El costo del sofware de baja calidad (CPSQ) en EE.UU.: 2022

Del problema a las soluciones

Herb Krasner
30 Marzo, 2023
hkrasner@utexas.edu

https://www.it-cisq.org/the-cost-of-poor-quality-software-in-the-us-a-2022-report/

Conferencista

Herb Krasner

- Autor de los informes del Costo del Software de Baja Calidad en EE.UU.: Informes del 2018, 2020, y 2022
- Profesor jubilado de Ingeniería de Software en UT (University of Texas).
- Miembro de la Directiva de CISQ.
- Científico y líder de investigación empírica para la Ingeniería de Software.
- Más de 50 años en la industria de consultoría de software. Ha ayudado a cientos de compañías a mejorar su desarrollo de software y capacidad de entrega.

La calidad del software no es gratuita, pero es más económica que el software de mala calidad

(mis disculpas a Phil Crosby)

El modelo Iceberg de CPSQ

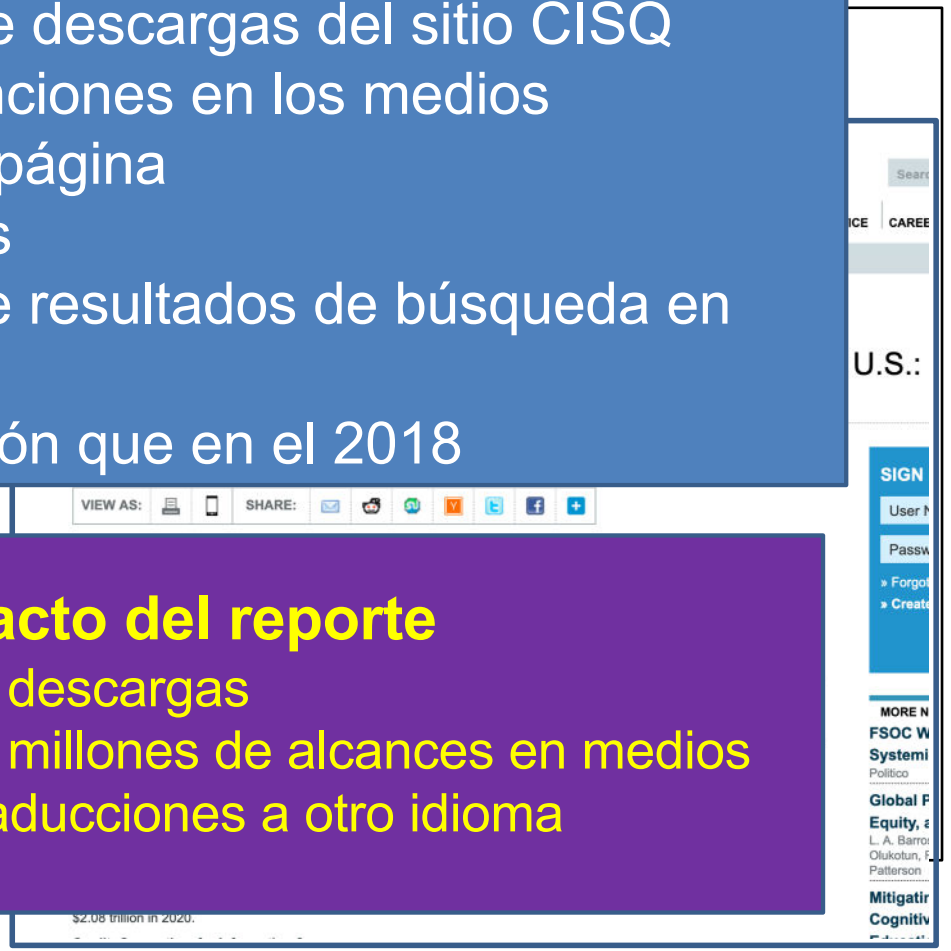
Costos ocultos = 6 a 50 veces los costos visibles.

