CIRI | CRITICAL INFRASTRUCTURE RESILIENCE INSTITUTE

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

UNT UNIVERSITY OF NORTH TEXAS

PI: Dan J. Kim

Jordan Bernot

Max Douglas

CIRI advisors: Casey O'Brien & Andrea Whitesell

Enhancing Cybersecurity KSATs (Knowledge, Skills, Abilities, and Tasks) for Cyber Education and Workforce Development

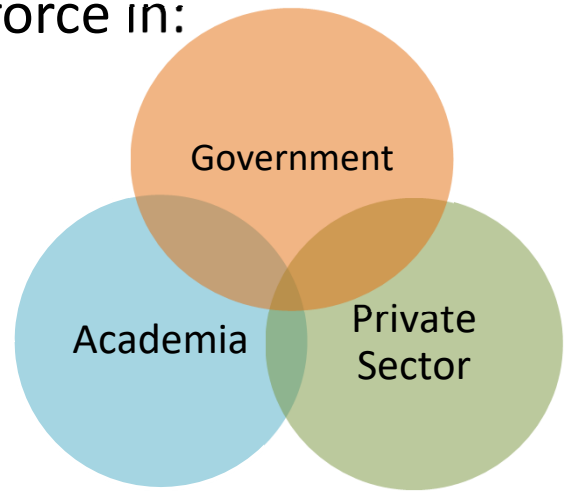
August 3, 2021

Agenda

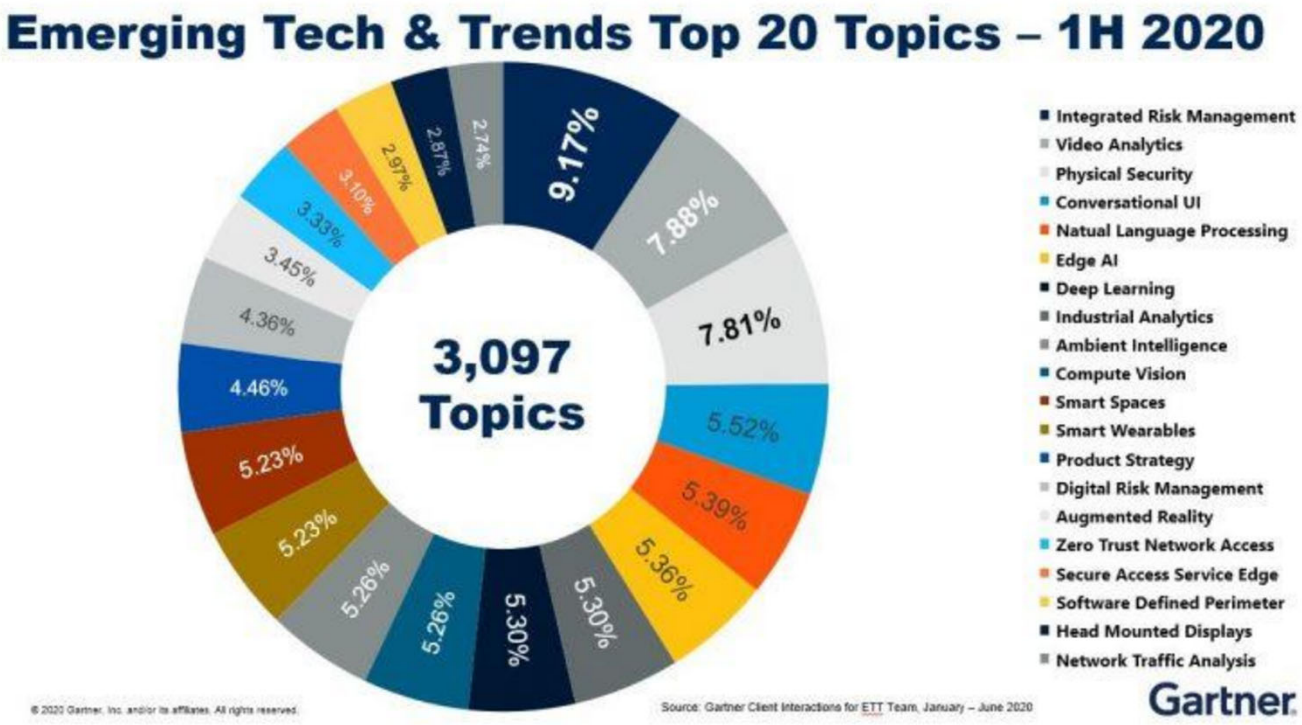
- Project overview
- Sub-project 1: “Analysis and Comparison of Cybersecurity Workforce Frameworks”
- Sub-project 2: “Exploratory Analysis of Cybersecurity KSATs from Job Posting Data Using NLP techniques”
- Discussion
  - How to strengthen workforce development and education
  - Future directions

