



WeightWatchers

Amateur Hour to Alerting Power:

Overcoming Challenges in Constructing a PII Observability Pipeline

Malav Modi

/

DevSecOps Days Washington, D.C.

October 12th, 2023

/

whoami

- ▶ Security Engineer @ WeightWatchers, Platform Operations
- ▶ Former disinformation researcher @ NCRI
 - <https://networkcontagion.us/reports/>
- ▶ Tennis 🎾 | Retro Games 🎮 | Vinyl Record 🎵 | Travel 🧳

Agenda

- ▶ TL;DR
- ▶ The Rebirth of DLP
- ▶ Evolution of PII Identification Algorithms
- ▶ Building an In-House Solution:
 - Business Context
 - Architectural Design Decisions / Challenges
 - The Future / Potential Improvements

TL;DR

- ▷ Discussing the evolution of DLP, and how it affects PII identification & observability
- ▷ The **trials and tribulations** of creating a PII observability pipeline with given business context

/The “Rebirth” of DLP

Times have changed

- ▶ Cloud-native architectural solutions are here to stay
- ▶ Open-source projects are rapidly improving
- ▶ CASBs just don't fit all the time



“Developing in-house capabilities in advanced analytics and artificial intelligence enables organizations to not only improve their own in-house data-management solutions but also better integrate vendor tools and gain a clearer picture of their data-loss risk, making incidents easier to prevent and contain.”

- McKinsey, 2022

#SecDataOps

- ▷ “...Security teams have to adopt strong data analysis, engineering and science processes from data collection and storage to dissemination and archiving. The goal of SecDataOps is to ensure that data is always finely curated and accessible, and that security decisions are made with high-fidelity data.”

- Jonathan Rau



- <https://blog.lightspin.io/guide-secdataops-vulnerability-management-aws>
- <https://www.darkreading.com/risk/why-secdataops-is-the-future-of-your-security-program>

/ Evolution of PII Identification Algorithms

Traditional Based DLP by Itself is 😞!



Rule-Based Matching / Checksums

- ▶ Regular Expressions
- ▶ Blocklists / Dictionaries
- ▶ Validating data integrity / tampering
- ▶ Conditional Statements



Contextual Analysis is



Named Entity Recognition

- ▶ Some overlap with rule-based methodologies, which require manual intervention
- ▶ Enhanced when using document preprocessing techniques like tokenization, POS tagging, and dependency parsing
- ▶ Models can require lots of training data depending on data annotation requirements

Me: *uses machine learning*
Machine: *learns*
Me:



Custom NER Model Example

The screenshot displays the IBM Watson Knowledge Studio interface. The main window shows a document titled "MQTT.docx" with ten numbered lines of text. Each line contains various entities highlighted in different colors, corresponding to the legend on the right. The legend lists 17 entity types: Actuator, Address, Attack, Attacker, Defensive_Means, Event, Host, Impact, Malware, Networking, Offensive_Means, Sensor, Software, Thing, User, Vendor, and Vulnerability. The document text includes references to CVEs (e.g., CVE-2018-8531, CVE-2018-19417, CVE-2018-18765), technical terms like "remote code execution vulnerability", "stack-smashing attack", and "heap-based buffer over-read", and specific components like "MQTT server", "Contiki-NG", and "MQTT packet-parsing functionality".

IBM Watson Knowledge Studio

Back to Task | Open document list | bogdan.iancu@gmail.com

View Details | Attribute View | Completed

Mention

Relation

Conference

MQTT.docx

1 CVE-2018-8531: A remote code execution vulnerability exists in the way that Azure IoT Hub Device Client SDK using MQTT protocol accesses objects in memory, aka "Azure IoT Device Client SDK Memory Corruption Vulnerability".

2 " This affects: IoT Device Client SDK, Azure IoT Edge.

3 CVE-2018-19417: An issue was discovered in the MQTT server in Contiki-NG before 4.2.

4 The function parse_publish_vhdr() that parses MQTT PUBLISH messages with a variable length header uses memcpy to input data into a fixed size buffer.

5 The allocated buffer can fit only MQTT_MAX_TOPIC_LENGTH! (default 64) bytes, and a length check is missing.

6 This could lead to Remote Code Execution via a stack-smashing attack (overwriting the function return address).

7 Contiki-NG does not separate the MQTT server from other servers and the OS modules, so access to all memory regions is possible.