Propósito original: Informar e inspirar a los lectores a buscar conocimiento de CPSQ dentro de sus organizaciones

Resultado: Una aproximación de primer orden de la magnitud de este enorme y poco reconocido problema

El impacto del informe 2020

- Decenas de miles de descargas del sitio CISQ
- **500,000,000** de menciones en los medios
- Miles de visitas a la página
- Traducido al japonés
- Más de 8 páginas de resultados de búsqueda en Google
- Mucho mayor atención que en el 2018



Impacto del reporte

- 400 descargas
- 229 millones de alcances en medios
- 2 traducciones a otro idioma

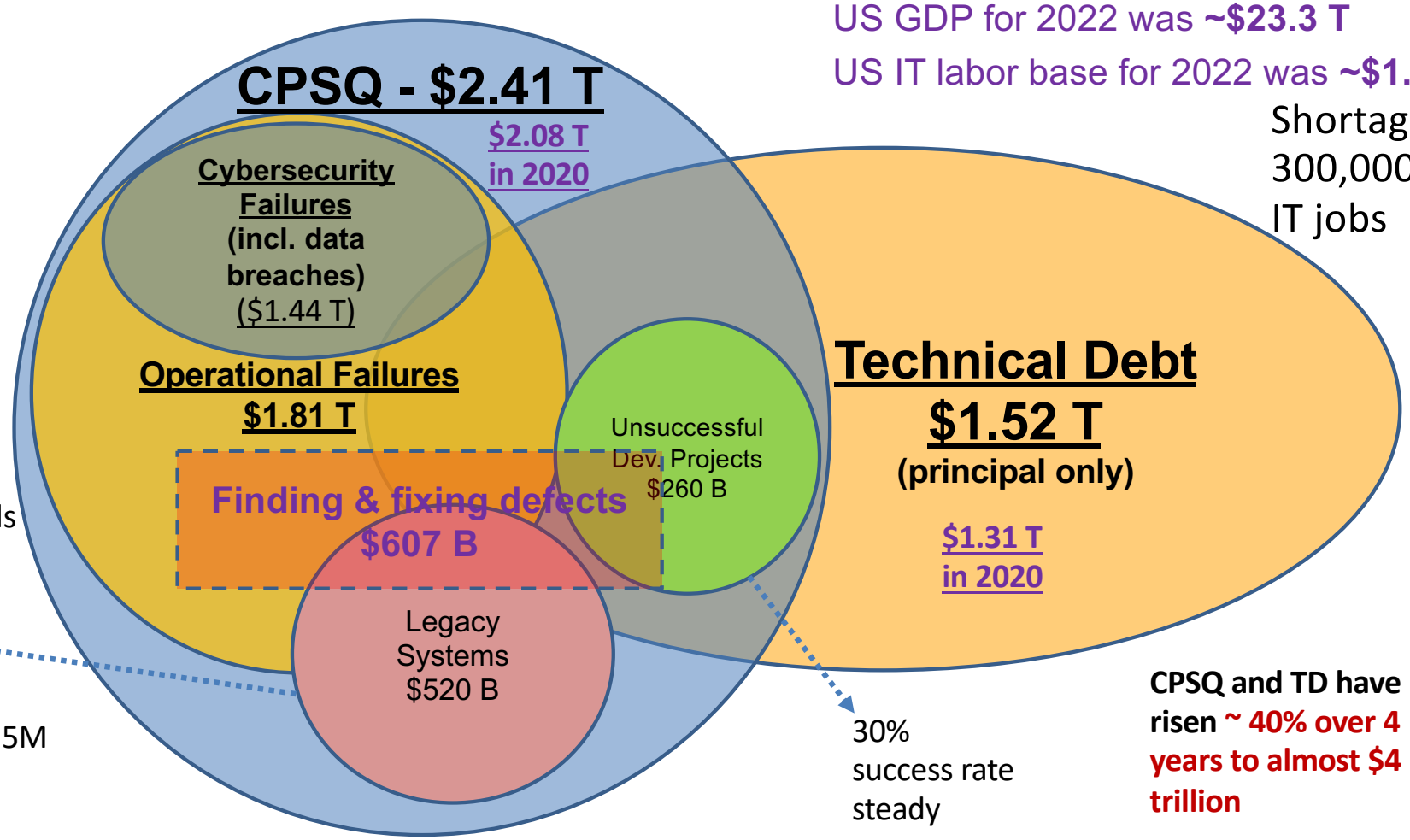
2022 Summary of CPSQ Estimates

US GDP for 2022 was **~$23.3 T**

US IT labor base for 2022 was **~$1.51 T**