Motivation

- Cybersecurity - the biggest fear for many organizations. (e.g., Colonial Pipeline ransomware attack, etc.) (Cowley 2018)
- Challenges between supply and demand of cybersecurity workforce
- Substantial initiatives to increase Cybersecurity workforce in:
  - Government
  - Academia
  - Private Sector



Cybersecurity is a moving target



Motivation (cont.)

- Gap between sophisticated cyber-attacks and cyber defense mechanisms
- Cybersecurity KSATs should be updated with emerging technologies

→ Need to identify and update new KSATs for Cybersecurity Workforce

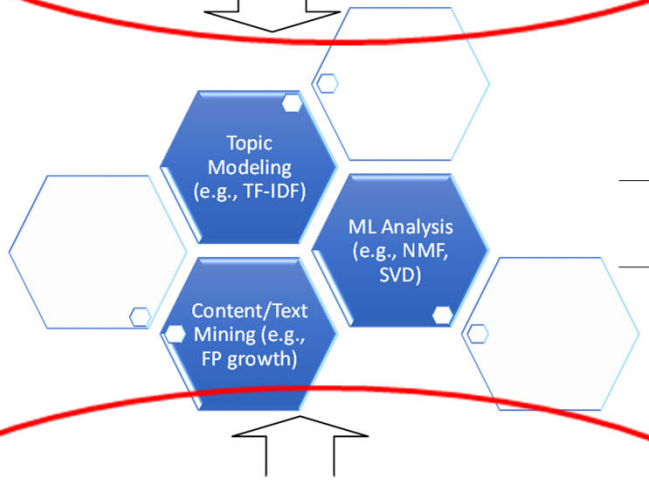
Research Purposes

- To analyze and compare existing cybersecurity workforce frameworks to identify detailed required elements
- To find new Cybersecurity KSATs from Job Posting Data through NLP techniques
- To provide a prototype to enhance current KSATs for cybersecurity education and workforce development

Comparison Analysis of Cybersecurity Workforce Frameworks

- NIST and new NICE Cybersecurity workforce framework,
- DoD Directive 8570 and 8140,
- CAE Cyber Defense Education Requirements
- Cybersecurity Competency Model, & others

Sub-Project #1



Outcomes

- Identified Cybersecurity job titles, job roles
- Key KSATs and new KSATs
- Roadmap of Career Paths With KSATs

Exploratory Analysis of Cybersecurity KSATs

- Job descriptions including knowledge, skills, abilities, tasks in job posting data
- Trend reports on emerging jobs and technologies

Sub-Project #2

Sub-project 1: "Comparison and Analysis of Cybersecurity Workforce (CSW) Frameworks"

Motivation

- Large complex corpus of frameworks across various industries
- Frameworks dedicated to cybersecurity workforce (CSW) development

Goals

- Identify differences and similarities between CSW frameworks
- Provide clarity

Method/Analysis

- Collect frameworks across industries
- Focus on CSW development
- Initial frameworks identified, reduced to six
  - Analyzed frameworks: 181 & 181r, COBIT, DoDD 8140, CCM, NCAE
- Develop criteria to evaluate frameworks
  - Key elements: publishing org., goal/objective, audience, mapping component

