AIツールの人気は悪意あるキャンペーンの好機？

目次

要旨

Siftが発表した最新の「Q2 2023 Digital Trust & Safety Index」によると、ユーザーの78%が、AIツールを悪用した詐欺について懸念しています。ChatGPTやGrammarlyを標的とした最近のサイバー攻撃に鑑みれば、彼らの心配は杞憂ではないかもしれません。

ブランド保護とフィッシング対策の観点から、WhoisXML APIとBayse Intelligenceはこのほど共同で調査を行い、「2023年最高のAI生産性向上ツール」の人気に乗じている可能性のあるサイバースクワッティングやフィッシングのプロパティを明らかにしました。

今回の調査の結果、以下を特定することができました。

- 人気のAI生産性向上ツールの名称を文字列に含む2,003個のドメイン名
- 信頼できるクラウドプロバイダーのインフラ内に潜伏しながら、人気がある複数のAIツールを標的にして活発に動き回っている1人（または組織）の脅威アクター

パート1：WhoisXML APIの分析

DNSにおけるサイバースクワッティングプロパティの検出

当社ではまず、ドメイン名の帰属を確認できるAI生産性向上ツールをピックアップすることから調査を始めました。37のツール開発者の公式サイトで使われているドメイン名をBulk WHOIS Lookupにかけ、ドメイン名登録者が以下のデータポイントのいずれかを示した8つのツールを選びました。

ツール	公式サイトのドメイン名	登録者データの種類	WHOISレコードのデータ

AgentGPT	agentgpt[.]reworkd[.]ai	Email address Name	contact.me.reworkd@gmail[.]com Reworkd AI
Bard	bard[.]google[.]com	Organization	Google LLC
EmailTree	emailtree[.]ai	Organization	TS Holding
Motion	motion[.]ai	Email address Organization	domain-groups@hubspot[.]com HubSpot, Inc.
ProWriting Aid	prowritingaid[.]com	Organization	123-Reg Limited
Runway	runway[.]ml[.]com	Email address Organization	domain.administrator@bankofamerica[.]com Bank of America Corporation
SaneBox	sanebox[.]com	Name Organization	S***** R***** SaneBox
Slidesgo	slidesgo[.]com	Organization	Freepik Company S.L.

注：*プライバシー保護のため、sanebox[.]comのWHOISレコードに含まれていた登録者名は一部非表示にしました。*

脅威アクターが今後のキャンペーンのために8つのツールのいずれかに狙いを定めている可能性を見るため、以下をキーワードとしてDomains & Subdomains Discoveryで検索を実行しました。

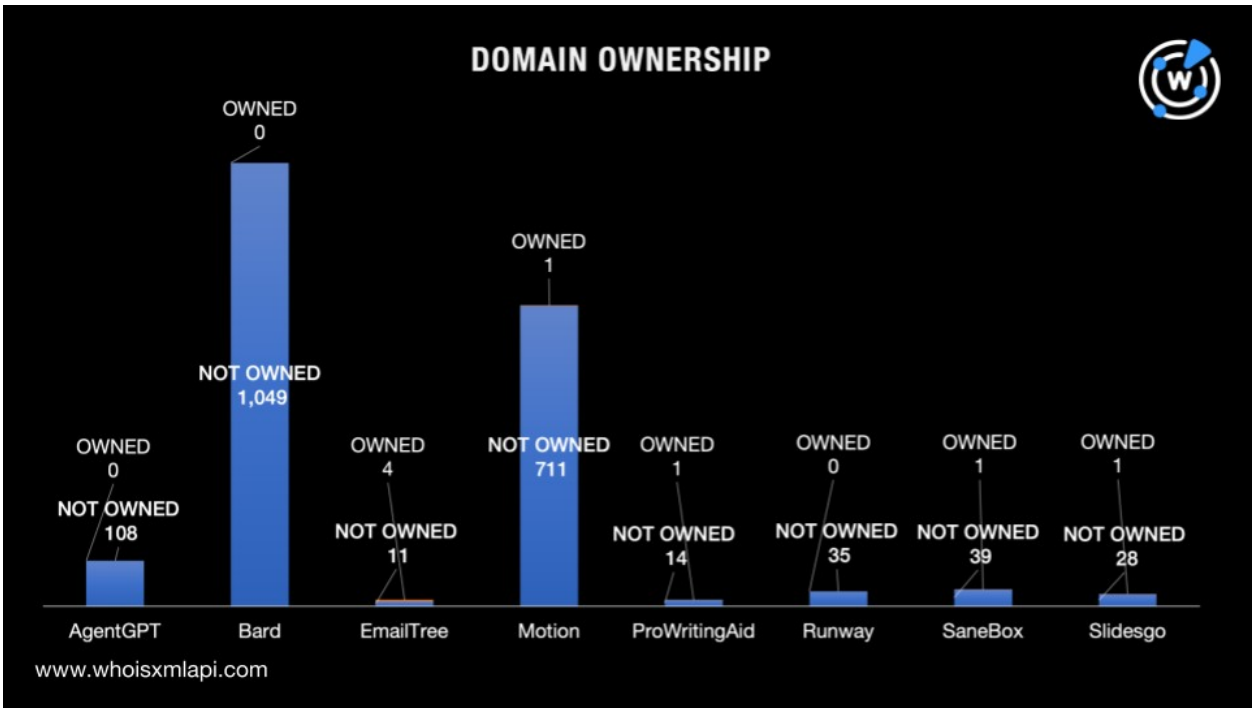
- **agentgpt**
- **bard + ai**
- **emailtree**
- **motion + ai**
- **prowritingaid**
- **runway + ml**
- **sanebox**
- **slidesgo**

