data quality issues further complicated interpretation. Several time series contained extreme outliers that could represent either legitimate activity bursts (such as during military exercises) or synchronization artifacts. While we implemented winsorization to minimize distortion, distinguishing between genuine behavioral signals and data artifacts remained challenging. Additionally, the analysis examined only individuals with both available movement data and documented risk events, potentially introducing selection bias as this subset may not represent the broader population of interest.

The fundamental limitation remains the absence of contextual data that could disambiguate the drivers of activity change. Without information about deployment status, duty requirements, geographic location, or physical health conditions, we cannot determine whether observed activity patterns reflect psychological state changes or external circumstances.

Implications

This analysis yields several important implications for the field of digital phenotyping in suicide prevention. First, it demonstrates that movement data from consumer devices, while readily available, may be too inconsistent and context-dependent to serve as a reliable standalone indicator of suicide risk. The high variability in both data quality and individual behavioral patterns suggests that any predictive model relying primarily on step count or distance metrics would likely generate unacceptable false positive and false negative rates.

However, the observation that risk events often cluster around periods of behavioral transition, regardless of direction, may indicate that movement data could contribute value as part of a multimodal assessment approach. When combined with other digital biomarkers such as sleep patterns, communication behaviors, or physiological metrics, activity transitions might help identify periods warranting closer monitoring or intervention. This finding aligns with emerging evidence that behavioral instability, rather than specific behavioral states, may signal elevated risk.

For researchers and clinicians working with military and veteran populations, these results underscore the critical importance of incorporating service-related context into any behavioral monitoring system. The same activity pattern could indicate psychological distress in a civilian but represent normal occupational demands in an active-duty service member. Future digital phenotyping efforts must account for these population-specific factors to avoid misinterpretation of behavioral signals.

Future Directions

Immediate next steps should focus on addressing the contextual gaps identified in this analysis. Integrating military service records, deployment histories, and duty status information would enable researchers to stratify movement patterns by external circumstances, potentially revealing risk-related signals currently obscured by occupational confounds. Additionally, expanding beyond simple step counts to examine more nuanced movement features such as circadian rhythm disruption, weekend-weekday differentials, or entropy measures of daily activity may yield more informative behavioral markers.

The inconsistent data availability highlighted by this analysis suggests that relying solely on passive consumer device collection may be insufficient for clinical applications. Future efforts might explore hybrid approaches combining passive monitoring with periodic active data collection or ecological momentary assessments to ensure more complete behavioral records. Alternatively, research-grade wearables with more reliable continuous monitoring capabilities could provide the data consistency necessary for robust risk modeling.

Longer-term vision should embrace federated learning approaches that preserve privacy while enabling population-level pattern discovery. By analyzing movement patterns across multiple cohorts and contexts without centralizing sensitive data, researchers could identify which behavioral signals generalize across populations and which remain specific to particular subgroups or circumstances. This approach aligns with the Black Box Project's broader mission of breaking down data silos while maintaining the highest standards of privacy protection.

TECHNICAL APPENDIX

This technical appendix documents the operational and infrastructure foundations that enable the Black Box Project to transform donated devices into actionable insights for suicide prevention. While the preceding sections of this white paper focus on findings and their implications, the work described here represents the critical bridge between raw device data and meaningful analysis. These technical components embody years of development, partnership, and refinement that make our work possible.

The device extraction and data processing procedures documented in this appendix reflect our unwavering commitment to both scientific rigor and respect for the families who entrust us with their loved ones' digital legacies. Every technical decision, from our choice of forensic tools to our data handling protocols, balances the competing demands of comprehensive data extraction, privacy protection, and analytical utility. The standardized processes we have developed ensure that each device receives the same meticulous attention, whether it represents our first donation or our hundredth.

Our cloud infrastructure architecture demonstrates how modern data engineering can serve humanitarian purposes. The platform we have built with our partners transcends typical commercial applications, creating a secure, scalable environment specifically designed for the sensitive nature of posthumous digital data. This infrastructure does more than store and process information; it preserves the dignity of those we study while enabling discoveries that could prevent future tragedies.

The partnerships that made these technical achievements possible deserve special recognition. Cellebrite and Magnet Forensics provided not just world-class forensic tools, but ongoing support and belief in our mission. Amazon Web Services contributed both infrastructure and expertise, helping us architect solutions that protect privacy while enabling breakthrough analytics. Pariveda demonstrated how technology consulting can serve a higher purpose, going beyond contractual obligations to ensure our systems truly serve the families and researchers who depend on them.