Outcome: Comparison Table

	NICE Cybersecurity Workforce Framework (NCWF) - NIST SP 800-181 in 2017	NICE CWF- NIST SP 800-181r1 in 2020	DoDD 8570/8140 Cyber Workforce Framework	National Centers of Academic Excellence in Cyber Defense/Cyber Operations (CAE-CD/CO)	Cybersecurity Competency Model	ISACA COBIT
Initiative organization	NIST (National Institute of Standard and Technology)	NIST	DoD (Department of Defense)	NSA	Department of Labor	Information Systems Audit and Control Association
Goal/Objective	Describes and categorizes cybersecurity work and identifies sample job titles, tasks, and KSAs (Knowledge, skills, and Abilities)	Describes cybersecurity work and learners implemented through a modular building block approach comprised of Tasks, Knowledge, and Skills	Provides the foundation for identifying education, training, and certification requirements to support cybersecurity personnel qualification	Promotes the focus of cybersecurity in higher education curriculum and research Aims to produces professionals with Cyber Operations (CO) and Cyber Defense (CD) expertise in various disciplines	Provide a comprehensive overview of cybersecurity workforce competencies, including roles, training, and career paths for job seekers.	Provide guidence for businesses to develop and organize, and implement strategies regarding information management and technology
Main Audience	• Employers • Current/Future Cybersecurity workers • Training/Certification Programs • Educators • Technology Providers	• Employers • Organizations, public and private • Educators/Trainers • Students and those looking to enter the field	• Military, civilian, and contractor cybersecurity personnel • Training vendors, certification bodies, colleges, and continuing education providers	2-year, 4-year, and graduate-level institutions in the U.S. seeking CAE designations	Current and future workforce (generally unemployed or under employed) participants, as well as employers, who need guidance on work role-KSA congruity	Businesses IT and Assurance, historically the financial sector, helps provide insight to senior management
Mapping Components	• 7 Categories (Securely, Provision, Protect and Defend, Investigate, College and Operate, analyze, Operate and Maintain, Oversee and Govern) • 33 Specialty Areas • 52 Work Roles (KSA)	(New TKS in Draft Previous NIST components valid until further publication) Nov 2021	• DoD work roles, tasks, functions and baseline KSAs • Under development awaiting DoD 8140.01 Manual • 33 Specialty Areas, 54 Work Roles	• 3 Foundational KUs • 5 Core Technical KUs • 5 Core Non-Technical KUs • 57 Optional KUs	• 5 base tiers - pyramid structure, tapering from general 'soft' skills and responsibilities to specialized competency areas at the top • Utilizes NICE Framework (NCWF), KSAs - 52 work roles	3 Focus areas: Devops, IT Risk, Security 6 Governance Principles 3 Framework Principles 5 Domains (Government and Management Objectives)
Highlights	• Backed by NIST • Large target audience • Biggest mapping component	• 181r provides a simplified and flexible approach to describe cybersecurity work • Able to be utilized by a wide range of users	• Small target audience • 8140.01 Manual not released yet • Strict guidelines and requirements	• 100% Academia focused • Technical/Non-technical KUs • 2nd Biggest mapping component	• Similar structure as NICE • Large target audience • Catered toward newer workforce participants	Main focus on management and enterprise governance Does not mention elements of workforce development