Shortage of 300,000 IT jobs

45% per year



A recent survey of 250 IT professionals stated that the majority of application modernization projects fail at an average cost of $1.5M and 16 months of wasted time

30% success rate steady

CPSQ and TD have risen ~ 40% over 4 years to almost $4 trillion

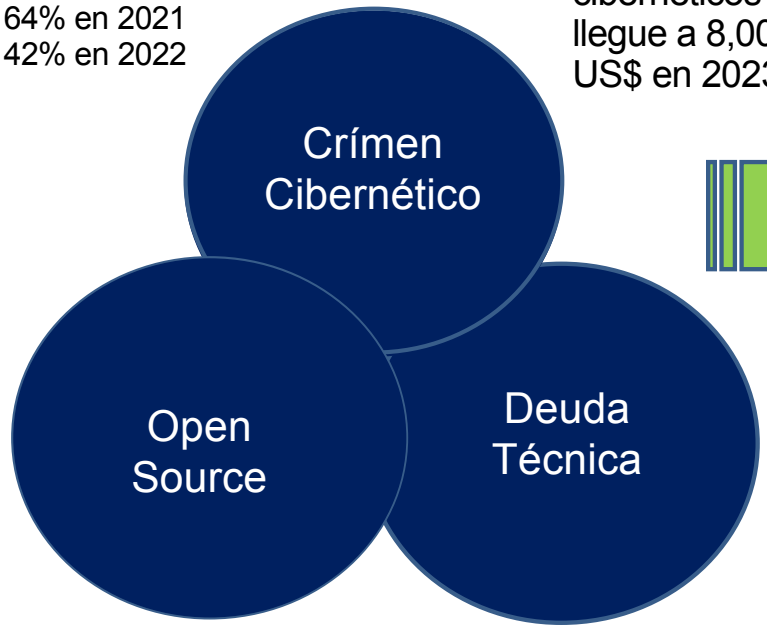
Puntos Importantes del Reporte 2022

Areas de problema

Costos incrementales debido a vulnerabilidad de SW:

- 64% en 2021
- 42% en 2022

El costo total de los crímenes cibernéticos se espera que llegue a 8,000 millones de US$ en 2023.



Las fallas debido a deficiencias en el *open source* subieron 650% en 2021

Es hoy en día el mayor obstáculo para el cambio de las bases de código existentes

