Introducing DarkLight ? Game Changing Artificial Intelligence for Cyber Security

Submitted by spriley on Tue, 12/06/2016 - 10:26am

A couple years ago, I wrote a popular research paper on the Science of Security for the Centre for Strategic Cyberspace + Security Science (CSCSS), a UK based cyber think tank, detailing how the artificial intelligence (AI) field of knowledge representation and reasoning (KR&R) could be applied to the enterprise cyber security ecosystem to help organizations develop a scientific foundation to their cyber security program. These types of evidence-driven, AI knowledge representation and reasoning solutions should not be confused with the statistical and mathematical modeling based machine learning solutions as the two are very different but complimentary approaches. In a future article, I'll highlight how these two different types of approaches can be applied to maximize the ROI from these solutions.

When knowledge representation and reasoning is applied to a specific domain like cyber security and the system is taught by cyber security experts, it can create what is known as an "expert system". An expert system is a computer system that emulates the decision-making ability of a human expert and are designed to solve complex problems by reasoning about the knowledge. Just like machine learning, expert systems have benefited from advances in technology, digitized data, information, and knowledge bases and our ability to apply these analytic solutions to real world problems has greatly increased over the years. In other words, these aren't the expert systems that dominated artificial intelligence 30 years ago in the 80s and pushed machine learning out of favor. They're more powerful, more agile, standards-based, and easier to use.

When we think about the core security science based functions in security operations such as threat intelligence, incident response, security monitoring, and other "analyst" positions, a good portion of their day to day activities are data-driven security processes where the analysts use the evidence in the data to make decisions and act on the evidence they are seeing. This type of AI-based expert system allows analysts to apply their domain expertise and experience in the organization to create prescriptive analytics that teach the AI how to make automated evidence-driven decisions and to orchestrate courses of action using the analyst's logical data-driven processes and reasoning.

I'd like to introduce the community to a new type of AI expert system for Security Operations, Analytics, and Reporting (SOAR) called DarkLight from Champion Technology Company, Inc. Gartner defines SOAR as:

"Security operations, analysis and reporting technologies support workflow

management and automation, analytics and reporting. This enables security operations teams to automate and prioritize security operational activities and report data to inform better business decision making."

DarkLight, an artificial intelligence software platform, allows analysts to codify their logical processes and run them at machine-speed, 24-hours a day. DarkLight's approach stems from decades of R&D at a National Laboratory working on semantic graphs, knowledge representation and advanced reasoning systems. Borrowing from the defense and intelligence community's revolutionary Object-Based Production (OBP) methodology, DarkLight's unique AI organizes what is known, infers what can be known, and provides means to discover the "unknown unknowns" hiding in our cyber ecosystems using Activity-Based Intelligence (ABI) tradecraft. This type of ABI tradecraft reasoning focuses on transactions, behaviors, and activities rather than signatures or mathematical algorithms to discover the unknown unknowns. DarkLight's patented analytic methodologies drive automated, evidence-driven decisions and orchestrated courses of action at machine speed to help organizations get ahead of the threats.

Let's take a closer look at DarkLight by breaking it down into two key areas, 'knowledge representation' and 'reasoning', and what benefits this approach brings to organizations. Since DarkLight is an AI-based solution it might help to think of DarkLight as a virtual analyst. Knowledge representation languages or ontologies are used to teach the AI about the world. In DarkLight's case, as an AI-based virtual analyst, that world is the cyber security ecosystem where DarkLight could be used in any of the data-driven analyst based security science areas in operations like threat intelligence, active defense, incident response, etc. DarkLight uses the W3C OWL2 standard as the description logic knowledge representation language since this is a mature, 2nd generation standardized language with increasing adoption across industries such as Healthcare and Biomedical, Financial Services, and the Defense and Intelligence Community.

The cyber security ontologies are the conceptual model for creating a cyber security knowledge and activity graph based on the knowledge they represent. Knowledge and activity graphs were made popular by technology giants like Google, Facebook, Microsoft and others who are using these semantic graphs to enable advance analytics on their massive datasets. The ontologies provide a unified conceptual model for a semantic graph where the objects (nouns) like people, places, things, and events are all represented as nodes in the graph and the attributes, associations, and activities (or verbs) of those objects are represented as edges. When the cyber security data and information is ingested into DarkLight and mapped to the ontologies the knowledge from the individual files is automatically organized into a cyber security knowledge and activity graph using the object-based production methodology.

DarkLight is working on updating and integrating over 100 modular cyber security ontologies that have been developed by the community over the past couple years to include standards-based knowledge representations of many of the government sponsored cyber security measurement and management

architecture standards like STIX, CYBOX, MAEC, CAPEC, CWE, CVE, and SCAP as well as new ontologies like the Insider Threat Indicators ontology from CMU. This will enable DarkLight to understand the meaning and context of the cyber security data and information coming from those standardized common languages and enable DarkLight to automatically organize that data into the cyber security knowledge and activity graph where it can then be taught to apply the knowledge as a virtual analyst and make evidence-driven decisions and orchestrate courses of action.

