BLAZE

INFORMATION SECURITY

What you see is NOT what you get.

When homographs attack.

BLAZE

INTRO

Julio Cesar Fort

Director of Professional Services
at Blaze Information Security

INTRO

Since the introduction of Unicode in domain names, a series of brand new security implications were also brought into light together with the possibility of registering domain names using different alphabets and Unicode characters.

Agenda

BLAZE

Internationalized Domain Names and how they work

Homographs: security risks and considerations

User agents and homograph attacks

Practical attacks

How to defend yourself

Conclusion

BLAZE

Internationalized Domain Names and how they work

BLAZE

INTERNATIONALIZED DOMAIN NAMES

EMERGENCE OF INTERNATIONALIZED DOMAIN NAMES:

- The Internet **was not designed to be multilingual**
- Domain names were **confined to Latin-based characters**
- However, **billions of people do not have Latin-based languages as their first language**

INTERNATIONALIZED DOMAIN NAMES

EVOLUTION OF IDN

- ICANN resolution version 1 – subsequent versions later
- Wide support for Unicode characters

INTERNATIONALIZED DOMAIN NAMES

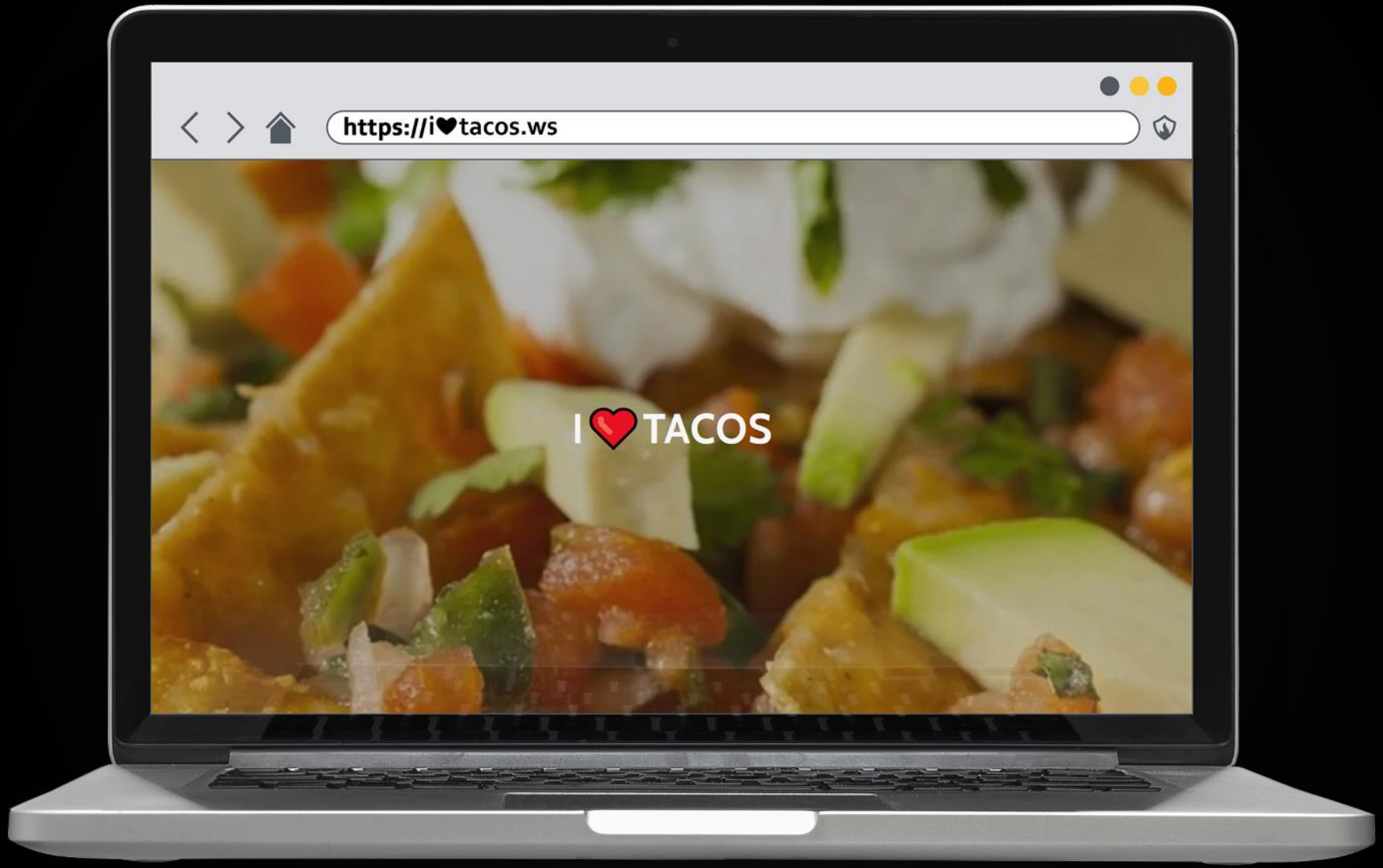
“HOUSTON, WE HAVE A PROBLEM“

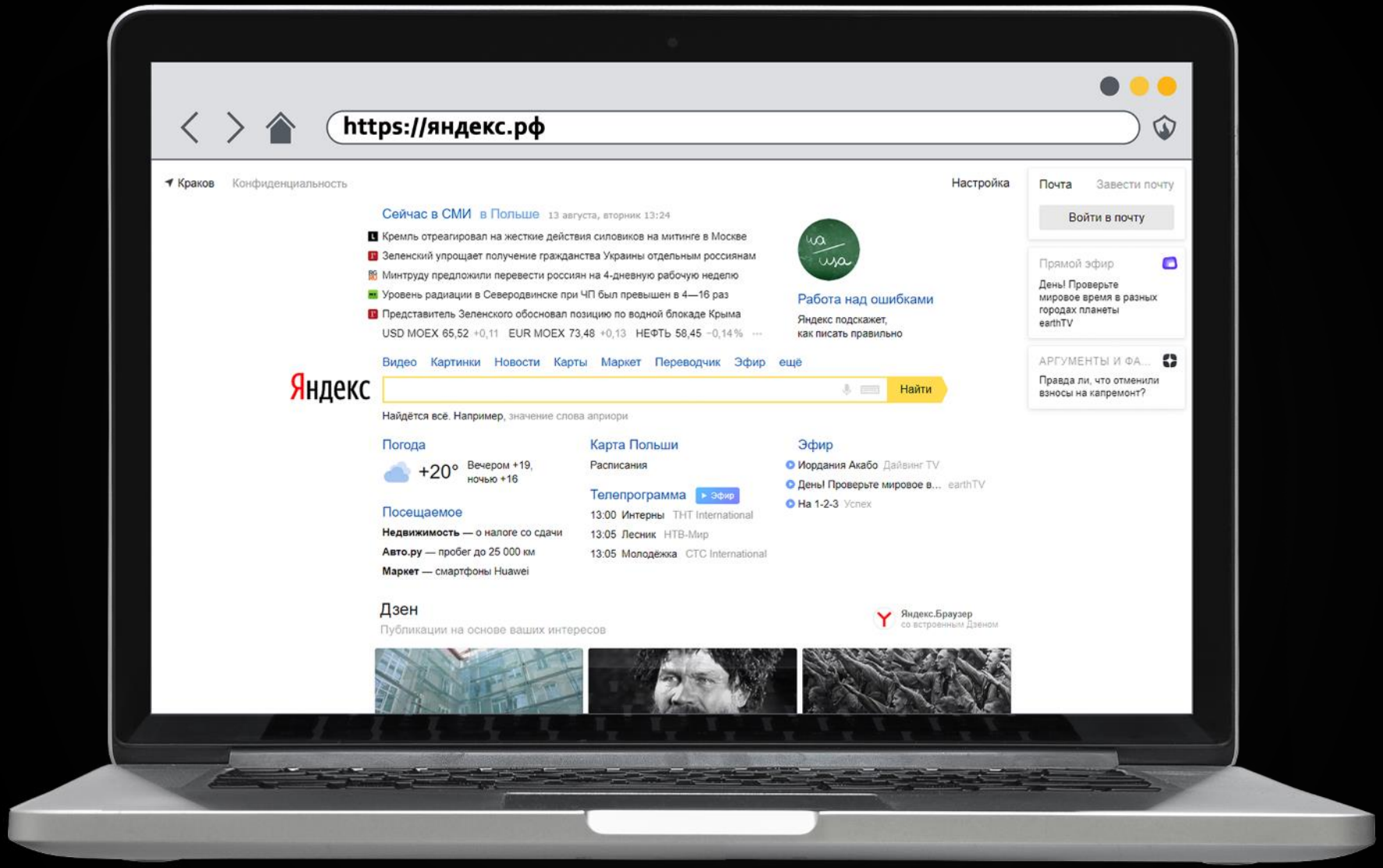
DNS is ASCII only (A-Z, 0-9 and „-”) and does not support Unicode

INTERNATIONALIZED DOMAIN NAMES

PUNYCODE

- **Translates Unicode into ASCII** using an algorithm known as IDNA2008
- Converts 🐮**.ws** into **xn--2o8h.ws**
- Or **öbb.at** into **xn--bb-eka.at**




https://яндекс.рф
Краков
Конфиденциальность
Настройка
Почта
Завести почту
Войти в почту
Сейчас в СМИ в Польше 13 августа, вторник 13:24