Areas de Solución

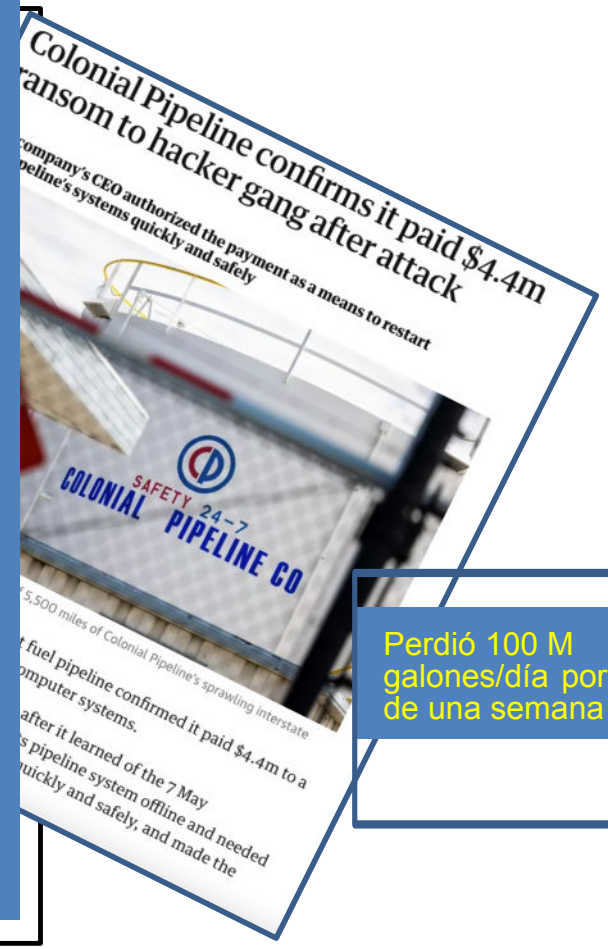
AI no generativo

Para el 2025, los ataques a la cadena de suministro de software se triplicarán con respecto al 2021.

Fallas Operacionales: $1.81T en 2022

Log4Shell

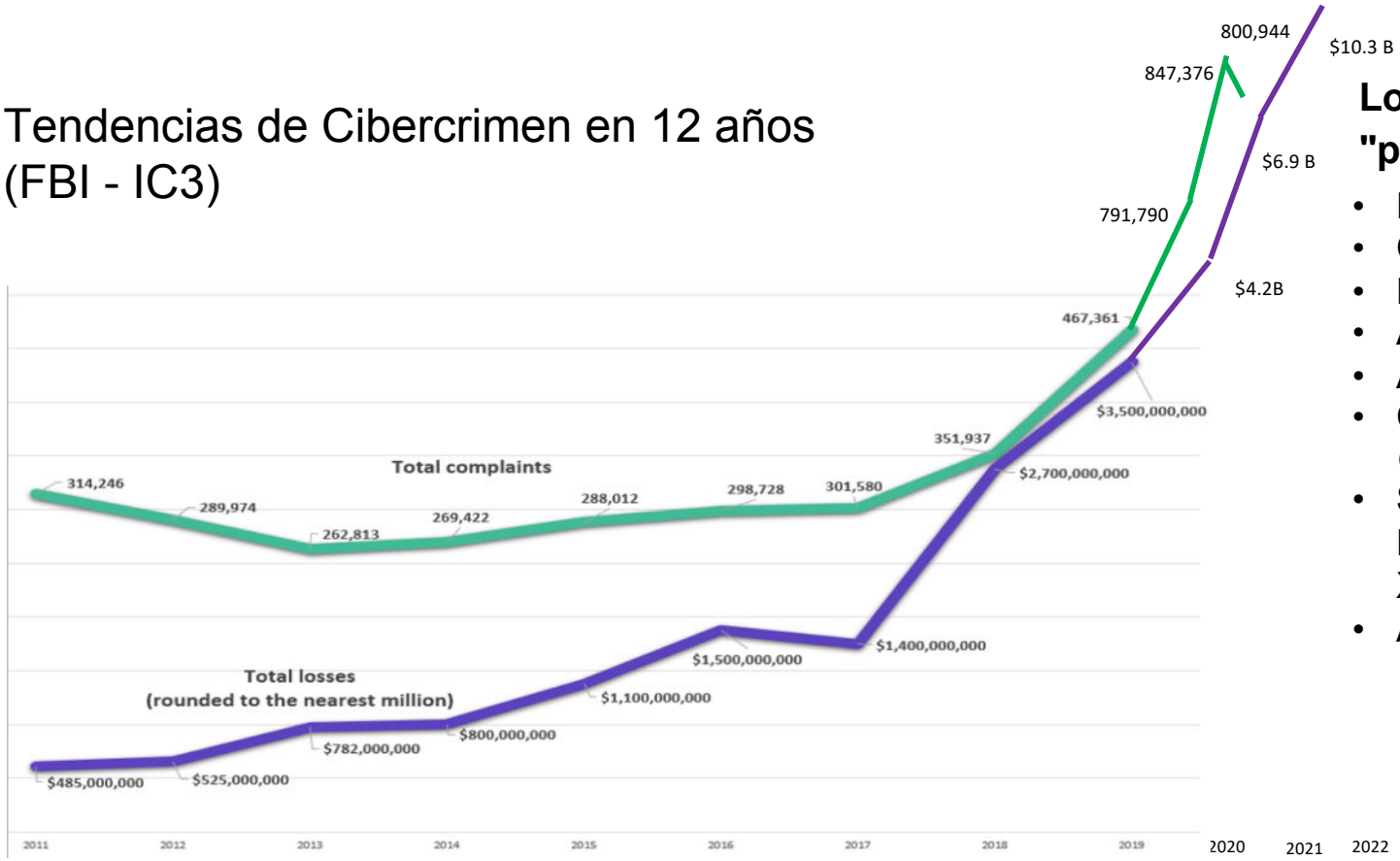
T-Mobile data breach
TikTok glitch
Twitter ID exposure
Facebook ranking bug
Grand Theft Auto
Wintermute crypto heist
Spyder Python IDE
BNB Chain
GIT /GitHub
OpenLightSpeed
FTX Crypto Hack
CommonSpirit Health
OpenSSL
Microsoft Exchange Server
Tesla recalls



