Conduct Search Engine Discovery Reconnaissance for Information Leakage (Bilgi Sızıntısı için Arama Motoru Keşif Keşif Keşif Keşifini Yapın)

Summary (Özet)

Arama motorlarının çalışabilmesi için bilgisayar programları (veya `robots`) düzenli olarak veri alın (web'deki milyarlarca sayfadan sürünme olarak anılır. Bu programlar, diğer sayfalardan gelen bağlantıları takip ederek veya site haritalarına bakarak web içeriği ve işlevselliğini bulur. Bir web sitesi adı özel bir dosya kullanırsa `robots.txt` Arama motorlarının gelmesini istemediği sayfaları listelemek için, orada listelenen sayfalar göz ardı edilecektir. Bu temel bir genel bakıştır - Google, bir arama motorunun nasıl çalıştığına dair daha derinlemesine bir açıklama sunar.

Testçiler, web sitelerinde ve web uygulamalarında keşif yapmak için arama motorlarını kullanabilirler. Arama motoru keşfi ve keşif için doğrudan ve dolaylı unsurlar vardır: doğrudan yöntemler, indeksleri ve ilgili içeriği önbelleklerden aramakla ilgiliyken, dolaylı yöntemler forumları, haber grupları ve pazarlama web sitelerini arayarak hassas tasarım ve yapılandırma bilgilerini öğrenmekle ilgilidir.

Bir arama motoru robotu taramayı tamamladıktan sonra, web içeriğini etiketlere ve ilgili niteliklere göre dizine dizine başlamaya başlar. `<TITLE>`, ilgili arama sonuçlarını iade etmek için. Eğer eğer varsa `robots.txt` Web sitesinin ömrü boyunca dosya güncellenmez ve robotlara içeriği
endekslememelerini söyleyen satır içi HTML meta etiketleri kullanılmamıştır, daha sonra indekslemelerin sahipleri tarafından dahil edilmesi amaçlanmayan web içeriği içermesi mümkündür. Web sitesi sahipleri daha önce bahsedilenleri kullanabilir

`robots.txt` HTML meta etiketleri, kimlik doğrulama ve bu tür içeriği kaldırmak için arama motorları tarafından sağlanan araçlar.

Test Objectives (Test Hedefleri)

- Uygulama, sistem veya kuruluşun hangi hassas tasarım veyapılandırma bilgilerinin doğrudan (kuruluşun web sitesinde) veyadolaylı olarak (üçüncü taraf hizmetleri aracılığıyla) maruz kaldığını belirleyin.

How to Test (Nasıl Test Edilir)

Potansiyel olarak hassas bilgileri aramak için bir arama motoru kullanın. Bu şunları içerebilir:

- ağ diyagramları ve konfigürasyonları;
- Yöneticiler veya diğer kilit personel tarafından arşivlenmiş gönderiler ve e-postalar;
- logon prosedürleri ve kullanıcı adı formatları;
- kullanıcı adları, şifreler ve özel tuşlar;
- Üçüncü taraf veya bulut servis yapılandırma dosyaları;
- hata mesajı içeriğinin ortaya çıkarılması; ve
- Geliştirme, test, Kullanıcı Kabul Testi (UAT) ve sitelerin evreleme sürümleri.

Search Engines (Arama Motorları)

Testleri sadece bir arama motoru sağlayıcısıyla sınırlamayın, çünkü farklı arama motorları farklı sonuçlar üretebilir. Arama motoru sonuçları, motorun son ne zaman tarandığına ve motorun ilgili sayfaları belirlemek için kullandığı algoritmaya bağlı olarak birkaç şekilde değişebilir. Aşağıdaki (alfabetik olarak listelenen) arama motorlarını kullanmayı düşünün:

- Baidu, Çin'in en popüler arama motoru.
- BingMicrosoft'un sahip olduğu ve işlettiği bir arama motoru olan Bing, dünya çapında en popüler ikinci kişi. Gelişmiş arama anahtar kelimelerini destekler.
- binsearch.info, ikili Usenet haber grupları için bir arama motoru.

- Common Crawl, “herhangi biri tarafından erişilebilen ve analiz edilebilen web tarama verilerinin açık bir deposu.”
- DuckDuckGo, birçok farklı kaynaktan gelen sonuçları derleyen gizlilik odaklı bir arama motoru. Arama sözdizimi destekler.
- Dünyanın en popüler arama motorunu sunan ve en alakalı sonuçları iade etmeye çalışmak için bir sıralama sistemi kullanan Google. Arama operatörlerini destekler.
- İnternet Arşivi Wayback Machine, “İnternet sitelerinden ve diğer kültürel eserlerin dijital bir kütüphanesini dijital formda inşa etmek”.
- Startpage, izleyiciler ve günlükler aracılığıyla kişisel bilgi toplamadan Google'ın sonuçlarını kullanan bir arama motoru. Arama operatörlerini destekler.
- Shodan, internete bağlı cihazları ve hizmetleri aramak için bir hizmet. Kullanım seçenekleri
  sınırlı bir ücretsiz planın yanı sıra ücretli abonelik planları içerir.