Кремль отреагировал на жесткие действия силовиков на митинге в Москве
Зеленский упрощает получение гражданства Украины отдельным россиянам
Минтруду предложили перевести россиян на 4-дневную рабочую неделю
Уровень радиации в Северодвинске при ЧП был превышен в 4—16 раз
Представитель Зеленского обосновал позицию по водной блокаде Крыма
USD MOEX 65,52 +0,11 EUR MOEX 73,48 +0,13 НЕФТЬ 58,45 −0,14%
Работа над ошибками
Яндекс подскажет, как писать правильно
Прямой эфир
День! Проверьте мировое время в разных городах планеты
earthTV
АРГУМЕНТЫ И ФА...
Правда ли, что отменили взносы на капремонт?
Видео Картинки Новости Карты Маркет Переводчик Эфир ещё
Яндекс
Найти
Найдётся всё. Например, значение слова априори
Погода
+20° Вечером +19, ночью +16
Карта Польши
Расписания
Эфир
Иордания Акабо Дайвинг TV
День! Проверьте мировое в... earthTV
На 1-2-3 Успех
Телепрограмма
Эфир
13:00 Интерны ТНТ International
13:05 Лесник НТВ-Мир
13:05 Молодёжка СТС International
Посещаемое
Недвижимость — о налоге со сдачи
Авто.ру — пробег до 25 000 км
Маркет — смартфоны Huawei
Дзен
Публикации на основе ваших интересов
Яндекс.Браузер
со встроенным Дзеном

INTERNATIONALIZED DOMAIN NAMES

PARTIAL IDNs

- **http://öbb.at**

FULL IDNs

- **http://президент.рф** (points to kremlin.ru)

Homographs: security risks and considerations

BLAZE

HOMOGLYPHS AND HOMOGRAPHS

Latin script, for example, can represent multiple languages (e.g., English, German, Spanish, French, Portuguese and more)

HOMOGLYPHS AND HOMOGRAPHS

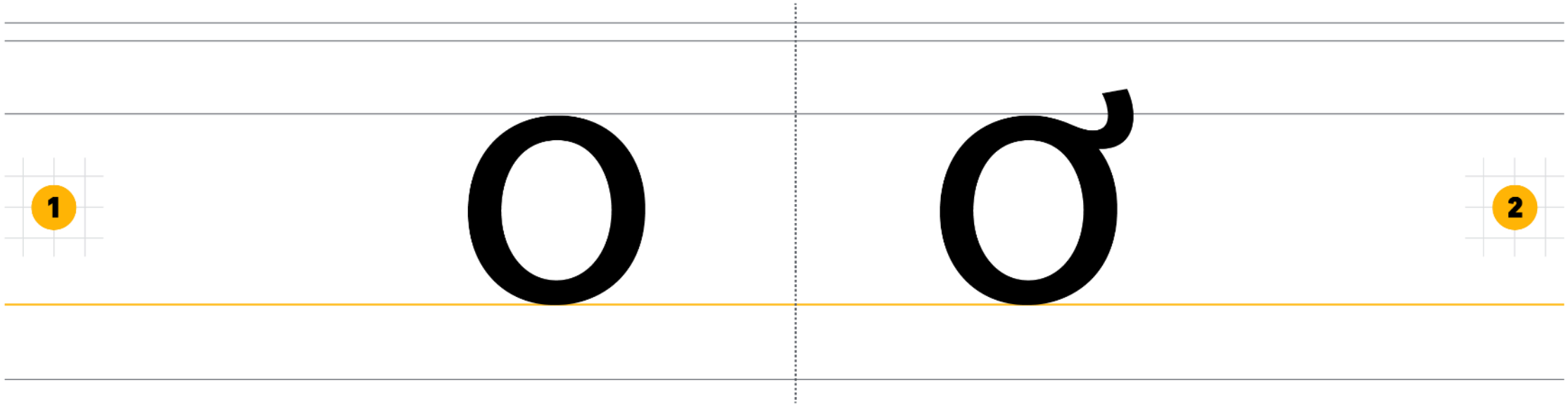
Different scripts share numerous characters that either look exactly similar or have a strong resemblance