その結果、合計2,003個のドメイン名がヒットしました。内訳は以下の通りです。

AIツール	ドメイン名数
Agent GPT	108
Bard	1,049
EmailTree	15
Motion	712
ProWritingAid	15
Runway	35

SaneBox	40
Slidesgo	29

WHOISレコードを比較したところ、これらのブランド名を含むドメイン名のうち、当社が選んだAI生産性向上ツール開発者のいずれかに帰属すると公開のWHOISデータで確認できたものは1%未満でした。

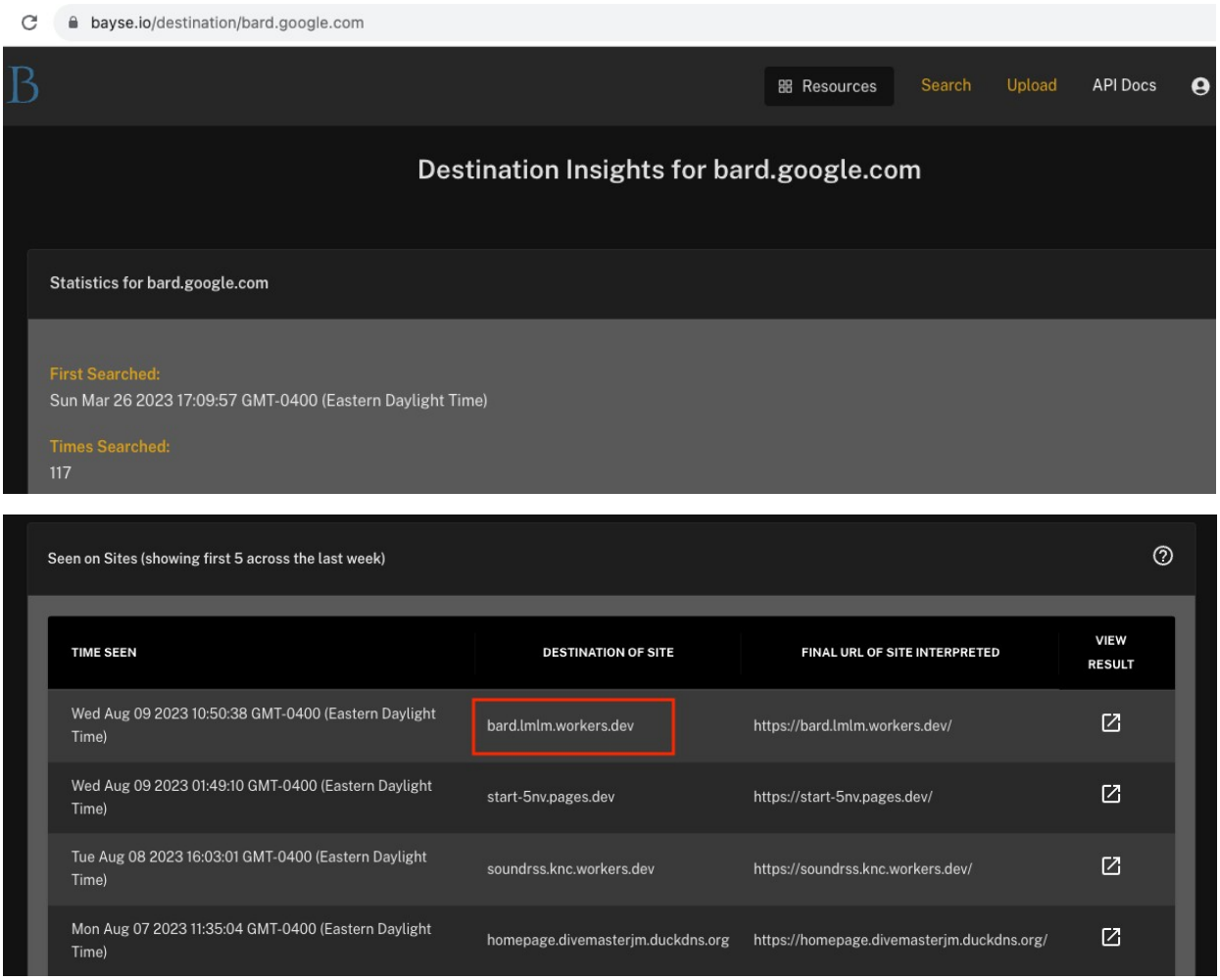


パート2：Bayseの分析

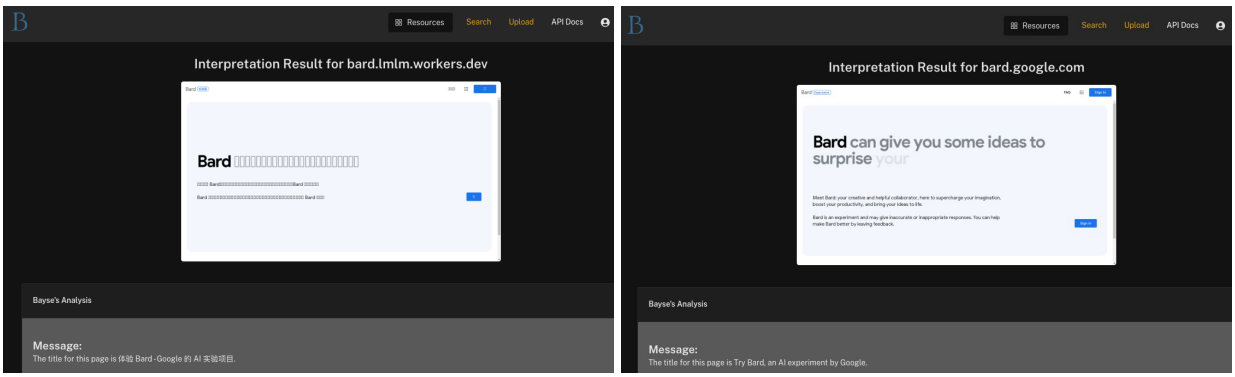
攻撃者が価値の高いウェブサイトになりすます主な方法の一つは、そのコンテンツを複製またはクローンすることです。この方法により、ユーザーが偽装サイトを正規のサイトと視覚的に関連付けるため、攻撃者が目的（認証情報や個人情報の収集、マルウェアのダウンロードなど）を達成する可能性が高まります。