8 CVE-2018-18765: An exploitable arbitrary memory read vulnerability exists in the MQTT packet-parsing functionality of Cesanta Mongoose 6.13.

9 It is a heap-based buffer over-read in mg_mqtt_next_subscribe_topic.

10 A specially crafted MQTT SUBSCRIBE packet can cause an arbitrary out-of-bounds memory read potentially resulting in information disclosure and denial of service.

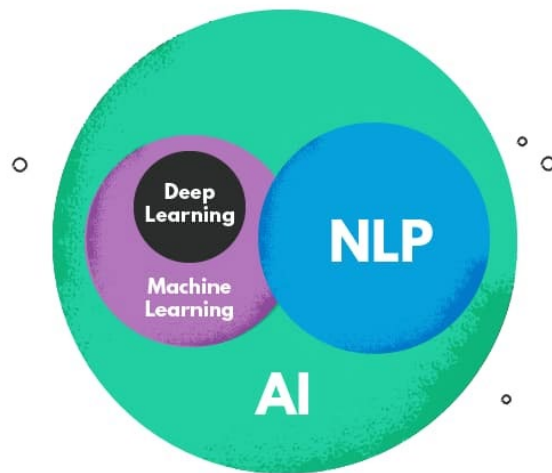
Entity | Mention

Type	Subtype	Role
-	Actuator	
-	Address	
-	Attack	
-	Attacker	
-	Defensive_Means	
-	Event	
-	Host	
-	Impact	
-	Malware	
-	Networking	
-	Offensive_Means	
-	Sensor	
-	Software	
-	Thing	
-	User	
-	Vendor	
-	Vulnerability	

Cookie Preferences

Buzzword / phrase(s) Lore!!

- ▶ Feature Extraction & Selection / Dimensionality Reduction
 - Word Embeddings
 - Word Segmentation
 - Topic Modelling
 - Tokenization
 - Lemmatization
- ▶ Language Models / Frameworks
 - LSTM
 - BERT
 - Word2Vec
 - SpaCY / NLTK
 - Cloud Providers
 - HuggingFace



eBuilding an In-House Solution

Business Context

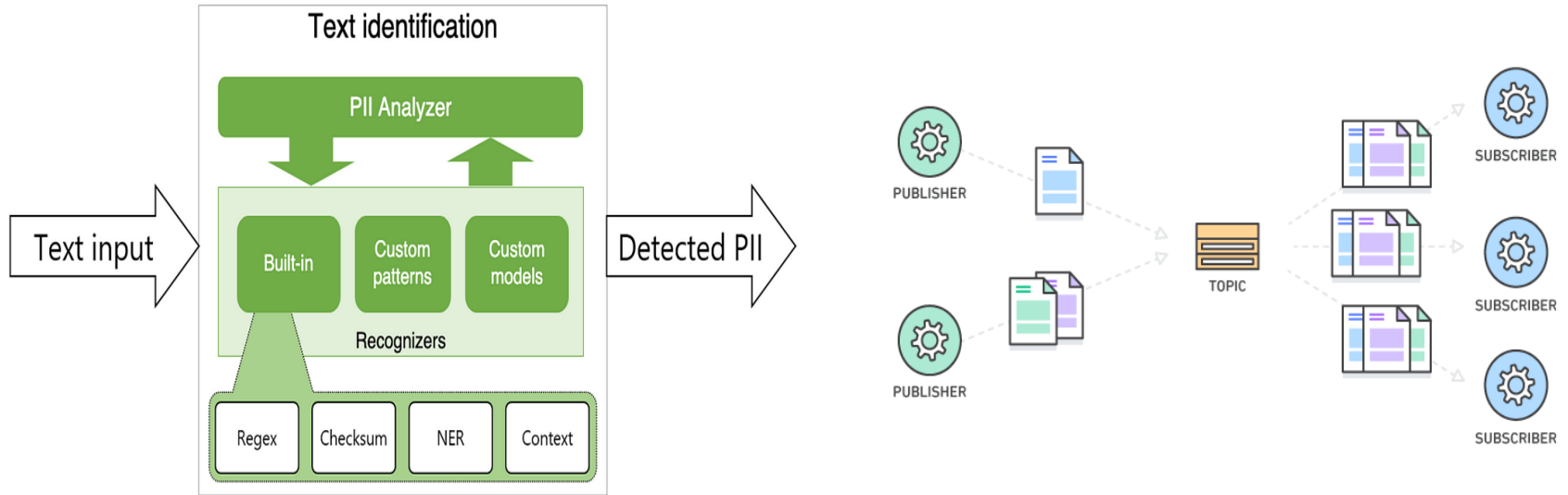
- ▶ Lots of member / subscriber calls about various issues (account & billing, general, q&a, deals, etc...)
 - On the magnitude of thousands a day
- ▶ Beyond vendor integrations, hard to oversee what is being said in call:
 - Need to understand scope of what is considered sensitive for monitoring (PCI DSS, SoX, etc...)