CONFUSABLE HOMOGRAPHS

1 **a** (U+0061) – latin

2 **a** (U+0430) – cyrillic

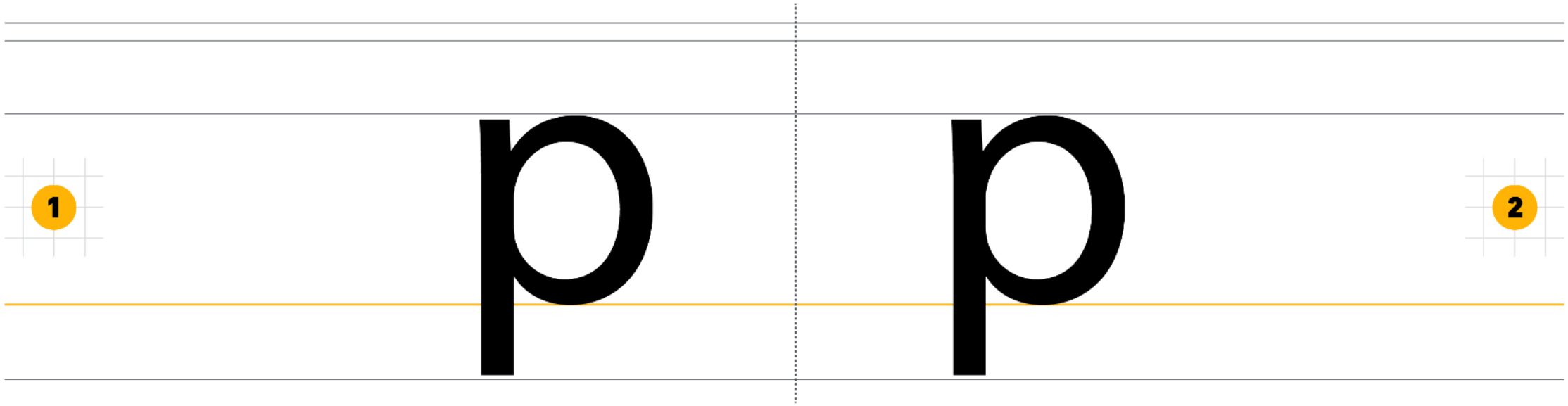
CONFUSABLE HOMOGRAPHS



1 o (U+006F) – latin

2 ơ (U+01A1) – “o” with a “horn” latin script

CONFUSABLE HOMOGRAPHS



1 **p** (U+006F) – latin “p”

2 **p** (U+0440) – “er” cyrillic

CONFUSABLE HOMOGRAPHS

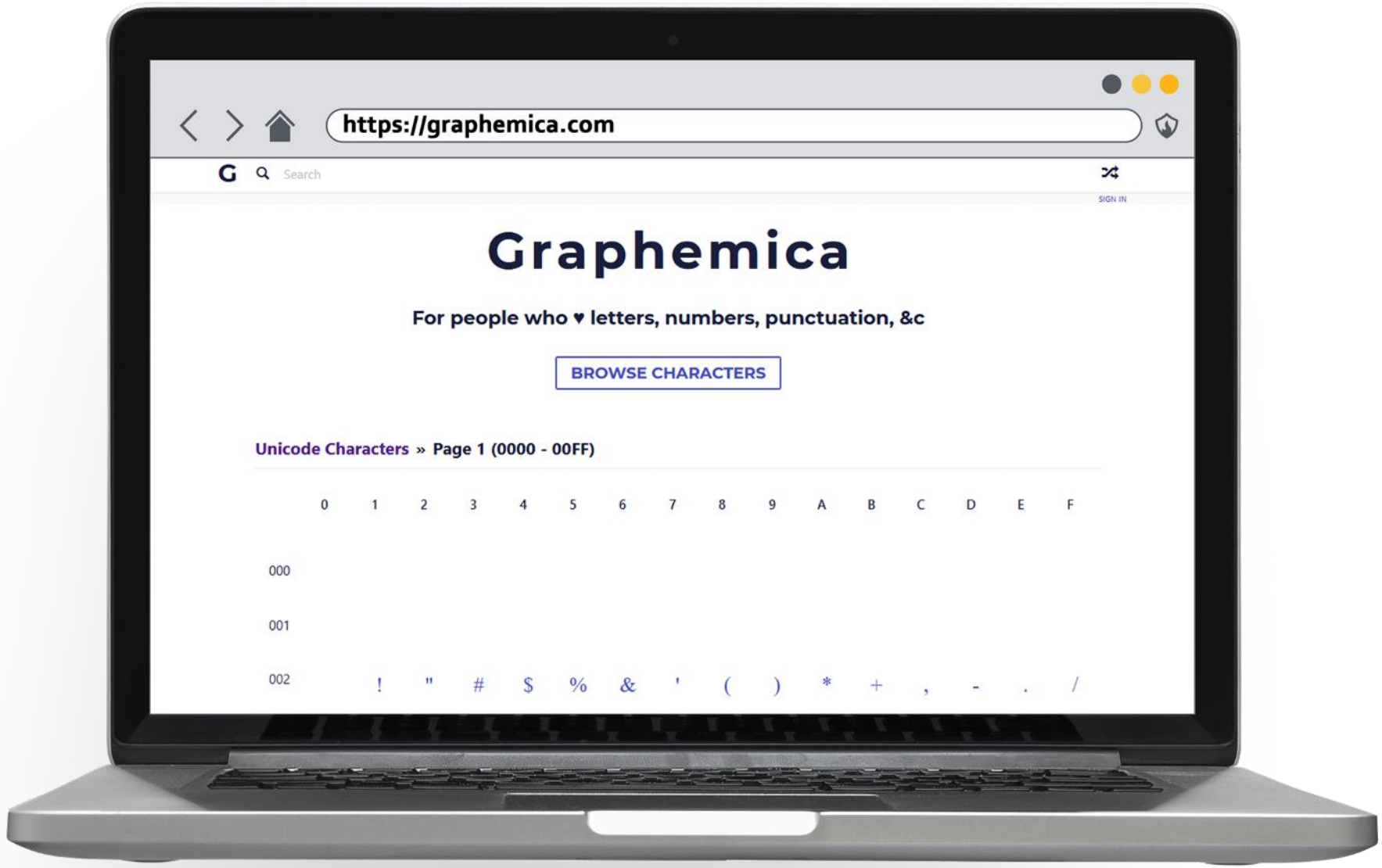
1 c (U+0063) – latin small ‘c’