このような手口はいくつかのAIツールで使われていますが、Bardを標的にしたものが圧倒的に多く見られました。

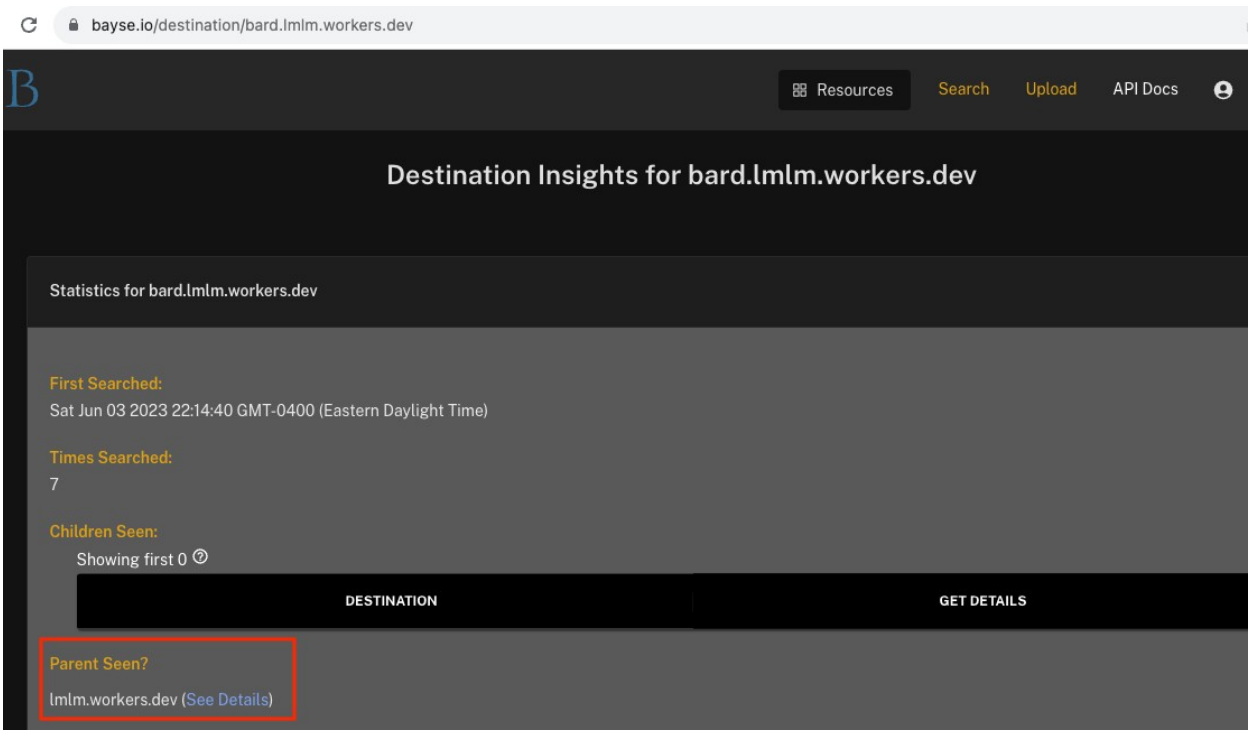
BardのサイトのURLをBayse Intelligenceに提出すると、以下のようにBardのアセットが参照される頻度、時期、場所を確認できます。