Perdió 100 M galones/día por más de una semana

Incremento de pérdidas por cibercrimen

Tendencias de Cibercrimen en 12 años
(FBI - IC3)

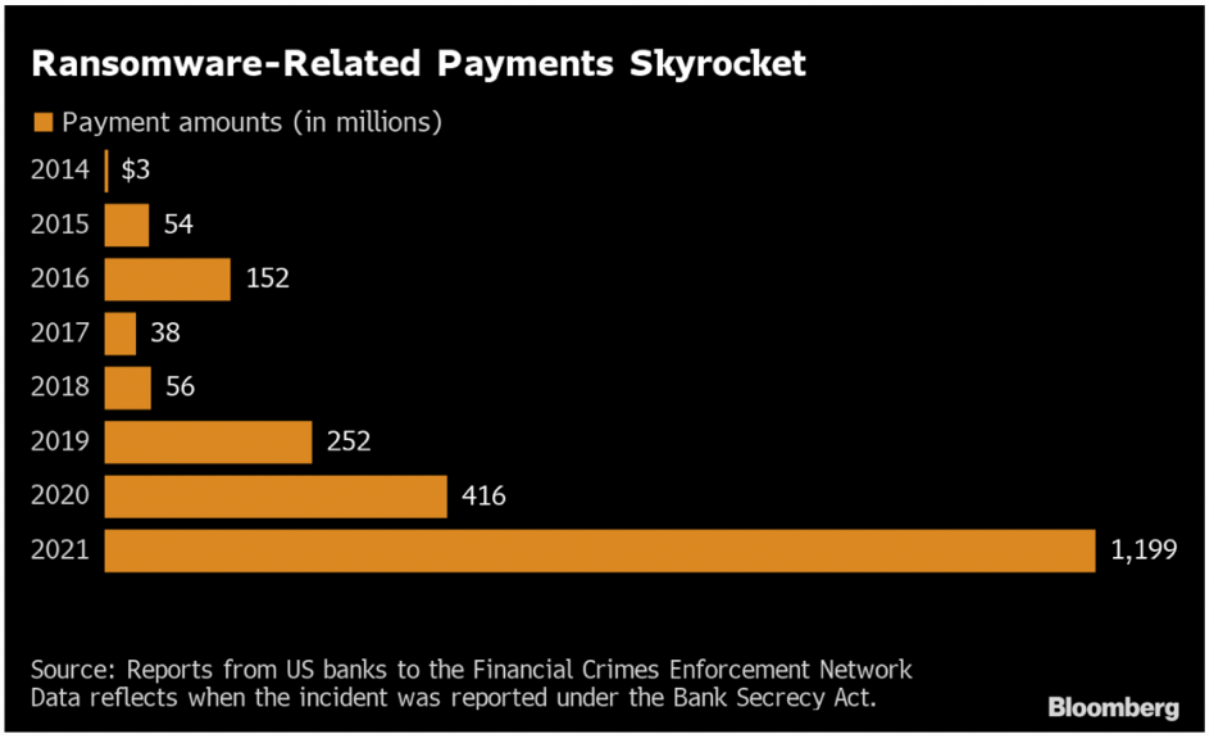


Los cibercrímenes más "populares" :