2 c (U+2CA5) – Coptic small letter ‘sima’

CONFUSABLE HOMOGRAPHS

1 **c** (U+0063) – latin small ‘c’

2 **c** (U+0441) – letter name in Cyrillic is ‘es’



See **www.graphemica.com** for more info

User agents and homograph attacks

BLAZE

FONT RENDERIZATION AND VISUAL SPOOFING

IMPORTANT FACTORS

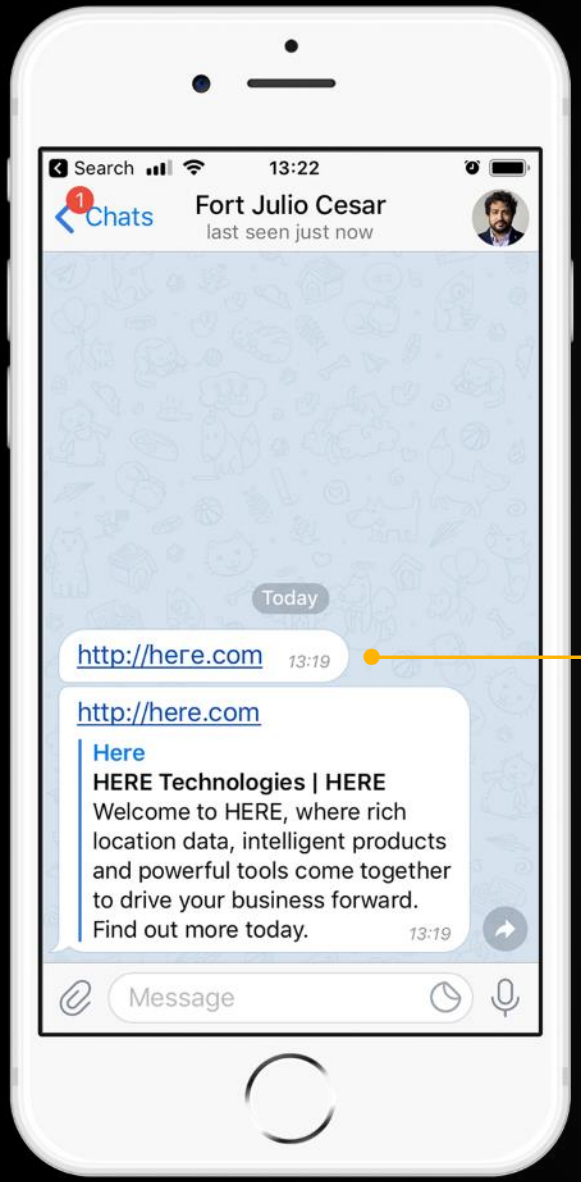
- Font type
- Font size
- The way it is rendered
- Even the display size


FONT: Tahoma Regular, 68pt (approx. render size @1920x1080, 72PPI)
LATIN
https://www.apple.com
CYRYLLIC
https://www.apple.com


FONT: Bookman Old Style Regular, 70pt (approx. render size @1920x1080, 72PPI)
LATIN
https://www.apple.com
CYRYLLIC
https://www.apple.com


FONT: Microsoft Yi Baiti, 96pt (approx. render size @1920x1080, 72PPI)
LATIN
https://www.apple.com
CYRYLLIC
https://www.apple.com


WIRE desktop for Windows (3.9.2928)
TELEGRAM desktop for Windows (1.8.1)
BLAZE INFORMATION SECURITY
JULIO CESAR FORT
Julio Cesar Fort
@400%
LEGITIMATE
http://here.com
HOMOGRAPH
http://here.com
VS
Julio
http://here.com/ and htt http://here.com
Julio Cesar Fort
last seen 3 minutes ago
@400%
LEGITIMATE
http://here.com
HOMOGRAPH
http://here.com
August 14
http://here.com/ 13:19
http://here.com
Here
HERE Technologies | HERE
Welcome to HERE, where rich location data, intelligent products and powerful tools come together to drive your business forward. Find out more today.
13:19
Write a message...



TELEGRAM iOS (12.3.1) client (version 5.10)

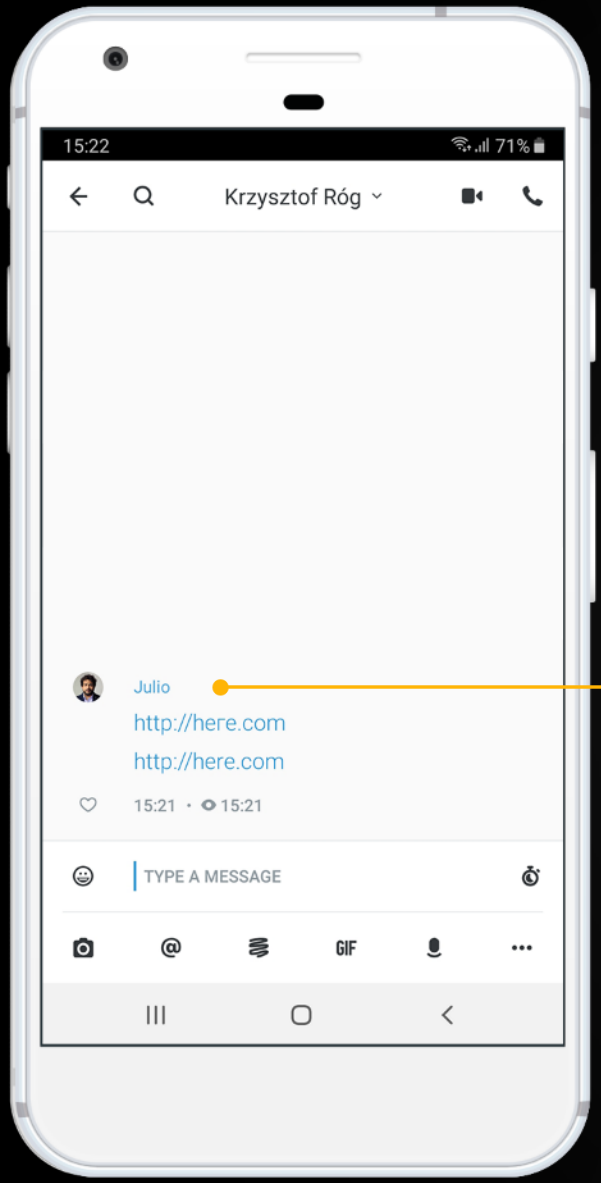
@400%

LEGITIMATE

http://here.com

HOMOGRAPH

http://here.com



WIRE

Android 9 (Pie; Samsung One UI ROM 1.1) client (3.35.814), **CVE-2019-15103**

@400%

LEGITIMATE

http://here.com

HOMOGRAPH

http://here.com

REGISTRATION OF HOMOGRAPH DOMAINS

RULES SEEM TO VARY DEPENDING ON THE gTLD:

- **.ws, .to:** all possible IDN languages allowed
- **.com, .net, .tv:** symbols from Portuguese, Romanian, Javanese, Thai, Sanskrit, Russian, etc.
- **.berlin:** Latin and Cyrillic scripts

REGISTRATION OF HOMOGRAPH DOMAINS

ICANN's IDN versions:

- ICANN's IDN **version 1 allowed** mixed scripts
- IDN **version 2 and 3 disallowed** mixed scripts

REGISTRATION OF HOMOGRAPH DOMAINS

PURE SCRIPTS CAN BE REGISTERED AND ARE TOTALLY FINE:

- **paypal.com**
- **apple.com**
- **opera.com**
- **yaһoo.com**
- **pһp.net**
- **һere.com**
- **faċebook.com**
- ...and many others

Practical attacks

BLAZE

PRACTICAL ATTACKS

HOMOGRAPH ATTACKS

- Original paper by *Evgeniy Gabrilovich* and *Alex Gontmakher* in 2001
- **Lately, phishers have taken notice** and we've seen a rise in such attacks

HISTORICAL AND RECENT BUGS

VARIOUS SOURCES

- **Firefox**: bugzilla ID 279099 filled in 2005 by *3ric of Shmoo* (P3 importance)
- CVE-2018-4277 in **Safari** (**ⅆ** interpreted as **d**)
- CVE-2019-11721 in **Firefox** (**ĸ** interpreted as **k**)

XUDONG ZHENG'S 2017 RESEARCH

- **Chrome** Issue 683314 (P1 importance)
- **Firefox**: bugzilla ID 1332714 (P3 importance)
- **Tor Browser** ticket 21961