Outcome: NICE CWF_2017 vs. NICE CWF_2020

	NICE Cybersecurity Workforce Framework (NCWF) - NIST SP 800-181 in 2017	NICE CWF- NIST SP 800-181r1 in 2020
Initiative organization	NIST (National Institute of Standard and Technology)	NIST
Goal/Objective	Describes and categorizes cybersecurity work and identifies sample job titles, tasks, and KSAs (Knowledge, skills, and Abilities)	Describes cybersecurity work and learners implemented through a modular building block approach comprised of Tasks, Knowledge, and Skills
Main Audience	• Employers • Current/Future Cybersecurity workers • Training/Certification Programs • Educators • Technology Providers	• Employers • Organizations, public and private • Educators/Trainers • Students and those looking to enter the field
Mapping Components	• 7 Categories (Securely, Provision, Protect and Defend, Investigate, College and Operate, analyze, Operate and Maintain, Oversee and Govern) • 33 Specialty Areas • 52 Work Roles (KSA)	(New KSATs in Draft Previous NIST components valid until further publication) Nov 2021
Highlights	• Backed by NIST • Large target audience • Biggest mapping component	• 181r provides a simplified and flexible approach to describe cybersecurity work • Able to be utilized by a wide range of users

Findings

- Identified similarities and differences
- Developed template for evaluation
- Common KSATs across frameworks

Sub-project 2: "Exploratory Analysis of Cybersecurity KSATs from Job Posting Data Using NLP techniques"

Motivation

- Emerging technologies may not be covered by existing KSATs
- Inconsistencies in job posting requirements and lexicon
- Difficult to identify necessary skills across disciplines

Goals

- Identify new KSATs related to emerging technologies
- Update KSATs for CSW development
- Provide a prototype to establish and improve new KSATs across CSW

Identified Job Roles

Different job roles (and related job titles)

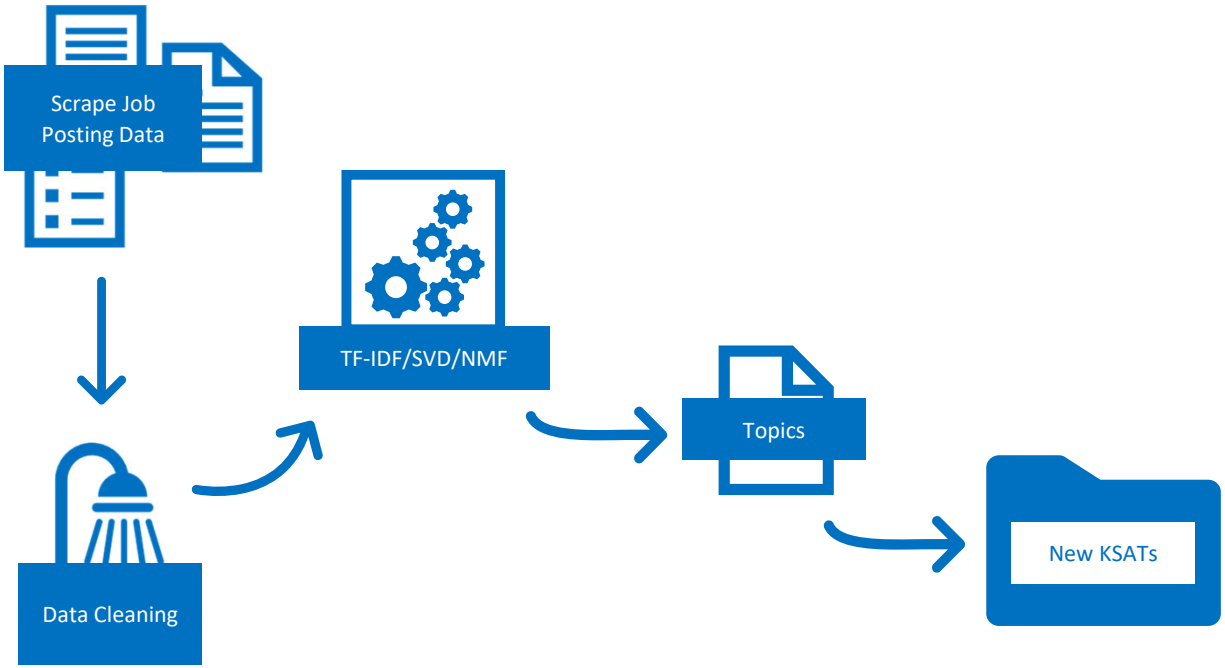
- Entry Level
  - 1.1 Cybersecurity Specialist/Technician
  - 1.2 Cyber Crime Analyst/Investigator
  - 1.3 Incident Analyst/Responder
  - 1.4 IT Auditor
- Mid-Level
  - 2.1 Cybersecurity Analyst
  - 2.2 Cybersecurity Consultant
  - 2.3 Penetration & Vulnerability Tester
- Advanced Level
  - 3.1 Cybersecurity Manager/ Administrator
  - 3.2 Cybersecurity Engineer
  - 3.3 Cybersecurity Architect
  - 3.4 Cybersecurity Director