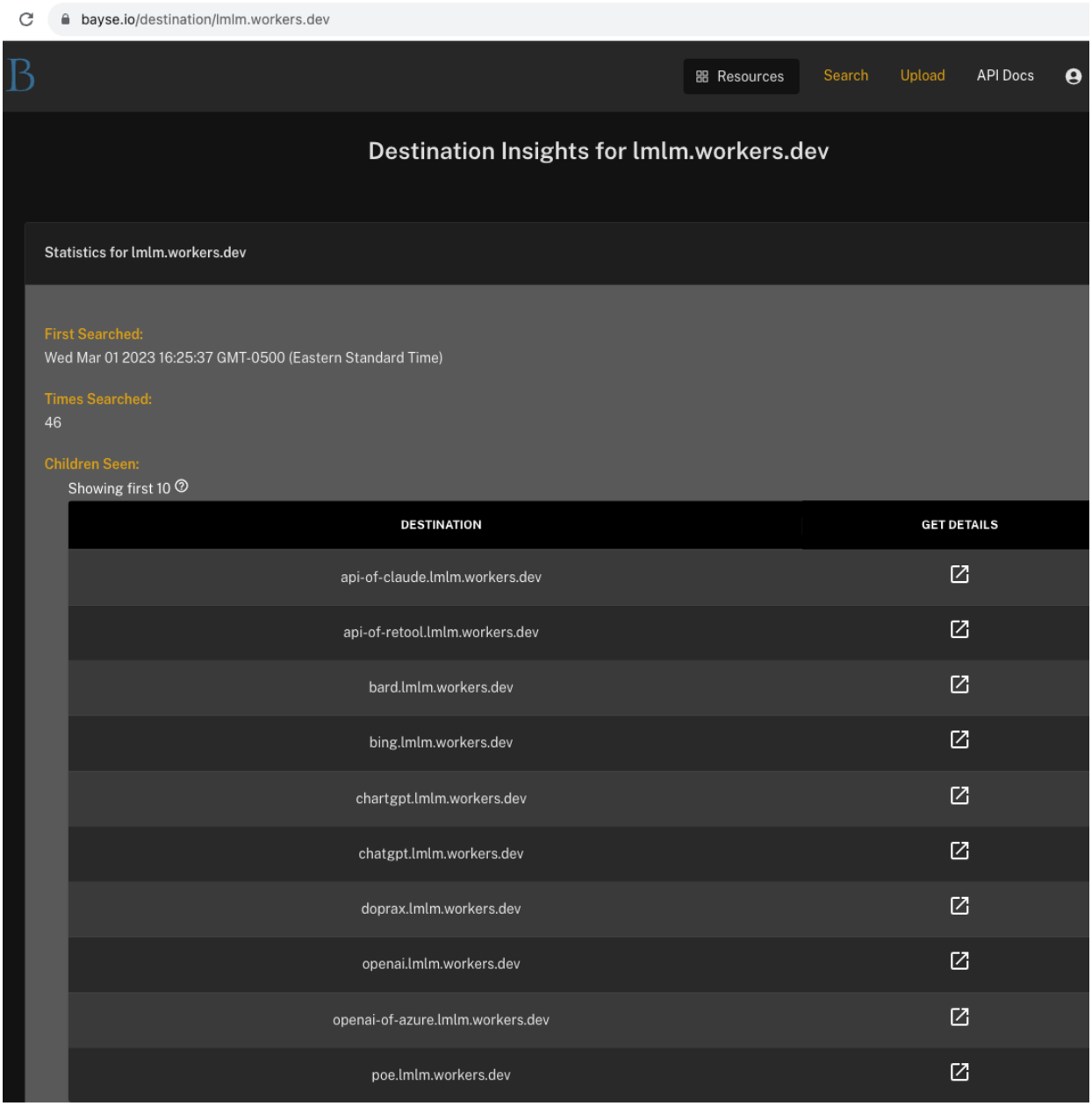
最近Bardにリンクを張ったサイトの一つ（上図のハイライト部分）は、明らかにBardになりすましています。