For readers seeking to understand the mechanics behind our findings, this appendix provides transparency into our methods. For potential collaborators considering similar initiatives, it offers a roadmap built on hard-won experience. For the broader research community, it establishes standards for handling sensitive digital data in service of public health objectives.

The technical details that follow may seem removed from the human tragedy of suicide, but they represent our concrete commitment to ensuring that no family's contribution is wasted, no analytical opportunity is missed, and no corner is cut in our pursuit of prevention. Behind every algorithm and architecture decision lies the same driving purpose: to honor those we have lost by protecting those who still live.

APPENDIX A

Overview

This appendix provides insight into the operational process employed by the Stop Soldier Suicide forensic team to extract data from donated devices and transfer it to our secure cloud infrastructure. This technical overview documents our standardized approach to handling devices from initial receipt through data ingestion into AWS, utilizing industry-standard forensic tools.

Device Receipt and Chain of Custody

When a device arrives at our forensic laboratory, our team follows rigorous evidence handling procedures to maintain data integrity and honor the trust placed in us by donor families. Each device is logged into our evidence management system, assigned a unique document number, and stored in our secure evidence repository. Our Digital Forensics Lab Manager oversees this process, ensuring every device is tracked through a detailed chain of custody that documents each interaction with the device.

The evidence custodian records essential information including the device type, date received, associated case information, and physical condition. This meticulous documentation serves dual purposes: maintaining the integrity of our process and providing families with assurance that their loved one's device is handled with the utmost care and professionalism.

Data Extraction Process

Our primary extraction tool is Cellebrite, a forensic platform widely used by law enforcement and digital forensics professionals worldwide. Cellebrite works with both Android and iOS devices, attempting to extract all available data including deleted files when possible. The tool creates two extraction views: a logical extraction representing the phone user's view, and a physical extraction that operates with root-level access to capture more complete data. Our team prioritizes physical extractions whenever technically feasible to ensure the most comprehensive data collection.

The extraction process captures diverse data types organized into two main categories: content and data. Content includes instant messages, bookmarks, web browsing history, and similar user-generated information. Data encompasses images, videos, documents, text files, and other media. Each category contains numerous subcategories. Our extractions typically include applications data, usage logs, archived files, audio recordings, calendar entries, call logs, chat conversations, configuration files, contacts, cookies, system level databases, device connectivity logs, event logs, device information, user accounts, locations data, network usage, passwords, searched items, social media content, text messages, timeline data, and web artifacts.

Data Export and Formatting

Once extraction is complete, the forensic software provides multiple export format options to accommodate different analytical needs. Our team typically utilizes the Excel workbook format, which organizes data types into separate tabs within a single file. This format proves most efficient for subsequent data processing and ingestion into our data lake infrastructure. The XML format, while comprehensive, presents challenges due to its verbose nature and repeated metadata, making it less suitable for our downstream processing requirements.

The exported data maintains all original information from the device, with metadata preserved for each data element. This includes crucial identifiers such as the device's IMEI number, which serves as a globally unique identifier linking all data to its source device. The export process does not result in any data loss. Even previously deleted files recovered during extraction are included in the output.

Transfer to Cloud Infrastructure

Following successful extraction and export, our team initiates the secure transfer process to AWS. Given that our forensic workstations remain disconnected from networks for security purposes, we employ a manual but secure upload protocol. The Digital Forensics Lab Manager oversees this critical step, utilizing secure file transfer protocols to move the extracted data from our air-gapped forensic environment to designated landing zones within our S3 storage infrastructure.

This manual process, while requiring additional time compared to automated transfers, ensures maximum security and provides an additional checkpoint for quality assurance. Each upload is verified for completeness and integrity before the local extraction files are archived according to our data retention policies.

Quality Assurance and Completeness

Throughout this process, our team maintains a commitment to extracting the maximum amount of data possible from each device. Our Digital Forensics Lab Manager, who holds industry-standard certifications in digital forensics, reviews each extraction to ensure completeness. If initial extraction attempts yield limited results, the team employs alternative forensic tools including Magnet Axiom, Oxygen Forensics, and other specialized software to maximize data recovery.

The extraction process respects the technical limitations of each device while pursuing comprehensive data collection. Some devices may have encryption, damage, or other factors that limit extraction capabilities. In such cases, our team documents these limitations while extracting all technically accessible data.