Identified Information

Level	Job Roles	Related Job Titles
1. Entry-level	1.1 Cybersecurity Specialist/Technician	Information Security Specialist, IT Security Specialist, IT Specialist Information Security, Information Technology Specialist Information Security, Mission Assessment Specialist, Cyber Defense Infrastructure Support Specialist, System Testing and Evaluation Specialist, Network Operations Specialist, Research & Development Specialist, Technical Support Specialist
	1.2 Cyber Crime Analyst/Investigator	Digital Forensics Analyst, Cyber Forensic Specialist, Cyber Security Forensic Analyst, Computer Forensics Analyst
	1.3 Incident Analyst/Responder	Senior Analyst, Information Security, Disaster Recovery Specialist, Network Technical Specialist, Audit Project Manager - Information Security, All-Source Analyst, Target Network Analyst, Multi-Disciplined Language Analyst, Threat/Warning Analyst, Vulnerability Assessment Analyst, Cyber Defense Analyst, Exploitation Analyst, Data Analyst, Systems Security Analyst
	1.4 IT Auditor	Senior IT Auditor, IT Audit Consultant, IT Audit Manager, IT Internal Auditor, IT Program Auditor

Major Cybersecurity Job Titles - Partial

Research Procedure – Outline



Data Collection

- Collect data from online job posting site
  - Develop tool to scrape job details
    - Description
    - Skills
    - Knowledge
    - Experience
    - Requirements

Data Collection (Cont.)

- Tool Developed Utilizing Python
  - Python Libraries
    - BeautifulSoup, NLTK, Pandas, Requests, Scikit-learn, TextBlob
  - Parse search terms
    - Example: cybersecurity, cyber security

```
Trying link #10: https://www.indeed.com/rc/clk?jk=95224602aca3f995&fccid=48ecd52
6e3aa3225&vjs=3
Gathering data about job #63....
Finished with job #63....
Trying link #11: https://www.indeed.com/rc/clk?jk=bf17fca1200822ca&fccid=bc18dc5
0336f5711&vjs=3
Gathering data about job #64....
Finished with job #64....
Trying link #12: https://www.indeed.com/rc/clk?jk=561e5c13c82faa34&fccid=7dc8be9
efe945d3a&vjs=3
Gathering data about job #65....
Finished with job #65....
65 new job postings for Cyber Crime Investigator retrieved from Indeed. Stored i
n Cyber Crime Investigator_results.xls.
Searching for Digital Forensics Analyst.
Obtaining job information..
Scraping job postings...
```