Proposed Architecture



*Accurate Depiction of my Brain's Frontal Lobe

Proposed Architecture (slight detail)



- <https://aws.amazon.com/what-is/pub-sub-messaging/>
- <https://towardsdatascience.com/building-a-customized-pii-anonymizer-with-microsoft-presidio-b5c2ddfe523b>

Tech Stack

- ▶ AWS Comprehend & Translate
 - NLP model trained on vast amount of PII data
 - Using translate service for non-english corpuses
- ▶ MongoDB - (Structured logging)
- ▶ Alerting Stack
 - Pub-sub Model for identification & triage
- ▶ Asyncio - Non-blocking API calls for 10x faster data preprocessing
- ▶ All Containerized via Docker
 - Run through Secure CI/CD Pipeline
 - Deployed on AWS EKS cluster



/Problems

Call Volume / API Reliability

- ▶ Processing thousands of calls via API can be computationally taxing
 - Need to leverage asynchronous / multiprocessing design patterns
- ▶ Accounting for server-side API issues:
 - HTTP 5XX can be really annoying!
- ▶ Structured Logging can be super helpful here!
 - Creating logs that are useful and actionable



```
python show_off.py
2022-10-15 18:12:27 [debug]   debugging is hard           a_list=[2, 2]
2022-10-15 18:12:27 [info]    Informative!             some_key=some_value
2022-10-15 18:12:27 [warning]  uh-oh!                    a_dict={'x': 42, 'y': 'foo'}
2022-10-15 18:12:27 [error]   error!                      what=someClass(x=1, y=2)
2022-10-15 18:12:27 [critical] wtf
2022-10-15 18:12:27 [error]   poor me

#users/home/FROSS/structlog/show_off.py:18: make_call_stack_more_impressive
#users/home/FROSS/structlog/show_off.py:18: make_call_stack_more_impressive
37 def make_call_stack_more_impressive(i:
38     try:
39         d = {'i': i}
39 |         print(SomeClass(d['i']), "foo")
39 |     except Exception:
39 |         log.exception("poor me")
39 |     log.info("all better now!", stack_info=True)
39 |
39 | locals = {}
39 | d = {'i': i}

KeyError: 'y'
2022-10-15 18:12:27 [info]    all better now!
Stack (most recent call last):
  File "/Users/home/FROSS/structlog/show_off.py", line 40, in <module>:
    make_call_stack_more_impressive()
  File "/Users/home/FROSS/structlog/show_off.py", line 37, in make_call_stack_more_impressive:
    log.info("all better now!", stack_info=True)
```

“Interesting” transcription format

- ▶ Voice transcriptions have numbers and special characters as the grammatical representation
- ▶ For a majority of pre-trained models, this is an edge case when it comes to training data!

1. One	11. Eleven	10. Ten
2. Two	12. Twelve	20. Twenty
3. Three	13. Thirteen	30. Thirty
4. Four	14. Fourteen	40. Forty
5. Five	15. Fifteen	50. Fifty
6. Six	16. Sixteen	60. Sixty
7. Seven	17. Seventeen	70. Seventy
8. Eight	18. Eighteen	80. Eighty
9. Nine	19. Nineteen	90. Ninety
10. Ten		100 - Hundred
		1,000 - Thousand

AWS Comprehend Side by Side

Personally identifiable information (PII) analysis mode

- Offsets
Identify the location of PII in your text documents.
- Labels
Label text documents with PII.

Analyzed text

Hello Zhang Wei, I am John. Your AnyCompany Financial Services, LLC credit card account 1111-0000-1111-0008 has a minimum payment of \$24.53 that is due by July 31st. Based on your autopay settings, we will withdraw your payment on the due date from your bank account number XXXXXX1111 with the routing number XXXXX0000.
Customer feedback for Sunshine Spa, 123 Main St, Anywhere. Send comments to Alice at sunspa@mail.com.
I enjoyed visiting the spa. It was very comfortable but it was also very expensive. The amenities were ok but the service made the spa a great experience.