- Ransomware
- Cryptojacking
- Deepfakes
- Ataques a videoconferencias
- Ataques IoT y OT
- Cadena abastecimiento/ataques *Open Source*
- Soluciones de Detección y Respuesta extendidos (aka XDR)
- Ataques a infraestructura crítica

IC3 – Internet Crime Complaint Center

Incremento de pérdidas por el cibercrimen



Source: Bloomberg

Le toma 277 días a una empresa promedio identificar y contener una fuga de datos, con un costo promedio de hasta US$4.24 millones (IBM)

Sólo en la primera mitad de 2022, más de 236 millones de ataques ransomware se reportaron en todo el mundo

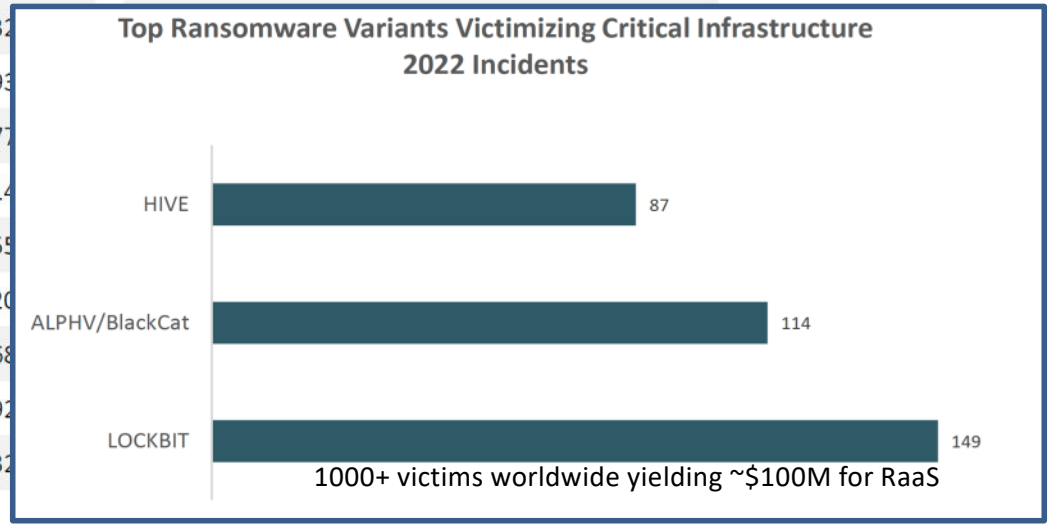
En 2022, el mercado global de cripto perdió 2,000 millones de US$

Tip: Revisar aplicaciones en Python que estén involucradas en transacciones cripto.

IC3 - Resumen de costos del Cibercrimen 2022

By Victim Loss			
Crime Type	*Loss*	*Crime Type*	*Loss*
Investment	$3,311,742,206	Lottery/Sweepstakes/Inheritance	$83,602,376
BEC (email compromise)	$2,742,354,049	SIM Swap	$72,652,571
Tech Support	$806,551,993	Extortion	$54,335,128
Personal Data Breach	$742,438,136	Employment	$52,204,269
Confidence/Romance	$735,882,192	Phishing	$52,089,159
Data Breach	$459,32		
	$396,93		
on-Delivery	$281,77		
k Fraud	$264,14		
ersonation	$240,55		
	$189,20		
	$117,68		
	$107,92		
	$104,32		
*Descriptors***			
Cryptocurrency	$2,496,196,530	Cryptocurrency Wallet	$1,349,090,883

¿Sabía usted que la calificación crediticia de su empresa baja, cuando sube el riesgo de ciberseguridad?