Data Collection & Detail

	A	B	C	D	E	F	G	H
1		link	title	location	qualifs	descr	posted	timestamp
2	0	https://ww	Incident Response /	MindPoint Group, LL(		Company I	MindPoint	2021-07-31 11:29:10
3	1	https://ww	Solutions Architect	RiskIQ5 reviewsCalif(		RiskIQ is t	RiskIQ1 da	2021-07-31 11:29:11
4	2	https://ww	Developer, Digital S	Phillips 66705 reviews		Phillips 66	Phillips661	2021-07-31 11:29:11
5	3	https://ww	IT SPEC (CUSTSP1	US Air National Guarc		DutiesSum	usajobs.gc	2021-07-31 11:29:12
6	4	https://ww	Security Operations	JPMorgan Chase Bar		Working in	JPMorgan	2021-07-31 11:29:13
7	5	https://ww	Sr. Information Sec(	The University of Chic		Departmer	The Univer	2021-07-31 11:29:14
8	6	https://ww	TITLE 32 SUPV IT (	US Army National Gu:		DutiesSum	usajobs.gc	2021-07-31 11:29:15
9	7	https://ww	SOC Analyst	Brown & Brown Insur:		Brown & B	Brown & B	2021-07-31 11:29:15
10	8	https://ww	Task Order Senior I	Peraton96 reviewsBe		Peraton dr	Peraton4 c	2021-07-31 11:29:16
11	9	https://ww	Systems Software I	VayuStreetsboro, OH		Duties:Res	Vayu30+ c	2021-07-31 11:29:17
12	10	https://ww	CSIS Open Source	Citi18,203 reviewsTa(		The Globa	Citi30+ da	2021-07-31 11:29:18
13	11	https://ww	Senior Cyber Fraud	Amazon Web Service		Bachelor's	Amazon.cc	2021-07-31 11:29:19
14	12	https://ww	Cyber Intelligence A	Array Information Tec		Cyber Inte	Array Infor	2021-07-31 11:29:19
15	13	https://ww	Global Financial Cri	Citi18,203 reviewsTa(		The Comp	Citi21 days	2021-07-31 11:29:20
16	14	https://ww	Mid Counterintellige	Prescient Edge Fede(		Job Descri	Prescient I	2021-07-31 11:29:21

Data Preparation - Cleaning

- Prepare text data for analysis
  - Remove punctuations, special characters, periods, dashes, etc.
  - Tokenize text
  - Remove stop words

Data Preparation - Preprocessing

- TF-IDF Vectorizer
  - Cleans Text
  - Creates N-grams
  - Measures the Importance of Terms

```
print(docterm_phrases.head())
```

```
   00 matchingdental  000 00  000 jmu  10 hour  10 years  13grade federal  \
0                0.0     0.0      0.0      0.0  0.000000              0.0
1                0.0     0.0      0.0      0.0  0.124776              0.0
2                0.0     0.0      0.0      0.0  0.000000              0.0
3                0.0     0.0      0.0      0.0  0.000000              0.0
4                0.0     0.0      0.0      0.0  0.000000              0.0

   13level competitive  14 gs  15 va  15 years  ...  years relevant  \
0                  0.0    0.0    0.0  0.000000  ...             0.0
1                  0.0    0.0    0.0  0.000000  ...             0.0
2                  0.0    0.0    0.0  0.000000  ...             0.0
3                  0.0    0.0    0.0  0.000000  ...             0.0
4                  0.0    0.0    0.0  0.044898  ...             0.0

   years security  years single  years user  years work  york atlanta  \
0             0.0           0.0         0.0         0.0           0.0
1             0.0           0.0         0.0         0.0           0.0
2             0.0           0.0         0.0         0.0           0.0
3             0.0           0.0         0.0         0.0           0.0
4             0.0           0.0         0.0         0.0           0.0
```