▼ Results

Entity	Type	Confidence
Zhang Wei	Name	0.99+
John	Name	0.99+
1111-0000-1111-0008	Credit debit number	0.99+
July 31st	Date time	0.99+
XXXXXX1111	Bank account number	0.99+
XXXXX0000	Bank routing	0.99+
123 Main St	Address	0.99+
Alice	Name	0.99+
sunspa@mail.com	Email	0.99+

Standard

Personally identifiable information (PII) analysis mode

- Offsets
Identify the location of PII in your text documents.
- Labels
Label text documents with PII.

Analyzed text

hello zhang wei i am john your anycompany financial services llc credit card account one one one one zero zero zero zero one one one one zero zero zero eight has a minimum payment of twenty four fifty three that is due by july thirty-first based on your autopay settings we will withdraw your payment on the due date from your bank account number xxxxxx one one one one with the routing number xxxxx zero zero zero zero customer feedback for sunshine spa one hundred twenty three main st anywhere send comments to alice at sunspa at mail dot com i enjoyed visiting the spa it was very comfortable but it was also very expensive the amenities were ok but the service made the spa a great experience

▼ Results

Entity	Type	Confidence
zhang wei	Name	0.99+
john	Name	0.99+
one one one one zero zero zero zero	Credit debit number	0.99+
one one one one zero zero zero eight	Credit debit number	0.99+
july thirty-first	Date time	0.99+
xxxxxx one one one one	Bank account number	0.99+
xxxxx zero zero zero zero	Bank routing	0.99+
one hundred twenty three main st	Address	0.99+
alice at sunspa at mail dot com	Email	0.99+

Janky

AWS Comprehend Model(s) Complications

▼ Application integration

API call and API response of ContainsPiiEntities API [Info](#)

API call

```
1 {
2   "Text": "my bank is one two three four five four
3     four six eight seven with email coolio at
4     gmail dot com",
5   "LanguageCode": "en"
6 }
```

Copy

API response

```
1 {
2   "Labels": [
3     {
4       "Name": "EMAIL",
5       "Score": 1
6     }
7   ]
8 }
```

Copy

▼ Application integration

API call and API response of DetectPiiEntities API [Info](#)

API call

```
1 {
2   "Text": "my bank is one two three four five
3     four four six eight seven with email
4     coolio at gmail dot com",
5   "LanguageCode": "en"
6 }
```

Copy

API response

```
1 {
2   "Entities": [
3     {
4       "Score": 0.9995856285095215,
5       "Type": "BANK_ACCOUNT_NUMBER",
6       "BeginOffset": 11,
7       "EndOffset": 60
8     },
9     {
10      "Score": 0.9999871253967285,
11      "Type": "EMAIL",
12      "BeginOffset": 72,
13      "EndOffset": 95
14    }
15  ]
16 }
```

Copy

1st "Pass"

2nd "Pass" (For Triaging)

PII (but foreign)

French

- ▷ Rule Based Match / NER #1
-

Portuguese

- ▷ Rule Based Match / NER #1
 - ▷ Rule Based Match / NER #2
-

Spanish

- ▷ Rule Based Match / NER #1
-

Swiss

- ▷ Rule Based Match / NER #1
 - ▷ Rule Based Match / NER #2
 - ▷ Rule Based Match / NER #3
-

Other

- ▷ Rule Based Match / NER #N

Life Hack: Just Translate!

- ▶ Creating custom programmatic solutions to account for N languages can be difficult
 - Language identification mechanisms come in clutch
- ▶ May lose a bit of context, but when scoped down to various entities, the context becomes general (credit card numbers, bank account numbers, etc...)