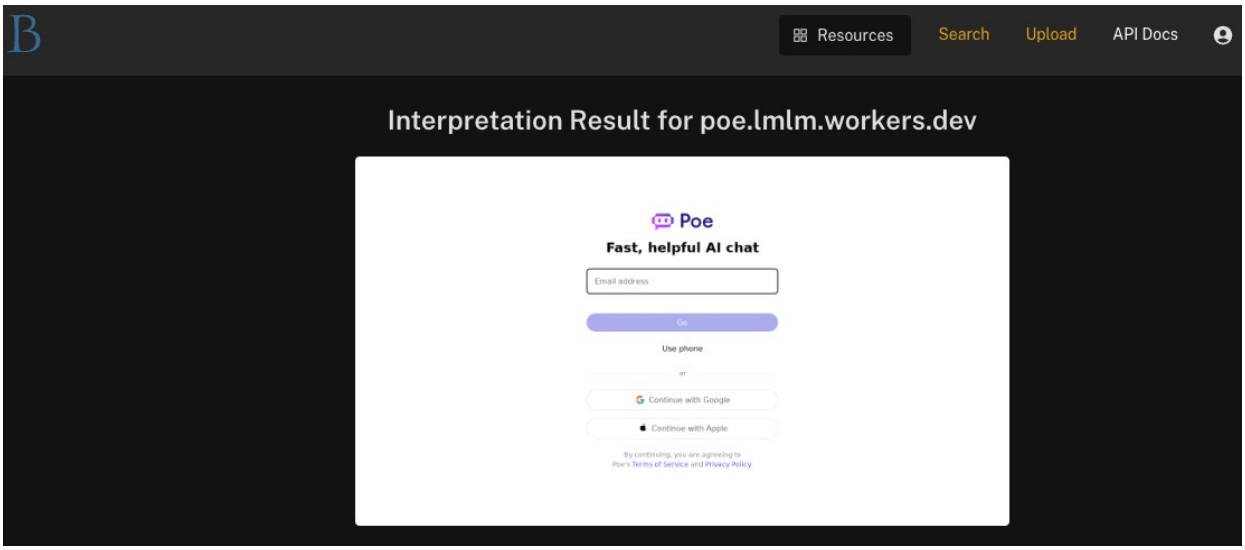
このサイトは過去2カ月間に複数回確認されており、親ドメイン（lmlm[.]works[.]dev）に関連づけられた他のサイトも確認されています。