Hem DuckDuckGo hem de Startpage, izleyicileri kullanmayarak veya günlük tutmayarak kullanıcılara daha fazla gizlilik sunar. Bu, test cihazı hakkında azaltılmış bilgi sızıntısı sağlayabilir.

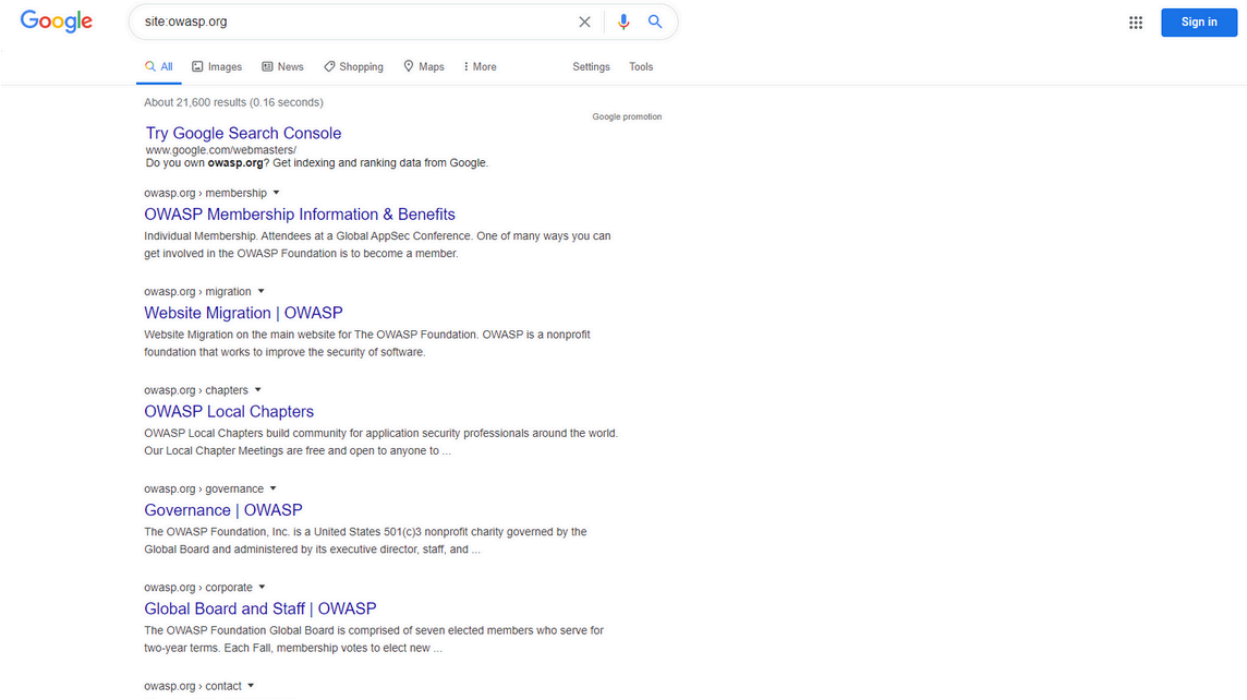
Search Operators (Arama Operatörleri)

Bir arama operatörü, düzenli arama sorgularının yeteneklerini genişleten ve daha spesifik sonuçlar elde etmeye yardımcı olabilecek özel bir anahtar kelime veya sözdizgesidir. Genellikle formlarını
alırlar
`operator:query`. . İşte yaygın olarak desteklenen bazı arama operatörleri:

- `site:` Aramayı sağlanan alan adı ile sınırlandıracaktır.
- `inurl:` Yalnızca URL’ye anahtar kelimeyi içeren sonuçları iade edecektir.
- `intitle:` Sadece sayfa başlığında anahtar kelimeye sahip olan sonuçları iade edecektir.
- `intext:` ya da `inbody:` Anahtar kelimeyi yalnızca sayfaların gövdesinde arayacaktır.
- `filetype:` Sadece belirli bir dosya tipi, yani png veya php ile eşleşir.