Data Analysis - Topic Modeling (N-Grams)

```
# Create a document-term matrix with only nouns Store TF-IDF Vectorizer
vect_phrases = TfidfVectorizer(ngram_range = (2,4), stop_words=stop_words_agg, max_df = .95, min_df = .01)
```

```
# Visually inspect Document Term Matrix
print(vect_phrases.get_feature_names())
```

```
bility moderate', 'ability multiple', 'ability natio
n', 'ability new', 'ability non', 'ability outstandin
g', 'ability perks', 'ability preferred', 'ability pre
sent', 'ability prioritize', 'ability sense', 'ability
sensitive', 'ability sound', 'ability switch', 'abilit
y team', 'ability technical', 'ability university', 'a
bility workflow', 'able develop', 'able multiple', 'ab
le new', 'able normal', 'able pm', 'able poc', 'ableid
t bbins', 'abstract ideas', 'academic environment', 'a
ccepts responsibility', 'access data', 'access excel',
'access geo', 'access mail', 'access management', 'acc
ess physical', 'access posting', 'access visual', 'acc
essibility citi', 'accommodation request', 'accommodat
ion requests', 'accommodation search', 'accommodations
accommodations', 'accommodations enable', 'accommodati
ons ext', 'accommodations nsa', 'accommodations religi
ous', 'account holders', 'account information', 'accou
nthealth insurancehealth', 'accounts compromised', 'ac
counts develop', 'accounts numbers', 'accuracy stron
```

```
# Visually inspect Document Term Matrix
print(vect_phrases.get_feature_names())
```

```
['00 matchingdental', '00 matchingdental insurancedisability', '000 00', '000 00 matchingdental', '000 00 matchingdental
insurancedisability', '000 jmu', '000 jmu information', '000 jmu information james', '10 hour', '10 years', '13grade fede
ral', '13grade federal service', '13grade federal service specialized', '13level competitive', '13level competitive candi
dates', '13level competitive candidates current', '14 gs', '14 gs 13grade', '14 gs 13grade federal', '14 gs 13level', '14
gs 13level competitive', '15 va', '15 years', '15 years cumulative', '15 years cumulative experience', '19 precaution',
'200th application', '200th application et', '200th application et vacancy', '30 000', '30 000 jmu', '30 000 jmu informat
ion', '3511 course', '3511 course army', '3511 course army 35e', '35e course', '35e course air', '35e course air force',
'351 course', '351 course army', '351 course army 3511', '37th birthday', '37th birthday maximum', '37th birthday maximum
entry', '40 hour', '40 hour workweek', '40 hour workweek possess', '401k contribution', '401k contribution programs', '40
1k plan', '401k plan careers', '401k plan careers ciphertrace', '50 dd214', '50 dd214 applicable', '50 dd214 applicable d
ocuments', '50 notification', '50 notification personnel', '50 notification personnel competitive', '50 performance', '50
performance appraisal', '50 performance appraisal submit', '80 dollar', '80 dollar investment', '80 dollar investment buy
ing', 'aa vet', 'aa vet disability', 'aa vet disability olgoonik', 'aa vet disability overview', 'ab sanctions', 'ab sanc
tions aml', 'ab sanctions aml fraud', 'abilities distance', 'abilities distance vision', 'abilities distance vision trave
l', 'abilities health', 'abilities health fairview', 'abilities health fairview application', 'abilities shall', 'abiliti
es shall english', 'abilities shall english technical', 'ability ability', 'ability ability concise', 'ability analytic',
'ability analytic methodologies', 'ability analytic methodologies ability', 'ability argument', 'ability argument evaluat
ion', 'ability argument evaluation analytic', 'ability complex', 'ability complex difficult', 'ability complex difficult
problems', 'ability concise', 'ability conduct', 'ability conduct user', 'ability conduct user feedback', 'ability data',
```

Data Analysis - NMF

- NMF (Non-negative Matrix Factorization)
  - Identify topics throughout corpus
  - Topics translated to KSATs

Data Analysis – Output (NMF)

Topic 5

user stories, stories saas, years user, agile product, product manager, saas b2b, cryptocurrency intelligence, stories agil e, dlt expertise, manage different

Topic 8

fraud investigations, fraud activity, investigations work, complex investigations, deep dive, aws fraud, ability high, ad ho c, dive investigations, investigations fraudulent

Topic 34

data analysis, digital data, foreign language, skills ability, diverse teams, applied mathematics, technical report, securit y learn, support professional, time experience

Outcomes

- New KSATs identified using NMF output (e.g., cyber cryptocurrency fraud analysis)
- Updated new KSATs related to job roles (e.g., Cyber Crime Analyst\Investigator)