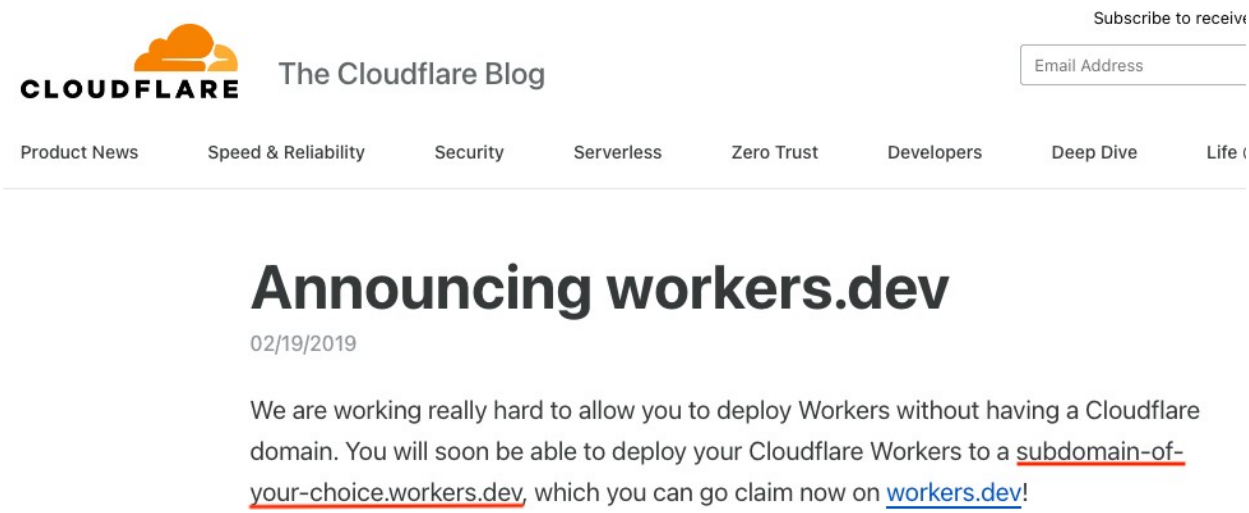
親ドメインの詳細に目を転じると、標的はBardだけではないことがわかります。2023年3月以降に標的とされた人気のAIやクラウド関連技術が他にも複数ありました。



これらのサイトのうち一部はダウンしていますが、その一方で、別の一部サイトについてはなりすましが現在進行中であることがわかります。



親ドメイン（lmlm[.]workers[.]dev）はCloudflareのウェブアプリホスティングプラットフォームでホストされており、これらのサイトは全て同じlmlmというサブドメインを共有していることから、これまでにハイライトしたサイトは全て、実際には同じ1人（または組織）の脅威アクターによって作成されたことになります。この証拠は、2019年のCloudflareの公式発表まで遡ることができます（以下）。