Örneğin, owasp.org'un web içeriğini tipik bir arama motoru tarafından dizine eklenen şekilde bulmak için, gerekli olan sözdizimi şunlardır:

```
site:owasp.org
```



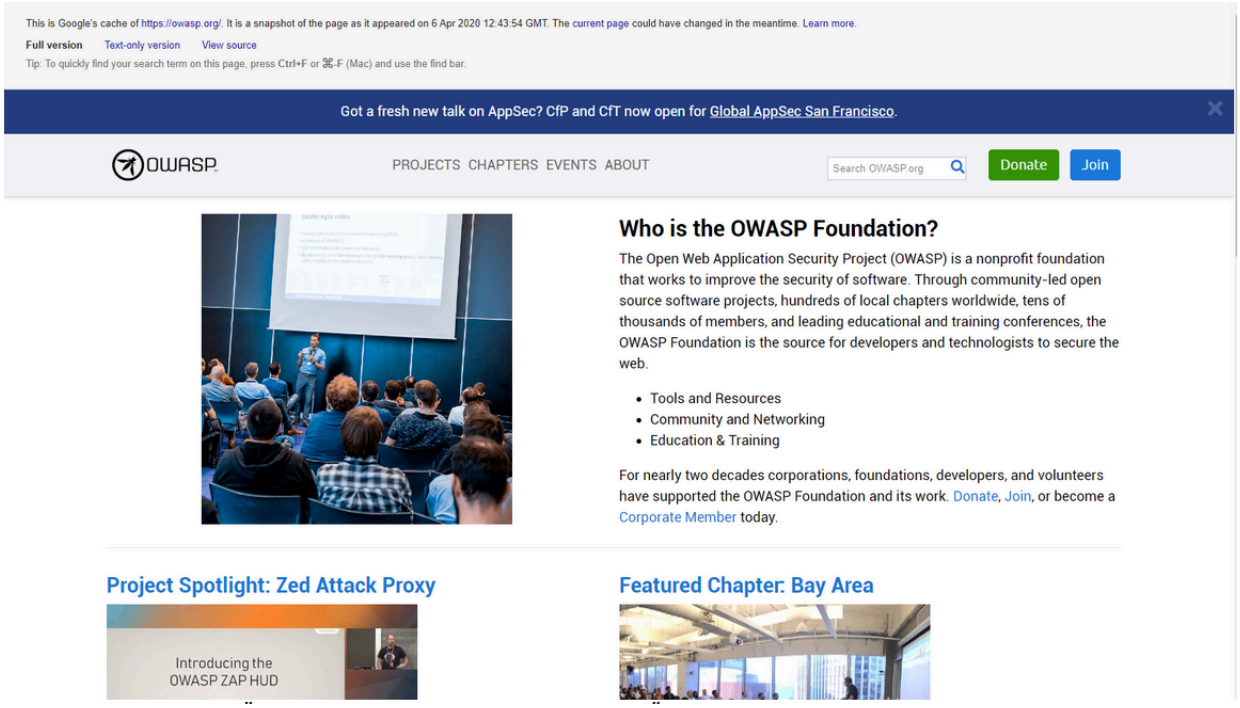
Şekil 4.1.1-1: Google Site Operasyonu Arama Sonucu Örneği

Viewing Cached Content (Önbelleğe Takdire Alınan Içeriği)

Daha önce indekslenmiş olan içeriği aramak için kullanın `cache:` Operatör. Bu, endekslendiği zamandan beri değişmiş olabilecek veya artık mevcut olmayan içeriği görüntülemek için yararlıdır. Tüm arama motorları arama için önbelleğe alınmış içerik sağlamaz; yazma sırasında en kullanışlı
kaynak Google'dır.

Görüntülemek için `owasp.org` Önbelleğe alındığı gibi, sözdizimi:

```
cache:owasp.org
```



Şekil 4.1.1-2: Google Önbellek Operasyonu Arama Sonucu Örneği

Google Hacking, or Dorking (Google Hacking veya Dorking)

Operatörlerle arama yapmak, test cihazının yaratıcılığı ile birleştirildiğinde çok etkili bir keşif tekniği olabilir. Operatörler, belirli hassas dosyaları ve bilgileri etkili bir şekilde keşfetmek için zincirlenebilir. Google hacking veya Dorking adı verilen bu teknik, arama operatörleri desteklendiği sürece diğer arama motorları kullanmak da mümkündür.

Google Hacking Veritabanı gibi bir çukurlar veritabanı, belirli bilgilerin ortaya çıkmasına
yardımcı olabilecek yararlı bir kaynaktır. Bu veritabanında bulunan bazı çukurlar kategorileri şunlardır:

- Kategori: Ayaklar
- Kullanıcıları içeren dosyalar
- Hassas Dizinler
- Web Sunucusu Algılama
- Savunmasız Dosyalar

- Savunmasız Sunucular
- Hata Mesajları
- Sulu bilgiler içeren dosyalar
- Şifre içeren dosyalar
- Hassas Online Alışveriş Bilgileri

Bing ve Shodan gibi diğer arama motorlarının veri tabanları, Bishop Fox'un Google Hacking Diggity Project gibi kaynaklardan temin edilebilir.

Remediation (Düzeltme)

Çevrimiçi yayınlanmadan önce tasarım ve yapılandırma bilgilerinin hassasiyetini dikkatlice düşünün.

Çevrimiçi olarak yayınlanan mevcut tasarım ve yapılandırma bilgilerinin hassasiyetini periyodik olarak gözden geçirin.