Real World Example

<p>Detection for Conversation ID: 'a5b8c4f4-219e-4dd3-9d58-8a652874e649' with potential label(s): SSN</p> <p><input type="checkbox"/> HIDE DETAILS</p>	PM
--	----

CUSTOM DETAILS	
communication_id	a28bd5d7-f779-4c30-93b8-fd1dbca4105e
conversation_id	a5b8c4f4-219e-4dd3-9d58-8a652874e649
labels_transcript_pii_result	[<div style="border: 1px solid black; padding: 2px; margin: 2px;"> <p>"Name": "BANK_ACCOUNT_NUMBER", "Score": 1</p> </div>
offset_transcript_pii_result	[<div style="border: 1px solid black; padding: 2px; margin: 2px;"> <p>"BeginOffset": 3866, "EndOffset": 3913, "OffsetDifference": 47, "Score": 0.9998678220375061, "Type": "SSN"</p> </div>

<p>051b5PnXubpNlFjQk3jW-KYIs3CYqEEXKsYxxQcTy3xg6JMFJmC3akazqhW07IZRS5nTYFN0HTYLCrUoMxDwK_KRp7TY7-8Bkqk4iyVQk9f6wjBkgGQuidL1Xcu3103mWZGA1YnLJXpPsKn4ZTvTq1m1se2NmehdoU12XRvTDRpwbJ6IehZ2iPxyV1r1Bvnr9nZHS01rbtclGuxg7LorX4WYGPdLZ1-</p> <p>Uih_uOaFtK37bVWmuo7YrK3xtp013Lr1Ej1918YRKTWcKymZ43Jn1j9Q0V2PC6V6m_wUTZ1_NKbeJ6WpM_66wFt0EAEQd9gkYao4Wq7T_KgnTwmFrP_uSetUCBwsGURKeME_BcYwL0LUQd0w9n0I4czLJZ0T7Wk-xdVaeTH6U1675BKB9h3EtACUd_MpZwJf85zapeMYr76k3JmV1yUk8s-gzWuKz-LDUALpT16mQAr_4PRBh8mC0gm1_PmoQ0k2RvZ12_DVUGL0meEMRBKPrfRdkiYVJ7SuaJAYdu5LDB1IcV8IenJyBySp98Htatu2S_u-NnLdLrFmPbVc9Pa7A0E8M8daNahuL1GcVq9jnzLgxsMPls8nVc_eeZ0DA_UglUeGVKXcq8lF0pV7b-nRAGaa15m3kT13174Byy7FbsN4FnoPn1n0JohWjYXtqjPynWmYnqPvZ2XppM1LooYwFHNh08a7v9rFVE36w6sL-1L6XfQp-mGuvx211Hj4p4dJ_c5v5HyLWf-Pm0UCCZ-</p> <p>K5ENRHFHrF1FbZ1n2EXu0956dAbWJVT8T1oyf0w9jCqXKEXVks5Jflet13714E4fPa5SM19f3u4MFC6G7vZ2A1-140XFRZtUJ5h6gs-w8E1CZQZLZoz27443TFM02wL_VcderLqXN1LQEDm142L5H1Yr6vLX01-g1R0BZ85K_rbfxylyZ68mQYfHYJ3wL-4_DmC7Pb0x8m1PHYB6A-196cmVuuEwH9pAodLhF0k0686YfegGr8aFXG5YFG68202vLD0R6EYfC10E_Grgh1VZ70hM935w08sBmW70UcJjCNR8F5AcwLies1wJ0gh125gDwJLEr68dAxx-EM40d9r9Kq0Jf0M9J17ZPE4WYegutL-dc3lRqI0zISd0bd-1EP2cWJp_cD7Kk02hJ1z3z_w8g0MmnuF9g6wurK8X8mP-gmshQd_r_XsbEG7M8q0s2v_1PmW7g8U5SMOYm2ifz6182rIGC0Kf1n5tYH89SN-DVY1Z1TSMGP009cCpWp_dod1L1M2z3q1c2C08H8FR5V0yVg6D10EF7ed0b9KcIG11JW9FN5VM4M7Zynxok33L0nfnNltgKNXfzyRvL1Luek0XU6G1P38LuhDz8Nk590P8mPzWmP6zC155wJ1FCvd6350z5dc2J50AR5M0gm8fRhy14vK042cuDKM-x0z2tERJx6GH418r-8u1U770GcN7FNbkqC1-Ye125rYEPZLCP1LGR3AZw8Rq1MmN4M35_UBDYdsuYnu0mBfkoz-HVku1INWb8Hm5sy-INP0puyNRCVfY3h2t7N1PLH7M15m7r9ek80J290z5FuPeMwIAzs18m8u1RrMtneg3C1C50x9331Xj7xvJenLaTPY-N1U9A0X28-ybgh75_yxkwbx2C1GP3k0h315E13DF8MEBmG00hYQ3mU8--b0y0E-AU2cekewoQnRAJUPrH255p5mrmkpl_WkR5nPDnJK-BMUMDM1r014U1pv55G5FTx0bVhJyYLaKxuuqeGkEkkxArNoi77wXXdBNx0sA9dGhA81Gpm_pumkqV0ydKARsdMk8mXhGL9IP9PpYsvY2nu51LpxnLkzkkL2VfooZVJAShZQR_F30wF6LCP8s1zeL6woehH1Rn7y6_hNriYDvkjFvHP_NX6FzWTPqUE8cELSKLAFEM-PMNQ01PuaibvE5_930x0laI8x1AnL0ooeFue7fUgw9M1PEBmfbnF0a_pudyr8P513tHz-1w6REX01jKEJNZ3j0HXkYV191b7dC6z6rYMDk89EK4FuzDqCBU_00XK5GAB9LusNxxzn36wJ6AbCnYgv9Y9TE1_CyJ3UbuXwdI0qeLk15x2fE_02yVjQh4CZK6J5t43N1Bl6gJ51QW0NdyVcs_ZMS506h8510aY610p-1pZKrpjVKh8YRrP5AHmNKR9v-Rdl0z9h085Dxt8R040sJ2n85Z4f-rtWJ1166vzX0fwgr_Yh0wQd0b0a8-M_Ud21q6pYRThUw7X0Gn0L9Yrann0mgzRzBtQjN8HegAwAmh3Z028140FrfYtCaZ0GujYjE02M6s711Yc0085mTGZy46WzQmY5_THEWj0z26_BhC7c25R02CCZf1R_h60ArfZ76fw_srz2f8v9y85PctqR0a6J8vJijauX38M119d0821j0b5jnbjy0M0Vns4c8r1Cn-632U9Rradu0cpsy6tjUjw0w-vbZ1REWk-wX-3C0F6f78R0arjBkF1JjpdZ2Y0B6cgy2vYe8071S1j-tj78Q019U9P7F8Cz24L0xbEP1C1-</p> <p>f-tK1TH7R5lddeP7LAlq0sJ4R41KH292as83_ZPXK0REEM8fdeYGen5_c0LkPv611a5Lu1fACM3G36tAX5enP7200v65SMVAmqmu_Fc9N10FbC454P86d3crYfGd17NhuFaj19a2_L4xG42mhkT4jaxDEferqu08H9FeusFh148HfkmwDhYFVhGcbj14qXpPbn91w8s8du57y9U1Q1D00w0h8SH418Ac_ZL3qozqfXiLl0AbDjpr1MK-AhQ10T742940yQGVZ2Hv3t5XwoZLum4N-0E6F51rUBPQY0E1NXXPxeKw2W7wqV0Wb2enTQw8CBA1VH10b2Zvz0E3JukLc5ZDZ0R1_E3kUeh8D_WEDAS5Wz4_Dnv8E23275knks_6873MkG8JmWkTGu5StHyK35tN0PyaIk9_UjTjrbj8VQV0UM9M90e6AYTEmgk81046R9p06F4R2DNE6Y5F3JdU3-52M4b4x9sLqY7HvN7rFmN-108yrzLSMV0U_170w28Vpaw8HE0cAgN2K6u70Jn9F4-as4K_LGfEb4VFD0_BBHM24BzYcN6PYg6p57na2VJ1Bo6cv-77fchZ2j7Yp5izrZ291h1tkT6D-5Zvp5Z5yZTD_pHJNyDe_XN0wzPeZab7hrp9a6Y21C_5QeRdWbMhM5m7z0t831B9q6Vx7E69dBN60aRNN650eP0N1_S0_M05Vhyo4DvPy_T7UE1z2wL20W6d1gR181deuV1_3dyJmHAcA6q7-r-NBNCCIrt_DCPBwXj rXkLaYnJf0_5X7fdvPM_dhF4jzntymawv0P4o-QKHJjhw421Le1McPSTXBAAT1qb8r08gFPmJ3185WpsQe6G1K0oJdxtdy0ZuZT_CAS1tbaE3A1mYrRlRnLamjJ30CEK5E2eg-TekbFR-D9Xk6e6LpY6B15P31Y4fYRZIG5g6HJ4cc23R1W-CC18yq8MNP2e_mbp48V5yve1G303p5Fb0v16Av150dwJ2H017E12_jd2YfEd1cPvLeJ8eVkmh11wDYR5510gW3dgcdficahJMu_61VG05Ekhqa-0A2B01_S1GAIS-lap0z27N4BzLu0200wycDPYLBWtWAG3AZ1v1z9m8f0rW6W1JhKxmpC1Nz1URKjD3EdarDFKCS1_SbpyZ1P5540xa8N6mP5EPf9Y0LxwCtXv0x0c9qRT1k1C1Vn3zA2kqg1A1S_cjVGA23s-rs48dXzax2R4X041AF1N1DE2cpMw8PAn0z5vY1E8ZHR-bpFvwt1X0p8m1w8uA9J3zLHMLW41Ucg8L0-s8f9rY_LNq7M9Pm9590L_52g9Z9g-BMwY1Z558j8KfPuhIHZ00h11P0KFC3L1Dg11r1f1RqQ13oFtUkUcXDU6Ej0Iaz95ShPMU7PznTAs7rHak1Q0k1hb1hA4-01jGd72MjM0t8B-z6shzk2V8c-c85mme1_abd7TPZK-pl5AK0MhkK30xvJdn3eM58jBu4k1f17z3f4Gj-rfF5jPmPK8M01cN8qKECn8zL1-fYp-2Zx2_tEYzKsp1PbX2N6kme_Ise_r57KTP6E3p5AyA1jYyBu86d2VQ1U1B5J1j5a889A00TWBInFmH3EAC4Csk48CB9hJqFRUHMz1eWzV1nQpT04KVL1E08bJyWEN3MVPz6f8hkhKc1X8N00525YTV00akwZUL_K5X247gwtCHZGHf6eEeZUdIBRSy5zjMTR14o5z-peespN1Xq1z5bTcN4Cj1Ujw9R4fC1SAG5009h4Mc4g15X6rdU78j417NEXdPug8b0197928farMjH7_Pp243NC_1H1_U1J0du0v5Y1LH29Mk1m6Ld8N8JH8CAMFU8_X1kcxZwaA218E0JnxDh0p1m1rd9e0f8ebvbw3cETaxF7zEMJ5emN86H0d67G6eMhQ1W0TrnyjK0E5SN-x81N0090ChtJJD05Y1CekJ0c18F9IKIN1rT46p104eIrU09H285P4v68DAm49A5a14BBFAFtYU8qF74K572QNMkU1fazC_kXUSpm1250sGM10Y1VMvM46AvHP5M2GZfInqAsXv0pJy1LwWc80Nyh8hLk7v6ZHD2g0X1B0d-RXFB-L2RH--56fPdeJepY8bqk7_L7yul1qJGoZrWkjE08Ldd1f3ymme_68Jnu5V81bK0JnFdtVfGP2sTCw9-1L1H-GhC31Pc1mF1-TsnJdx_6Ydsks3P00D_16HfAA0a4uKn0yeyz185VnQ7YcANBmWpKcPEP1JkrapZbaqwg48x95rYh46SRMK7P-ycnB1234Ceb07--kt_771Mc6TFVWF5JrE3HfARUe6870a36Eh1cHrBh-Z6H0P6yEcmWFtXzujk18-1aCnnTQ17fF6cE2ts4jx0J7W9H8Xr2u6WvXw35G1A6m--lVPZ1A1HJfjBvR10qup976V1c9p95YPr8E40dmg8R1R16vL8Rz5s--PfbhoAfn2519eZ2n2sCEdUev37Vg84v07-MBErEN3KPNKcCV5qP0tv9j-1B8X00h1bWdMgfsW0PjVzrd_1zrxXdnNB05p5wAbThJVEou31cWUPM0r73U2ectc2jpm1-tfBuoE050-BWKRH5t8qK8m661Cx7E5Z9X99e7Yr6Y9yXqR3959dZalyLzheMcrZBR115aPmZNPmKorXgHmgIep088LrzaZ2x10T-</p>	<p>translated_transcript</p>
---	------------------------------

General Information

Encrypted Transcript



Potential Improvements

- ▶ Fully event driven / serverless
 - Current architecture currently relies on some API polling -> semi state management
- ▶ More automated actionability / remediation
 - AWS Actions
 - Custom Webhook
 - Flagging / Deleting affected calls
 - Visualize metrics / findings in operational dashboard
- ▶ Relying more on open source
 - Microsoft Presidio, Octopii, EarlyBird, etc...
 - Dependent on having cleaner transcription format, which will help reduce cost



WeightWatchers

Thank You!



WeightWatchers

Q & A