これが意味するのは、現在Cloudflareのインフラ上でコンテンツをホストしている脅威アクターが存在すること、そして、その脅威アクターが5カ月以上にわたって、非常に人気が高い多数のAIベースおよびクラウドベースのツールの利用者を標的にしていた可能性が高いということです。このサブドメイン（lmlm[.]workers[.]dev）およびその下のドメインで行われているいかなる活動も極めて疑わしいものとして扱われるべきであり、完全にブロックすることが推奨されます。

同様の調査をご希望のお客様、または本調査で使用したツールにご興味をお持ちのお客様は、ぜひ[whoisxmlapi.com](whoisxmlapi.com)または[bayse.io](bayse.io)をご参照ください。

免責事項： *当社は、潜在的な危険から企業を守るために役立つ情報の提供を目指しており、脅威の検出に対して厳格な姿勢で臨んでいます。そのため、当初「脅威」または「悪意がある」とみなされたエンティティが、さらなる調査やその後の状況の変化により最終的に無害と判断される可能性があります。ここで提供されている情報を確認する補足調査の実施を強くお勧めします。*

付録：アーティファクトの例

ブランド名を含むドメイン名の例

- agentgpt[.]digital
- agentgptstudio[.]com
- agentgpt[.]pt
- agentgpt[.]team
- agentgptexpert[.]com
- agentgpt[.]red
- agentgpt-website[.]com
- agentgptcn[.]online
- agentgpt[.]com[.]br
- agentgpt[.]info
- agentgpt[.]finance
- agentgpt-p7pnrkh44-rogerthiede[.]vercel[.]app
- agentgpt[.]asia
- bardai[.]ai
- bardrail[.]ai
- bard-maintain[.]fr
- bard[.]ai
- bardaiklaipeda[.]lt
- bardaitraining[.]com
- bardaisanism[.]faith
- bardsaimailing[.]com
- barde[.]ai
- bardo[.]ai
- bards[.]ai
- bardavilaitaim[.]com[.]br
- bardai[.]uk
- emailtree[.]co
- emailtree[.]net
- emailtreefrog[.]ca
- emailtree[.]in
- emailtreeai[.]com
- emailtree[.]club
- emailtree[.]pw
- emailtree[.]ai
- emailtrees[.]com
- emailtree[.]icu
- emailtree[.]uk
- emailtree[.]com
- emailtree[.]co[.]uk
- motionaid[.]training
- motion[.]ai
- motions[.]ai
- motiong[.]ai
- motionit[.]ai
- motionos[.]ai
- motioniq[.]ai
- motionai[.]eu
- motionstakeairports[.]email
- motional[.]ai
- motionai[.]cn
- motionai[.]io
- motionce[.]ai
- prowritingaid[.]tk

- prowritingaid[.]com
- prowritingaid[.]app
- prowritingaid[.]cn
- prowritingaide[.]com
- prowritingaid[.]ca
- prowritingaid[.]net
- prowritingaids[.]com
- prowritingaid[.]org
- prowritingaid[.]co
- prowritingaid[.]info
- prowritingaid[.]nl
- prowritingaid[.]co[.]uk
- runwayml[.]ml
- runwayml[.]ai
- runwayml[.]cn
- runwayml[.]fr
- runwayml[.]eu
- runwayml[.]it
- runwayml[.]de
- runwayml[.]co
- runwayml[.]net
- runwayml[.]xyz
- runwayml[.]com
- runwayml[.]vip
- runwayml[.]top
- sanebox[.]me
- sanebox[.]cloud
- sanebox[.]co
- saneboxoffst[.]gq
- sanebox-support[.]com
- sanebox[.]fr
- sanebox[.]net
- sanebox[.]rocks
- sanebox[.]com[.]au
- saneboxpartners[.]com
- sanebox[.]in
- sanebox[.]se
- sanebox[.]nl
- slidesgo[.]xyz
- slidesgoogle[.]com
- slidesgo[.]net
- slidesgoogle[.]gq
- slidesgoai[.]com
- slidesgo[.]com[.]de
- slidesgo[.]net[.]cn
- slidesgo[.]com
- slidesgo[.]cm
- slidesgoo[.]com
- slidesgo[.]ru
- slidesgod[.]com
- slidesgo[.]cn