BROWSERS HANDLING OF IDNs

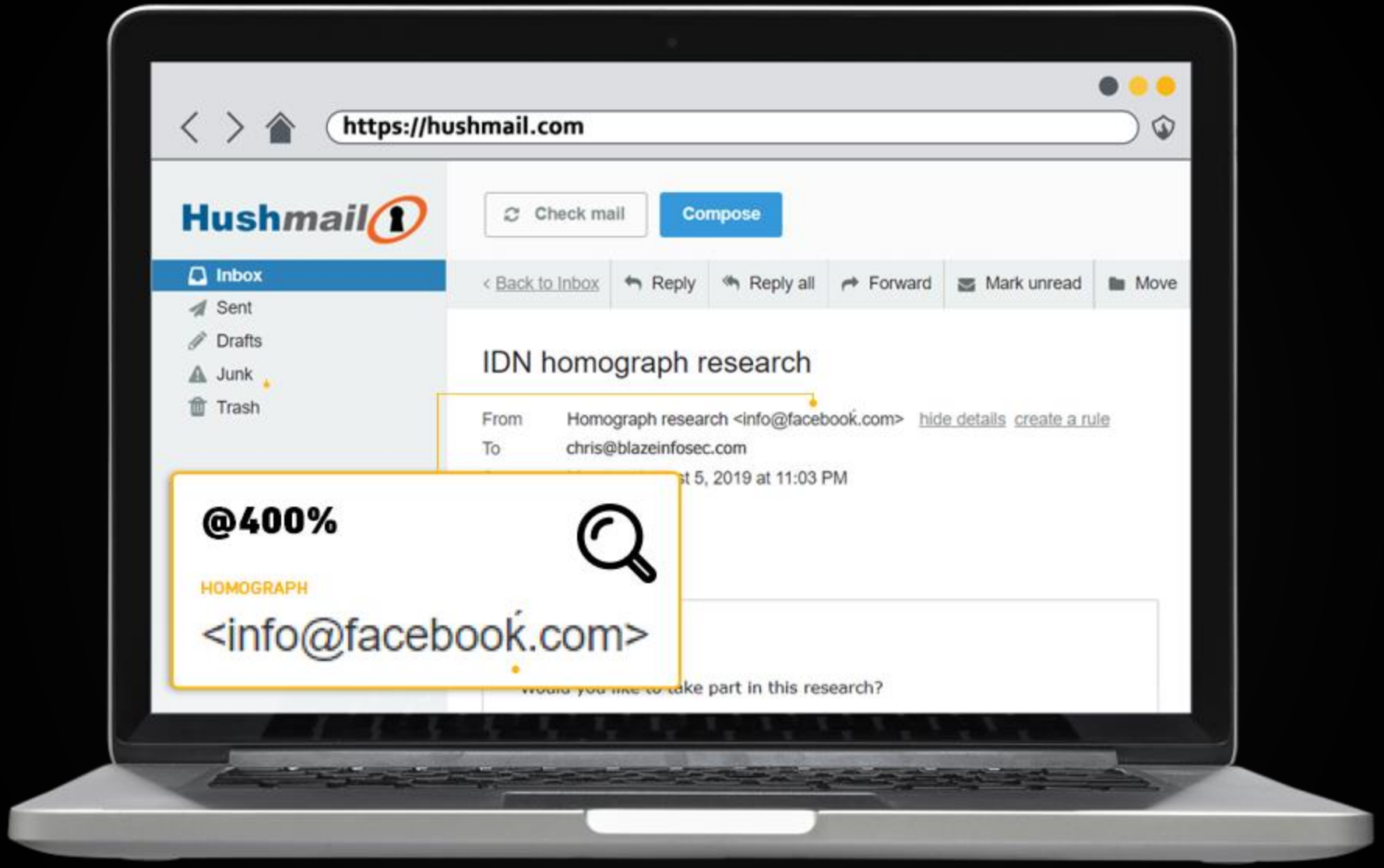
WHAT WE KNOW

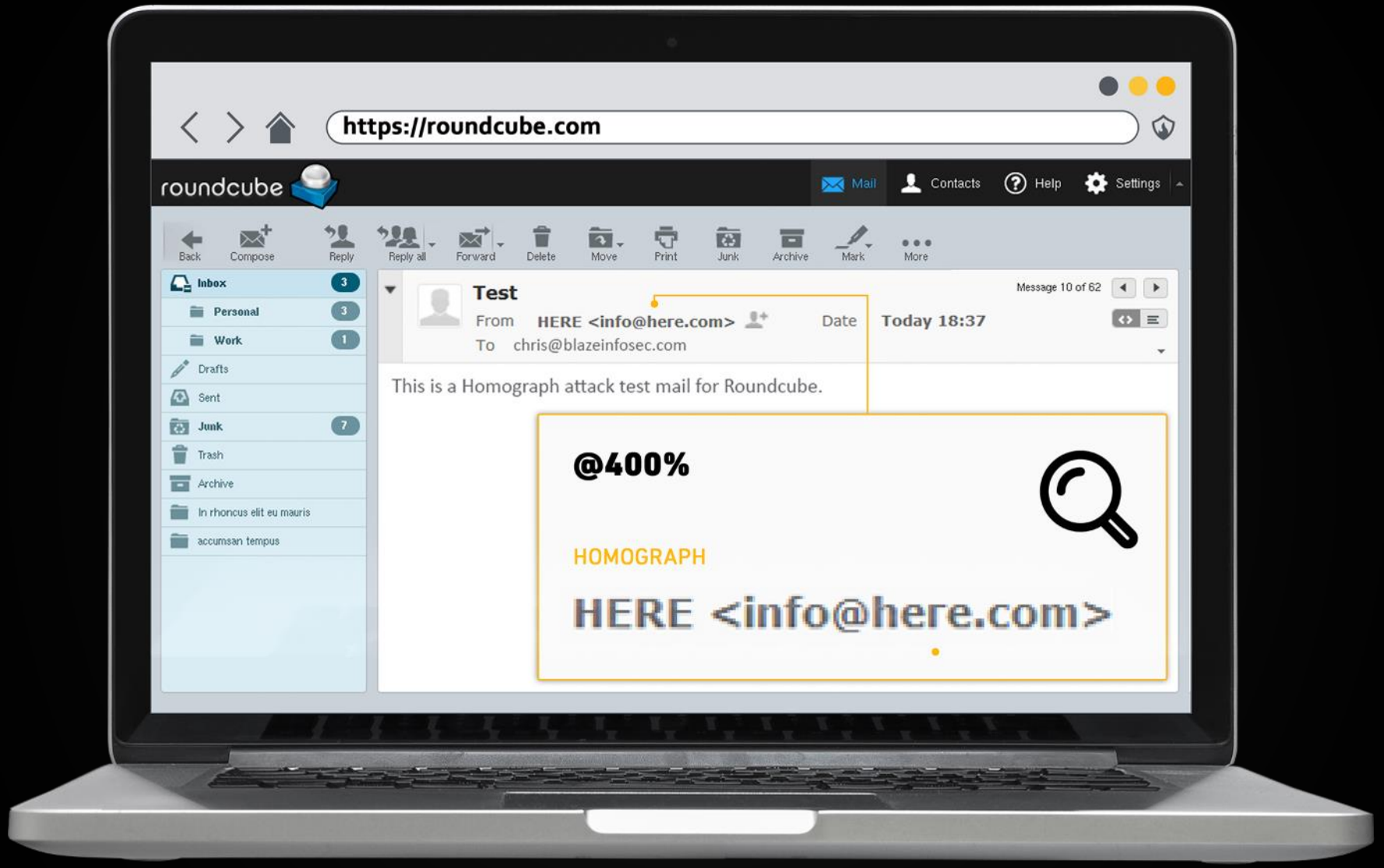
- **Chrome:** Has a quite complex policy to display IDNs
- **IE/Edge:** surprisingly, never seemed to suffer any issues
- **Firefox/Tor Browser:** will display Unicode characters in their intended scripts, even if they are confusable
- **Opera and Brave:** seems similar to Chrome

EMAIL CLIENTS AND WEBMAILS

“BACK-STABBING FRIEND”

- **For the sake of user-friendliness**, some clients and webmails translate convert from **punycode.toUnicode** to **punycode.toASCII**
- Often, **no checks for confusables are made**


https://hushmail.com
Hushmail
Check mail
Compose
Inbox
Sent
Drafts
Junk
Trash
Back to Inbox
Reply
Reply all
Forward
Mark unread
Move
IDN homograph research
From Homograph research <info@facebook.com> hide details create a rule
To chris@blazeinfosec.com
@400%
HOMOGRAPH
<info@facebook.com>
take part in this research?

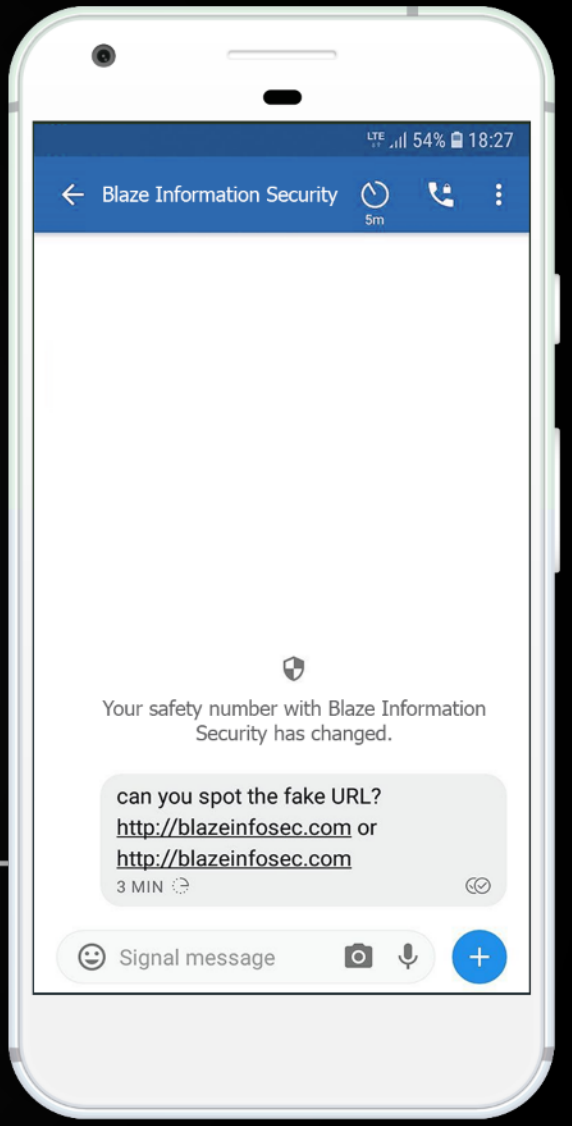


CVE-2019-15237

SECURE MESSENGER APPS:
SIGNAL

(for Android and Windows)

