From Lure to Alert:
Implementing Canary Tokens in M365

Ryan O’Donnell

February 12, 2026

Day 12

11:47 PM

Alert! Outlook Canary Access

Agenda

1. Problem
2. Build
3. Results

whoami

- Ryan O’Donnell
- Senior Security Engineer @Microsoft
- Pentesting, Purple Teaming, Red Teaming
- OSCP, OSEP, GREM, GCFA
- Disclaimer*

Cybersecurity Reality

136%

Increase in cloud intrusions
in 2025[1]

"Attackers are bypassing the endpoint entirely… operating purely in the cloud."
– Florian Roth (@cyb3rops)

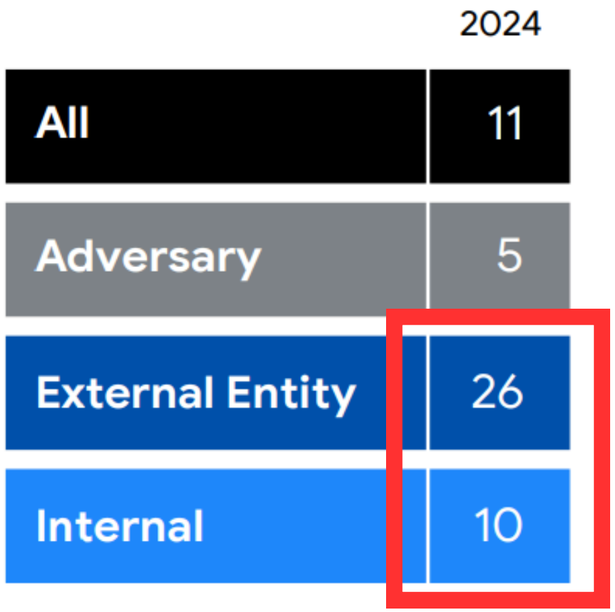
[1] CrowdStrike 2025 Threat Hunting Report

Proposal

Canary Tokens can provide effective **early alerting** on attacker **post-compromise** activity in M365.

Dwell Time

Median Dwell Time by Detection Source, 2024



The median adversary notification time was just five days, while external partners notified in a median of 26 days. This discrepancy is not surprising given that the vast majority of adversary notifications originate from extortion actors who benefit from monetizing intrusions quickly.

16 day difference!

[1] Mandiant M-Trends 2025 Report

Post-Compromise

Real-world example of APT post-ex behavior

> Additionally, Microsoft observed Storm-2372 using Microsoft Graph to search through messages of the account they've compromised. The threat actor was using keyword searching to view messages containing words such as username, password, admin, teamviewer, anydesk, credentials, secret, ministry, and gov.

[1]Storm-2372 conducts device code phishing campaign

Real-World Implementation

Building a native Outlook Canary

Existing Approach

- Relies on links being clicked
- Alert on **resource access**, not interaction
- Challenge:
  - Remove the “link click” dependency
  - Utilize native telemetry

Generate Lures

Craft email with known search terms of interest

Topics	Keywords	Context
 Creds	[pass] [admin]	Shared system
 Finance	[invoice] [wire]	Payment
 HR	[payroll] [w-2]	New employee

Research

Hidden mail folders

The default value of the `isHidden` property is `false`. You can set **isHidden** only once when creating the mailFolder. You can't update the property using a PATCH operation. To change the **isHidden** property of a folder, delete the existing folder and create a new one with the desired value.

Hidden mail folders support all operations that are supported by a regular mail folder.

- Deploy via Graph API[1]
- **15-min** hide delay

[1]Soteria: Hidden Mailbox Folders

Hidden Folder, Still Searchable

Pilfering inbox via GraphRunner[1]

```
PS /opt/tools/GraphRunner> Invoke-SearchMailbox -Tokens $tokens -SearchTerm "credentials"
[*] Using the provided access tokens.
[*] Found 9 matches for search term credentials
Subject: Info | Sender: finance@azpurple.com | Receivers: Adele Vance | Date: 06/28/2025 17:59:30
| Message Preview: Credentials Password Information ...
================================================================================
Subject: Dev Environment | Sender: /O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPD
LT)/CN=RECIPIENTS/CN=E286EAE0D96F4FC89B9FD61675ED286F-903988D0-C8 | Receivers: ippsec | Date: 03/0
9/2025 19:31:32 | Message Preview: ... Below are the temporary credentials to the developer enviro
nment. User: mark Pass: MDRisTh3best! Host: 192.168.1.12  Thanks, Adele ...
================================================================================
[*] Do you want to download these emails and their attachments? (Yes/No)
Yes
```

