Omelas Use Cases White Paper

Measurement and Evaluation

Abstract

According to a 2017 report by the Government Accountability Office, "no cohesive strategy with measurable outcomes has been established to guide the multi-agency CVE effort towards its goals." This echoes a 2016 metastudy conducted by National Consortium for the Study of Terrorism and Responses to Terrorism (START) that "identified only five published studies that gathered prevention program outcome data" and "did not find any CVE evaluations based on experimental design".

The lack of measurable outcomes in CVE gives practitioners little sense of the efficacy of programs, let alone how to improve them. The issue stems not only from intrinsic difficulties in measuring states of mind, but also from a lack of empiricism within the field. This section will look at how similar fields have addressed these issues and propose solutions enabled by new advances in technology.

Challenge

There are numerous challenges to measuring and evaluating countering violent extremism campaigns. Campaigns intended to reduce radicalization must define and measure radicalization. The START metastudy found that programs collecting data used exclusively attitudinal surveys, or questionnaires about opinions, to do so, with little consistency among different programs. Attitudinal surveys are subject to numerous response biases, ranging from survey takers giving answers they believe the survey giver wants to hear to survey takers responding based on fleeting emotion rather than long held beliefs. Setting up an experimental design, which would control for those biases as well as other confounding factors, did not occur for any test.

The lack of experimental design presents another major challenge. Without setting up control and test populations and controlling for confounding factors, the few metrics that do exist provide no insight into causation. The lack of consistent metrics hinders the ability to ensure that a control and a test group have the same level of radicalization before a program begins.

The unquantified nature of the field and the inconsistency of measures raises a major obstacle to the general acceptance of a single metric to quantify radicalization. As no one within the field is accustomed to a single a measure, and standardization across program is extremely rare, there is still room for significant debate about the best possible metrics to use.

Product Design

The preponderance of publicly available data through social media and more has enabled other fields to begin to address the issues with response biases. Psychiatry has led the way in this regard. Works published in the

Trends and Applications in Knowledge Discovery and Data Mining and *Proceedings of the 5th Annual ACM Web Science Conference* show that social media posts, in which subjects tend to be more candid, can detect chronic depression.

Feelings of allegiance to a certain organization is a well studied field, as well, led by advertisers in the corporate world and management scientists in academia. Studies published in the 8th International Strategic Management Conference, International Journal of Information Management, The Journal of Interactive Marketing, and more show that social media can be a powerful signal for measuring loyalty and sentiment towards an organization. That has become standard in the advertising industry, where industry giants like Facebook prominently promote their capacity to use social media to measure opinion.

Like CVE, advertising faced the same issue of lacking a unified metric to measure brand awareness. Advertising networks resolved this concern through a commitment to transparency and the provision of multiple metrics. For instance, Google suggests using at least four metrics to understand changes in brand loyalty.

Omelas empowers CVE practitioners to use a full set of metrics to understand changes in radicalization and offers a platform designed to bring experimental design to CVE programs. Omelas provides numerous metrics to understand radicalization. The flagship metric is radicalization score, which measures the similarity between an individual or a group's online behavior and that of a known violent extremist or violent extremist group. Other metrics, such as engagement with known violent extremist accounts, posts originating from extremist websites, and number of posts containing extremist content, allow a full and nuanced view of someone's level of radicalization and affinity for extremist groups.

Omelas allows clients to split individuals into test segments and controls. By providing metrics on those segments, clients can run tests with the assurance that the test and control are similar enough to control for confounding factors.

Conclusion

Omelas's offering addresses three of the biggest hurdles facing the measurement and evaluation of CVE: 1) Lack of metrics, 2) Response biases associated with the few metrics observed, and 3) Lack of scientific rigor in experimentation. Omelas addresses the lack of metrics in the field by presenting a wide variety of metrics that enable clients to view radicalization from multiple viewpoints. Omelas controls for response biases by following in the footsteps of other fields that have used the candor of social media to find more reliable data on how people think and feel. And Omelas offers easy segmentation into control and test groups to isolate confounding variables and implement experimental design.

Omelas empowers CVE practitioners to measure and evaluate their programs with a new level of accuracy, providing benchmarks for performance and clear understanding of how to become more effective at combating violent extremism.

Open Source Intelligence

Abstract

According to a 2007 report by the Congressional Research Service, "A consensus now exists that OSINT [open source intelligence] must be systematically collected and should constitute an essential component of analytical products". With nearly a hundred times as much data produced per year as when the report came out, the potential of open source intelligence has never been greater.

Despite this, collecting OSINT, especially on violent extremist organizations, remains a manual, subjective, and fragmented process. This section will discuss the causes of these and how other fields have dealt with similar issues, as well as a review of existing tools.

Challenge

The challenges facing OSINT gathering can be broken into three parts: lack of automation in tracking, the scattering of useful data across the web, and conclusions colored by personal experience and with little empirical backing. While a significant portion of OSINT focuses on state actors, this section will focus on extremist content alone.

Tracking public extremist networks and content is currently a manual process. Content discovery and extremist relies on a number of best practices to guide, but relies heavily on repetitive work by analysts. Discovering networks requires manual recording of those networks. With tech giants like Google and Facebook continuously trying to remove extremist content, discovering a network can be an ephemeral accomplishment, and, according to analysts surveyed, hours out of their week are spent finding new networks after old accounts and content has been removed.

This is especially challenging because much of that work is collecting data from disparate sources across the web. Collecting data sucks time away from analysis, while inconsistent systems to track this data leads to friction.

Finally, OSINT relies heavily on qualitative analysis, with analysts providing summaries and samples of their findings. The process of determining what to include and exclude, while guided by best practices, is inevitably colored by the personal experiences of the analyst and their team. We found no study with quantitative support for what to include in OSINT reports and few reports that included metrics of any kind. OSINT that includes metrics rarely provides analysis over time and even rare is controlling for variables. This leave OSINT open to significant subjective interpretation and two analysts can deliver significantly differing reports on the same subject.

Product Design

There exist a number applications for assisting in parts of OSINT gathering. For instance, hunch.ly assists in recording websites while tracking. Maltego and FOCA map out connections between data and individuals. Maltego collects metadata on organizations for users as well but requires significant technical skills and labor to do so.

Omelas delivers the largest reduction in workload for analysts and enables the greatest shift from research to analysis. Omelas automates the detection of social media accounts, blogs, forums, and Telegram channels spreading extremist content. Not only is the process of mapping a network fully automated, but also analysis is provided on connections in the network, such as an individual account's influence and closeness to other accounts.

Included in that analysis is the integration of data from sources across the web. Omelas organizes this information in a readily accessible and easily understood way, as well as giving analysts the ability to customize how and what information is displayed. Omelas is the first OSINT tool to integrate multiple social media platforms in one dashboard.

Omelas quantifies the most important measures in understanding the way extremist networks function online. We demonstrate how content spreads through a network, the most important influencers within a network, and how multiple networks connect to each other. We provide automated visualizations to demonstrate network and content flows, each easily customized and ready to insert into reports.

Conclusion

Omelas's offering addresses three major hurdles that have impeded the ability of OSINT gathering to keep up with the expansion of OSINT potential: 1) manual nature of open source intelligence gathering; 2) fragmented nature of data sources; and 3) the lack of empirical backing for findings.

Omelas automates the process of intelligence gathering including integrating data sources from dozens of places online. Omelas provides empirical backing for conclusions and analysis, providing statistical backing to what is now left to instinct.