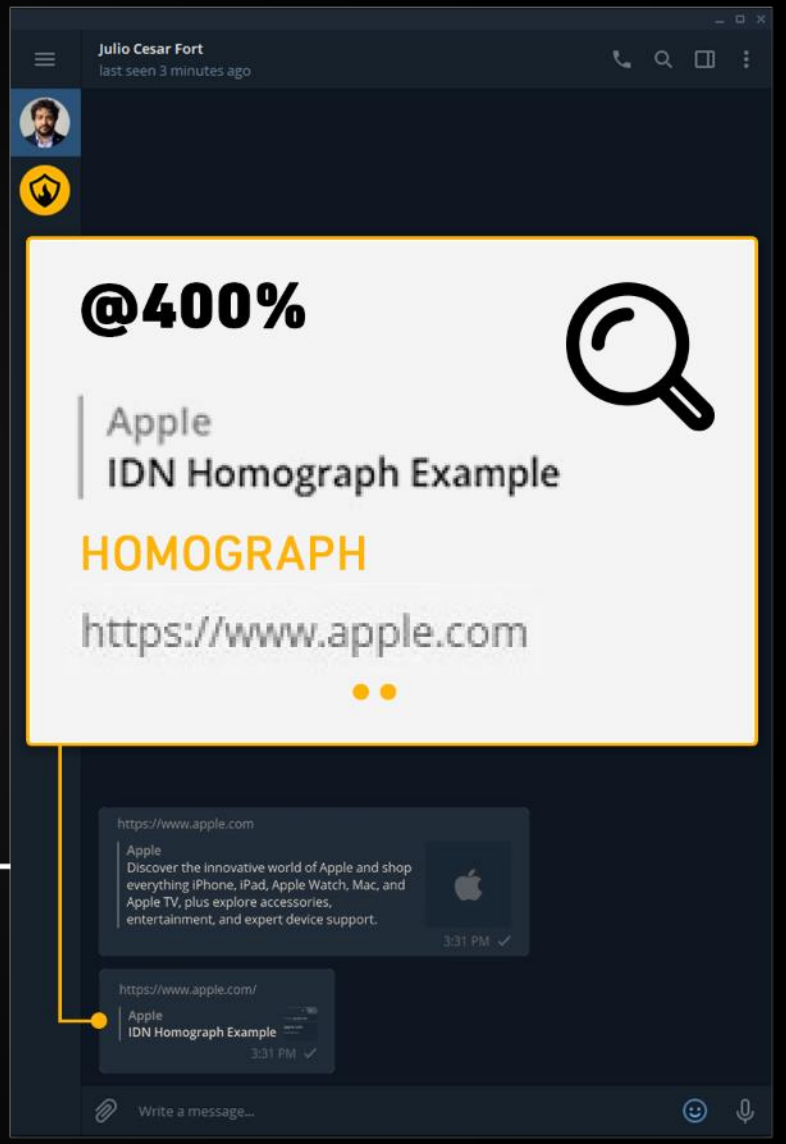
CVE-2019-9970



SECURE MESSENGER APPS:
TELEGRAM

(for Android, Windows and Linux)