[1]GraphRunner: Graph API post-ex toolset

Monitor for Access

Use MailItemsAccessed to investigate compromised accounts

A compromised user account (also called an *account takeover*) is a type of cyberattack where an attacker gains access to a user account and operates as the user. These types of cyberattacks

```
CloudAppEvents
| where Timestamp > ago(3d)
| where Application has "Exchange"
| where ActionType =~ "MailItemsAccessed"
| where tostring(RawEventData) has CanaryMessageId
| order by Timestamp desc
```

Alert and Triage


```
"AADSessionId":  "00401e29-2d0f-1197-87ce-6b4235eb9041",
"APIId":  "00000003-0000-0000-c000-000000000000",
"ClientAppId":  "d3590ed6-52b3-4102-aeff-aad2292ab01c",
"IssuedAtTime":  "2025-04-20T15:59:18",
"UniqueTokenId":  "0jpPuyxmikOdByzOaIcXAA"
```


1) Alert
MailItemsAccessed operation (UAL)

→

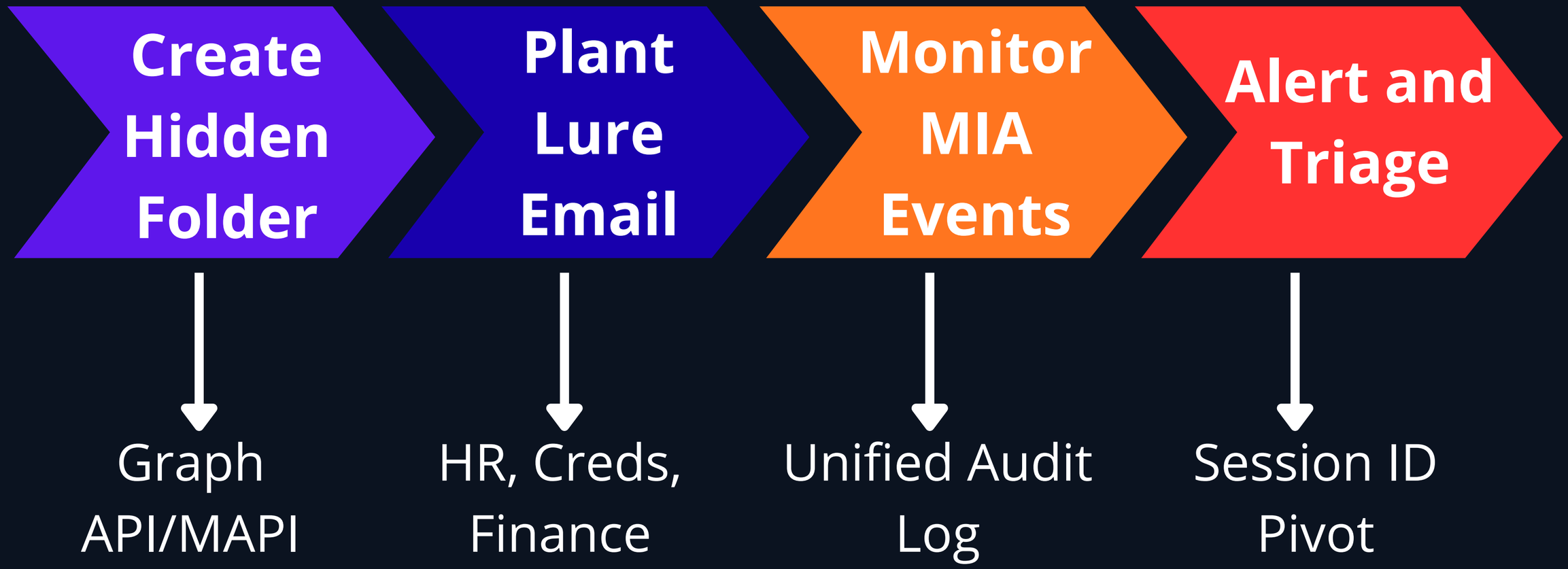
2) Identify
AADSessionID
UniqueTokenId

→

3) Scope
UAL operations
Sign-in logs

[1]Entra Linkable Identifiers

Overall Process



Scope & Coverage

✅ Detects

- Canary item access
- Search-driven mailbox collection
- **Bind** MIA Events

❌ Doesn’t Detect

- Exfil that never accesses canary
- Compliance eDiscovery
- **Sync** MIA events

⚠️ Prerequisites

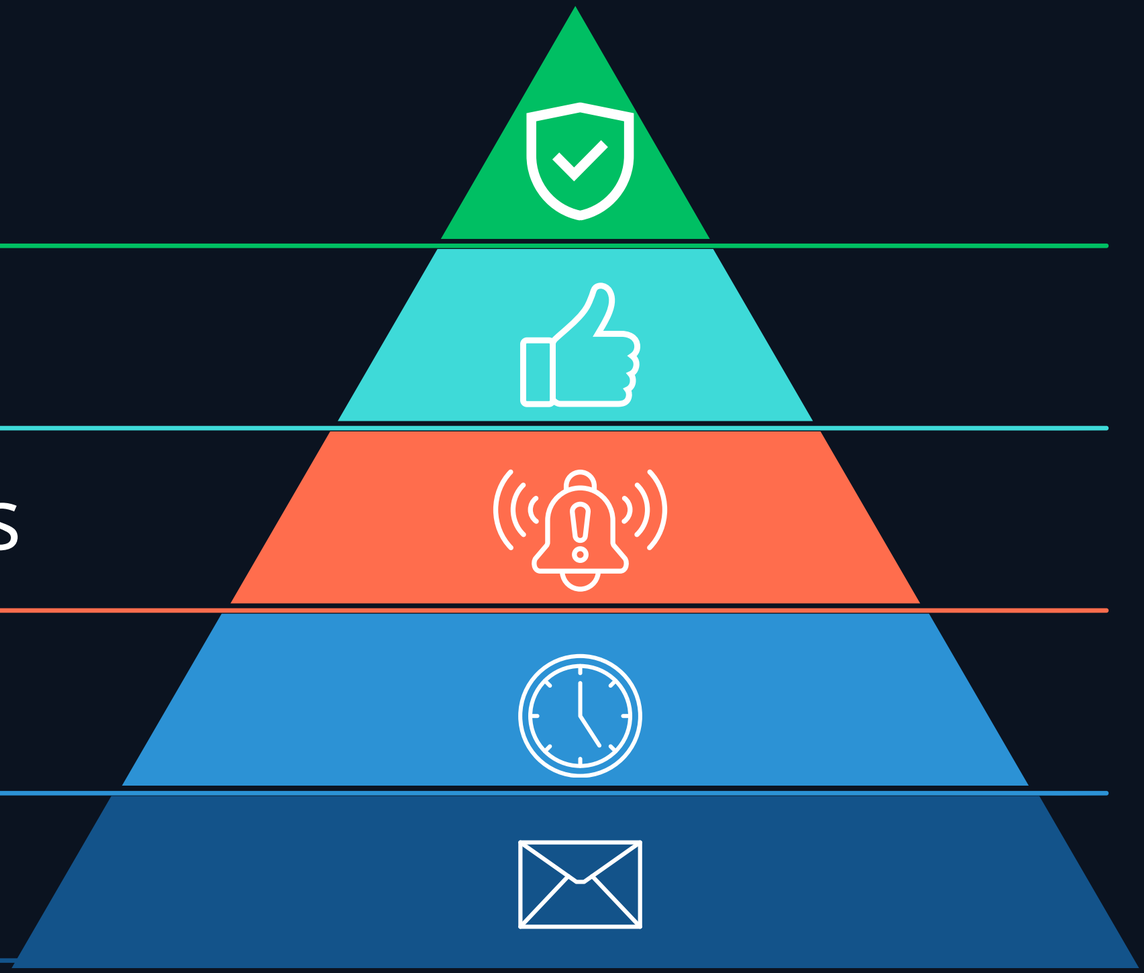
- Unified Audit Log retention, Mailbox auditing enabled

From Design to Deployment

Does it trigger in production?