A differentiator for DarkLight is how it uses the knowledge representation languages to create an Enterprise Contextual Knowledge graph that allows the AI-based virtual analyst to understand organization's people, processes, and technologies. DarkLight also creates an Adversarial Contextual Knowledge graph that allows the AI-based virtual analyst to understand the threat actors, campaigns, TTPs, indicators and other known threat information from incident response, threat intelligence, and threat sharing. This makes DarkLight ideal for helping organizations maximize the ROI from the current technologies deployed in their cyber ecosystem through automation and orchestration because DarkLight has the contextual knowledge to make sense of and act on what the other security technologies and sensors are seeing and reporting. Cyber is a team sport and if the AI-based virtual analyst is going to be a team player for the organization they need to know who the different teams are, their respective playbooks, and the environment they are playing in.

The ontologies also provide an obvious means for pursuing a declarative approach to cyber security data and information integration while capturing provenance and data governance information. The ontologies define the concepts used daily in cyber security and provide a mapping in DarkLight to the actual datasets containing those concepts. The ontologies then in turn provide an enterprise data model for the organization cyber security program since the knowledge is represented in a way that the human experts, data stewards, and the AI can both understand and use.

The cyber security ontologies in DarkLight provides surface learning for the AI-based virtual analyst so it can understand the meaning and context of the cyber security objects, attributes, associations and activities in the organization's cyber security knowledge and activity graph. Deep learning for DarkLight's AI-based virtual analyst is provided by cyber security experts who share their knowledge about their logical step by step, data-driven processes and the reasoning required to make evidence-driven decisions and automated courses of action based on those decisions.

The cyber security analysts in our organizations are the experts who know our organization's data-driven, intelligence-based processes and what actions to take in the organization's cyber ecosystem. DarkLight provides easy to use wizards for analysts to create programmable reasoning objects (PROs) based on their day to day data-driven processes to create AI-based prescriptive analytics that can automate evidence-driven decisions and orchestrate automated courses of action

in the organization's cyber ecosystem. Once an analyst has learned how to use DarkLight, it's expected senior analysts will produce an average of 1 PRO per day, mid-level analysts will average few PROs a week, and junior analysts will average about 1 PRO per week.

Unlike a black box machine learning approach, all the logic is exposed, defendable, and can be used as a learning method to educate junior analysts. The conceptual modeling approach in DarkLight is much more transparent than the statistical and mathematical modeling approaches of machine learning based solutions because they are the domain specific concepts and terms cyber security professionals use in their daily jobs. This approach doesn't require data scientists or mathematicians to create the analytics, it puts the power of self-service analytics in the hands of the analysts and business users to easily create their own prescriptive analytics based on their domain expertise and operational experience in the organization.

DarkLight enables organizations to share ontologies for different datasets and PROs for the AI-based prescriptive analytics and automated courses of action they create for threat hunting, insider threats, false positive reduction, etc with other organizations through ontology and PRO sharing in the same way organization might share threat intelligence and IDS or AV signatures today. This enables organizations to share analytic tradecraft with each other and enables the AI-based virtual analyst to learn from human experts at other organizations who might be more advanced in their tradecraft knowledge across different functional areas.

The ability to share the ontologies and PROs for AI-based prescriptive analytics with other organizations is a game changer. Suddenly the Information Sharing and Analysis Centers (ISACs) and Information Sharing and Analysis Organizations (ISAOs) can move from sharing actionable intelligence to sharing actionable intelligence and PROs that can automate the decisions and actions that should be taken on that intelligence.

DarkLight empowers organizations, ISACs, ISAOs, security vendors, and CERTs to create custom prescriptive analytic PROs based on their domain expertise and experience and share the PROs with other organizations to help the community move from sharing knowledge about threats to sharing step by step machine-readable instructions to automate and orchestrate what to do with that knowledge. I believe this can significantly shift the balance of power from the adversaries to the defenders.

Additionally, this enables the ability to evolve analytics incrementally as our tradecraft knowledge evolves and scale those analytics with the data by simply adding more PROs encoded with the cyber security analytic tradecraft knowledge of human experts. The goal isn't to replace the human analysts defending our organizations but to make them more effective and efficient by giving them an army of AI-based virtual analyst experts to assist them in defending our enterprises. Ultimately organizations can apply DarkLight where they need the

help and the evidence-based decisions will help them feel more comfortable about enabling AI-based automation since it's the same evidence and data-driven processes their human experts use.

When you think about it, by creating PROs in DarkLight, human analysts are capturing their analytic tradecraft knowledge in a machine-readable, sharable format that also serves to document this tradecraft as part of the organizational knowledge to protect the organization from possible "brain drain" should the analyst leave. By capturing this knowledge in a PRO, DarkLight can perform the data-driven analysis, make evidence-based decisions, and orchestrate automated courses of action in additional to being a learning tool for new analysts to help them rapidly get up to speed on how analysis is done in their new organization.

If you are faculty or staff overseeing a cybersecurity program, ask us about the DarkLight Cyber Security Educational Consortium (CSEC) and utilizing DarkLight in the classroom.

As part of our committment to support global cybersecurity education, for qualified Academic intstitutions, DarkLight will donate its next generation, patented cyber security software as a no-cost academic license to all CSEC partner schools to be used in classrooms as a curriculum tool to support cybersecurity-degree seeking students.

For more information: https://www.darklightcyber.com/darklight-university

---

? Not so productive Grad School Links WEIS 2016 - Trip ?


Science of Security Topics