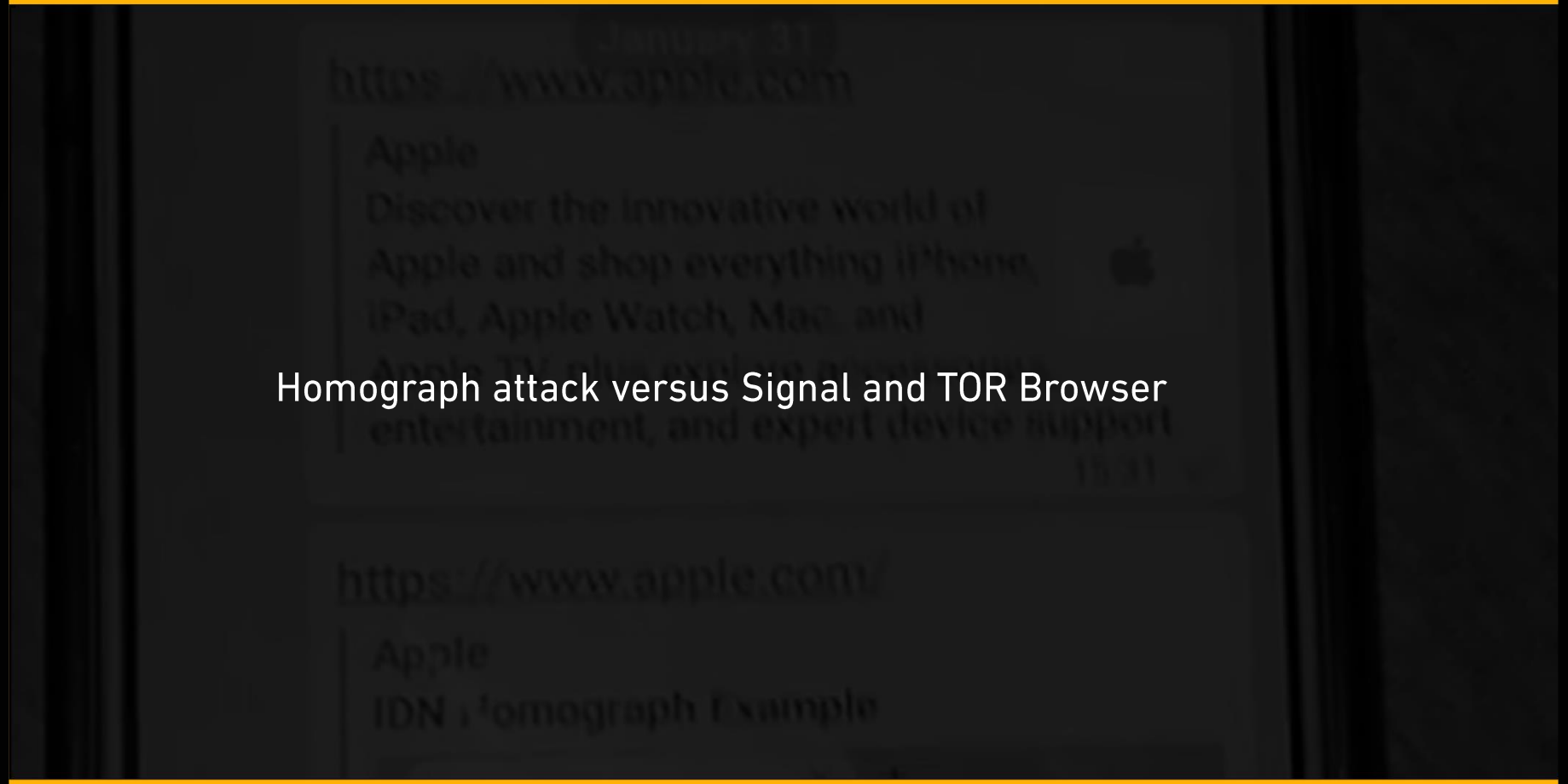
CVE-2019-10044



BLAZE

BLAZE

HOMOGRAPH ATTACK DEMONSTRATION



How to defend yourself

BLAZE

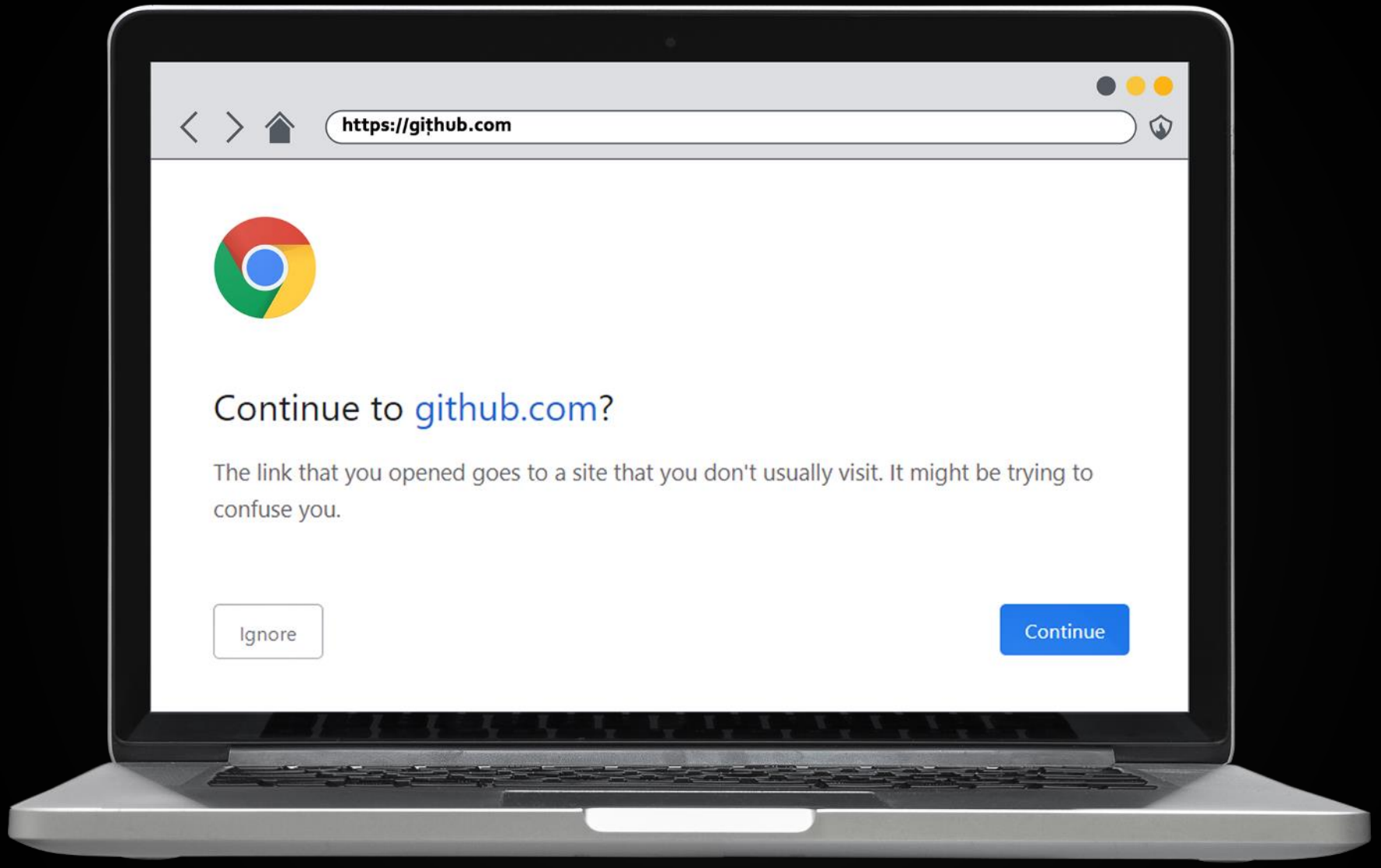
DEFENSES

FOR BROWSERS

- **Preferably, use Google Chrome**
- **Phish.ai** Chrome extension
- **Firefox: about:config** - switch idn_show_punycode to “true”

FOR EMAIL

- **Outlook, ProtonMail, Tutanota**, are fine; **Mailbird** and **Thunderbird** are also good.
- Other popular ones, not so much



Google Chrome Browser

DEFENSES

HUMAN EYE PERSPECTIVE

- Coloring confusable characters: a proposal that never took off

APPLICATION DEVELOPERS

- Use libraries to check for confusables when converting from punycode to Unicode

Conclusion

BLAZE

CONCLUSION

Confusable homographs have been around for a while, yet are **frequently overlooked** and little has been discussed about them

CONCLUSION

These issues are not always part of the threat model of some applications, as they are oftentimes considered as social engineering

CONCLUSION

Application security teams can do more and be **proactive at preventing such threats** instead of relying on users to be vigilant or ICANN to come up with a magic solution for the issue

REFERENCES

RESEARCH REFERENCE

- https://www.chromium.org/developers/design-documents/idn-in-google-chrome
- https://www.xudongz.com/blog/2017/idn-phishing/
- https://dev.to/logan/homographs-attack--5a1p
- https://blog.blazeinfosec.com/what-you-see-is-not-what-you-get-when-homographs-attack/
- Large Scale Detection of IDN Domain Name Masquerading (Elsayed and Shosha)
- An Assessment of Internationalised Domain Name Homograph Attack Mitigation Implementations (Peter Hanay)
- The Homograph Attack (Gabrilovich and Gontmakher)

Be secure. Be ahead. Be Blaze.

THANK YOU!

Questions?

BLAZE

www.blazeinfosec.com

info@blazeinfosec.com

Brazil
Rua Visconde de Jequitinhonha
279. Office 701. Recife

Portugal
Praça Bom Sucesso 131
Península, Office 206, Porto

Poland
Rynek Główny 28
33-332, Kraków

BR: +55 81 3071 7148
PT: +351 222 463 641
PL: +48 792 436 755

BLAZE