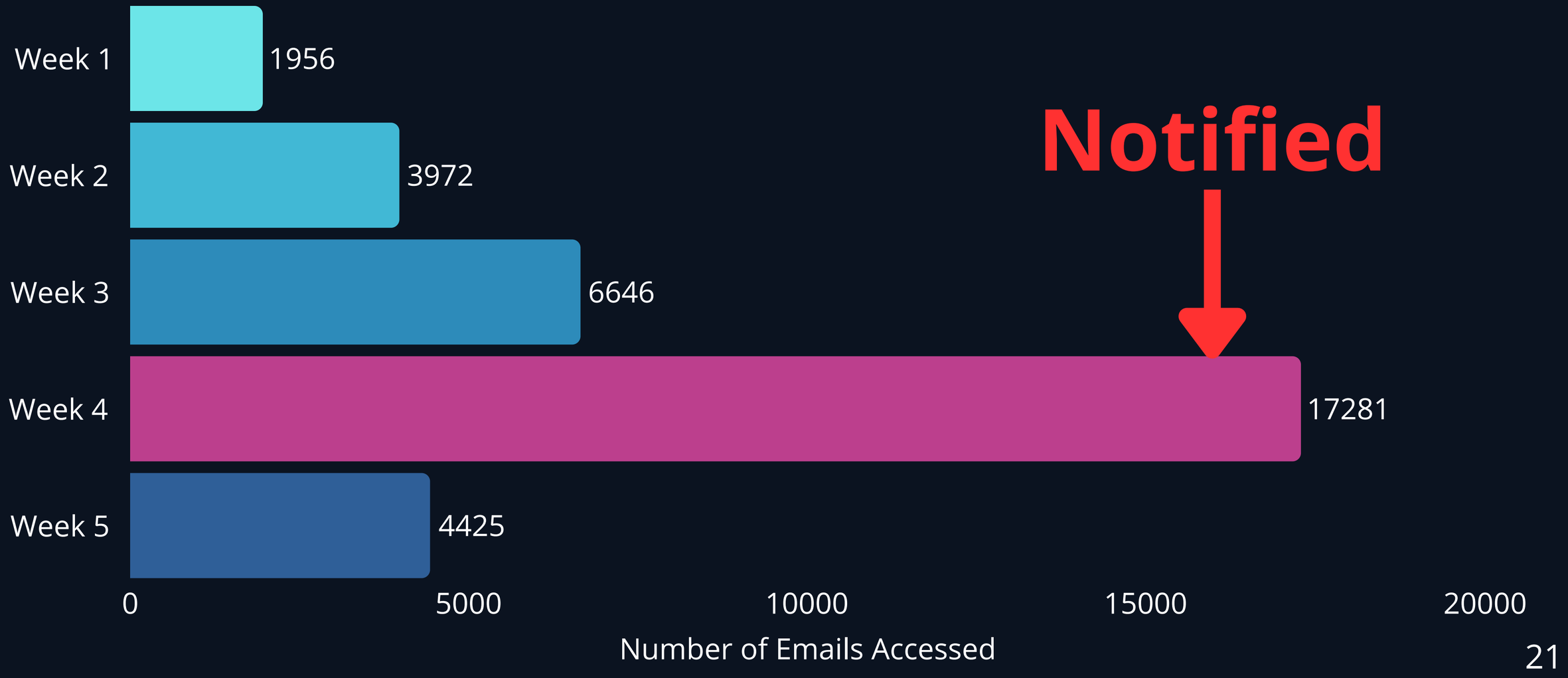
Production Pilot 1

- 1 Real Incident
- 2 Benign Alerts
- 3 Canary Interactions
- 3-month Duration
- 50 Outlook Canaries



First Canary Trip

Events by Week



Unexpected Win: Insider Risk

Canaries trigger on **collection behavior,** not identity

Search → **Export** → **Exfil**

PRESS RELEASE

Former Google Engineer Found Guilty of Economic Espionage and Theft of Confidential AI Technology

Friday, January 30, 2026

Pilot 2: What We Tuned Out

Active Canaries

payment

password

admin

root

credentials

api keys

payroll

SSN

Tuned Out

invoice

wire

What They Actually Searched


```
{
  "CreationTime": "2025-11-17T18:23:08.0000000Z",
  "Operation": "SearchQueryInitiatedExchange",
  "Workload": "Exchange",
  "UserId": [redacted],
  "ClientIP": "38.69.8.29",
  "QuerySource": "Email",
  "QueryText": "invoice"
}
```

```
{
  "CreationTime": "2025-11-17T18:38:52.0000000Z",
  "Operation": "SearchQueryInitiatedExchange",
  "Workload": "Exchange",
  "UserId": [redacted],
  "ClientIP": "38.69.8.29",
  "QuerySource": "Email",
  "QueryText": "wire"
}
```


2 searches. Total.

invoice

wire

The Lesson

Assumption was wrong:

Attackers **don’t always behave** like **pentesters.**

Tuning is a bet. Sometimes you lose.

Closing Summary

Key findings and results

Takeaways

1. Canaries catch collection external OR insider
2. Tuning is a bet. Sometimes you lose.
3. Native M365 telemetry can be enough

Questions?

Contact:

linkedin.com/in/odonnell-ryan

Feel free to reach out!