1.2 CYBERCRIME ANALYST/INVESTIGATOR

Description: - plan, implement, upgrade, or monitor security measures for the protection of networks and systems - create, teste and implement network disaster recovery plans - install firewalls, data encryption and other security measures - performing risk assessments and testing of data processing systems - investigate security breaches and other cyber security incidents. - uses data collected from a variety of cyber defense tools (e.g., ids alerts, firewalls, network traffic logs) to analyze events that occur within their environments for the purposes of mitigating threats.
Related Job Titles: Cyber Analyst, Cyber Security Analyst, Data Security Analyst, Information, Assurance Analyst, IT Security Analyst, Network Security Analyst
NICE relevant factors: - Category: Analyze, Collect And Operate, Investigate, Protect and Defend - Specialty Areas: cyber defense infrastructure support, digital forensics, cyber investigation, incident response, cyber defense analysis, vulnerability assessment and management, exploitation analysis, cyber operations
Desired Knowledge/ Skills/ Abilities/Tasks: - computer networking concepts and protocols, and network security methodologies, - network traffic analysis methods, operating systems, system and application security threats vulnerabilities - buffer overflow, mobile code, cross-site scripting, Procedural Language/Structured Query Language [PL/SQL] and injections, race conditions, covert channel, replay, return-oriented attacks, malicious code - threats and vulnerabilities, authentication, authorization, access control methods, IDS/ IPS, cryptocurrency fraud analysis, Deepfake analysis, language skills
Desired Certifications: CISSP, CEH, CCNA, CISM, CASP, CNDA, CISA, Security+
Related Degree: - BS: Security and Information Assurance, Computer Science, Management Information Systems, Information Technology - MS: Cyber Security, Network Engineering, Management Information Systems, Computer Science
Other Experiences: Cyber Security Analysist, Network Engineering, Software Analyst, System, etc.

Discussion

Contributions

- To develop important tools that continuously update new KSATs
- To provide opportunities to utilize the outcomes and developed tools
- To predict future workforce needs and to fill the shortage of qualified cybersecurity professionals

Discussion (cont.)

Implications

- For job candidates, it gives a better understanding of the emerging KSATs needed for their future careers
- For employers, it provides new skill requirements for recruitment and workforce development
- For institutions, it helps to develop new content for their cybersecurity programs

Future Plan

- To extend the analysis including more data sets
- To incorporate detailed education and training units based on the identified new KSATs.
- To disseminate findings
  - Presenting “Analyzing Cybersecurity KSATs through Text Analysis” at DSI 52nd Annual Conference on Nov. 17-20, 2021
  - Submitting journal outlets such as CACM, CAIS, etc.

Selected References

- National Institute of Standards and technology, (2020, November), *NIST Special Publication 800-181 Revision 1,* https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf
- National Institute of Standards and technology, (2021, April 13), *Task Knowledge Skill (TKS) Statements Authoring Guide for Workforce Frameworks*, https://www.nist.gov/system/files/documents/2021/04/13/TKS_Authoring_Guide13apr2021-508Compliant.pdf
- National Institute of Standards and technology, (2021, March), *NICE Framework Competencies: Assessing Learners for Cybersecurity Work,* https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8355-draft.pdf
- Debortoli, S., Müller, O., Junglas, I., vom Brocke, J. (2016, July) Text Mining For Information Systems Researchers: An Annotated Topic Modeling Tutorial 7-2016, Volume 39 – Article 7
- Lovejoy, C. (2020, May) job-scraper (https://github.com/chris-lovejoy/job-scraper)
- Elliot (2021, May) Word Clouds in Python (https://onebyzero.org.in/2021/05/17/word-clouds-in-python)

Acknowledgement

This research was performed under an appointment to the U.S. Department of Homeland Security (DHS) Science & Technology (S&T) Directorate Office of University Programs Summer Research Team Program for Minority Serving Institutions, administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and DHS. ORISE is managed by ORAU under DOE contract number DE-SC0014664. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DHS, DOE or ORAU/ORISE.