Conclusion

This standardized process from device receipt through cloud ingestion represents our operational foundation for the Black Box Project. By maintaining rigorous procedures and utilizing professional-grade forensic tools, we ensure that each device's data is extracted comprehensively, documented thoroughly, and transferred securely. This systematic approach honors both the technical requirements of our analysis and the trust placed in us by the families who have donated their loved ones' devices to advance suicide prevention efforts.

APPENDIX B

Overview

The Black Box Project's cloud infrastructure leverages a modern data lakehouse architecture built on Amazon Web Services (AWS) to process, analyze, and derive insights from digital device data. The platform is designed to handle sensitive data with appropriate security controls while enabling scalable analytics and machine learning capabilities.

Architecture Components

Data Ingestion Layer

The infrastructure employs AWS Transfer Family to provide a secure SFTP endpoint for forensic analysts to upload extracted device data. These Excel files containing parsed device information are automatically transferred to designated S3 buckets in the landing zone, initiating the data processing pipeline through S3 event notifications.

Data Processing Pipeline

The processing pipeline orchestrates data transformation through several key components:

AWS Step Functions coordinates the multi-stage workflow, ensuring reliable execution and error handling across the transformation process. Lambda functions serve as the primary compute engine for initial data processing, converting Excel files into PySpark dataframes and registering them in the AWS Glue Data Catalog.

AWS Glue ETL jobs handle critical personally identifiable information (PII) redaction using the AWS Glue EntityDetector library, ensuring compliance with privacy requirements before data moves to the cleansed zone. For compute-intensive feature engineering tasks, Amazon ECS clusters process Pandas-based transformations, interfacing with Apache Iceberg tables through Amazon Athena for read/write operations.

Data Lake Architecture

The platform implements a three-zone data lake structure:

1. Landing Zone (Raw): Stores original device data in its native format
2. Cleansed Zone (Analytics-Ready): Contains PII-redacted, standardized data in Apache Iceberg format
3. Curated Zone (Purpose-Ready): Houses feature-engineered datasets optimized for specific analytical use cases

All data assets are cataloged in AWS Glue Data Catalog with governance controls managed through AWS Lake Formation, ensuring proper access permissions and data lineage tracking.

Storage and Format Strategy

The infrastructure utilizes Amazon S3 as the primary storage layer, with all analytical data stored in Apache Iceberg format. This choice provides:

- ACID transaction support for reliable data updates
- Time travel capabilities for data versioning Efficient handling of schema evolution
- Optimized query performance through partition pruning and file compaction

To optimize costs while maintaining Apache Spark compatibility, the team developed a custom Lambda container image capable of running Spark workloads within Lambda's runtime constraints, eliminating the need for persistent EMR clusters.

Analytics and Machine Learning Layer

Processed data in the curated zone feeds into:

- Amazon QuickSight for business intelligence dashboards and visualizations
- Amazon SageMaker for machine learning model development and training
- Amazon Athena for ad-hoc SQL queries and exploratory data analysis

Infrastructure as Code

The entire infrastructure is provisioned using AWS Cloud Development Kit (CDK) with deployment managed through AWS CodePipeline. This approach ensures:

- Version-controlled infrastructure definitions
- Consistent deployments across development, staging, and production environments
- Automated testing and validation of infrastructure changes
- Manual approval gates for production deployments

Security and Compliance Considerations

The architecture implements multiple security layers:

- Encryption at rest for all S3 buckets using AWS KMS
- Encryption in transit for all data transfers IAM roles with least-privilege access policies
- VPC endpoints for private connectivity between services
- CloudTrail logging for audit compliance Automated PII detection and redaction in the data pipeline

Scalability and Performance

The serverless-first approach provides automatic scaling based on workload demands. Lambda functions scale horizontally to handle concurrent file uploads, while ECS tasks can be dynamically provisioned for intensive feature engineering operations. The use of Apache Iceberg format ensures query performance remains consistent as data volumes grow.

Cost Optimization

Several architectural decisions optimize operational costs:

- Serverless compute (Lambda, ECS Fargate) eliminates idle infrastructure costs
- S3 lifecycle policies move older data to cheaper storage tiers
- Spot instances for non-critical ECS tasks
- Query result caching in Athena reduces redundant compute

This cloud infrastructure provides Stop Soldier Suicide with a robust, scalable foundation for the Black Box Project's data processing needs while maintaining the security and compliance requirements necessary for handling sensitive personal data.