Open Source Software (OSS)

Algunos datos del SW *Open Source:*

- Casi todas las empresas utilizan *Software Open Source (ya sea que lo sepan o no)*
- *Una aplicación mediana (< 1 millón de SLOC) conlleva de 200 a 300 componentes de terceros (en promedio)*
- *El 80% de las organizaciones reportaron un incremento en el uso de SW Open Source.*
- *Del 68% al 75% de las organizaciones que utilizan SW Open Source reconocen que han sido impactadas por vulnerabilidaes en los últimos dos años.*
- *Solo el 46% de las organizaciones están realizando revisiones de seguridad para identificar vulnerabilidades del SW Open Source*

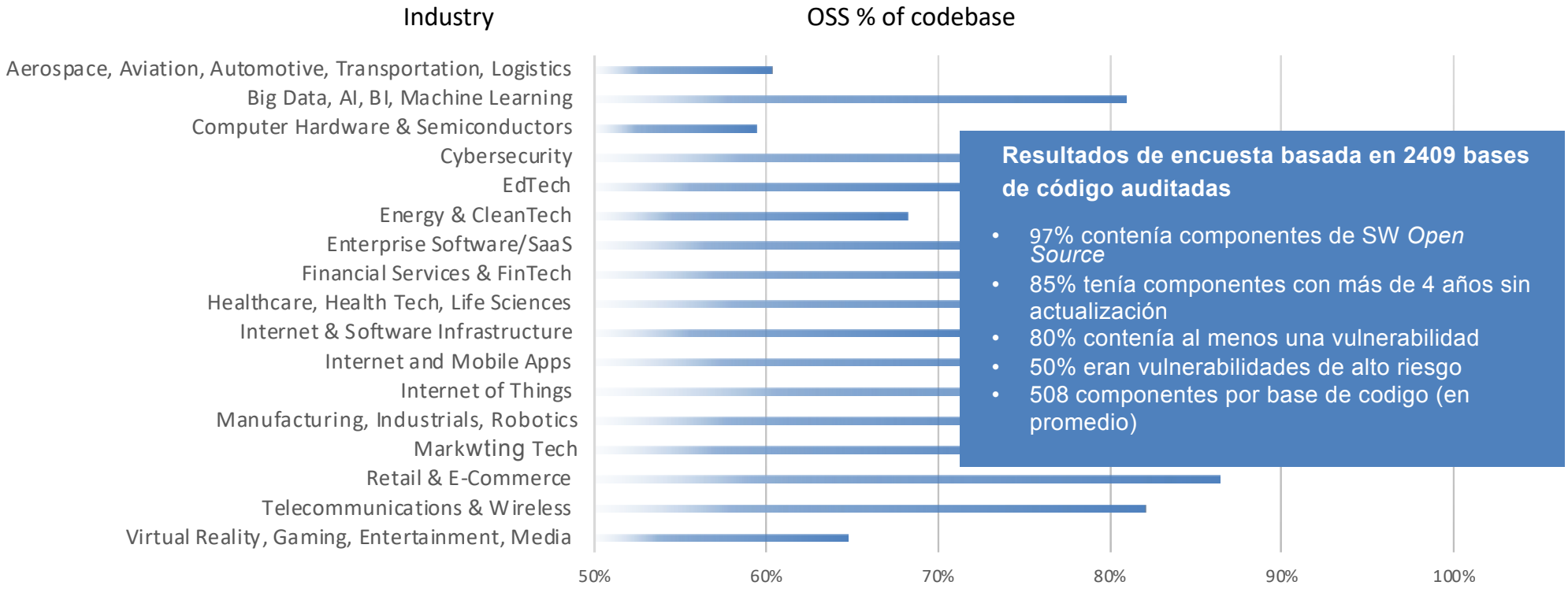
Las 10 vulnerabilidades más críticas y altas del SW *Open Source* fueron:

- Denegación de servicio (DoS)
- Ejecución remota de código (RCE)
- Deserialización de datos que no son de confianza
- Inyección SQL
- Prototipo de contaminación
- Archivo temporal inseguro
- Directorio/ruta transversa
- Escalada de privilegios
- Denegación de servicio de expresión regular (ReDoS)
- Desreferencia de puntero NULL

* de Snyk scans

Synopsys: Encuesta de componentes de SW Open Source

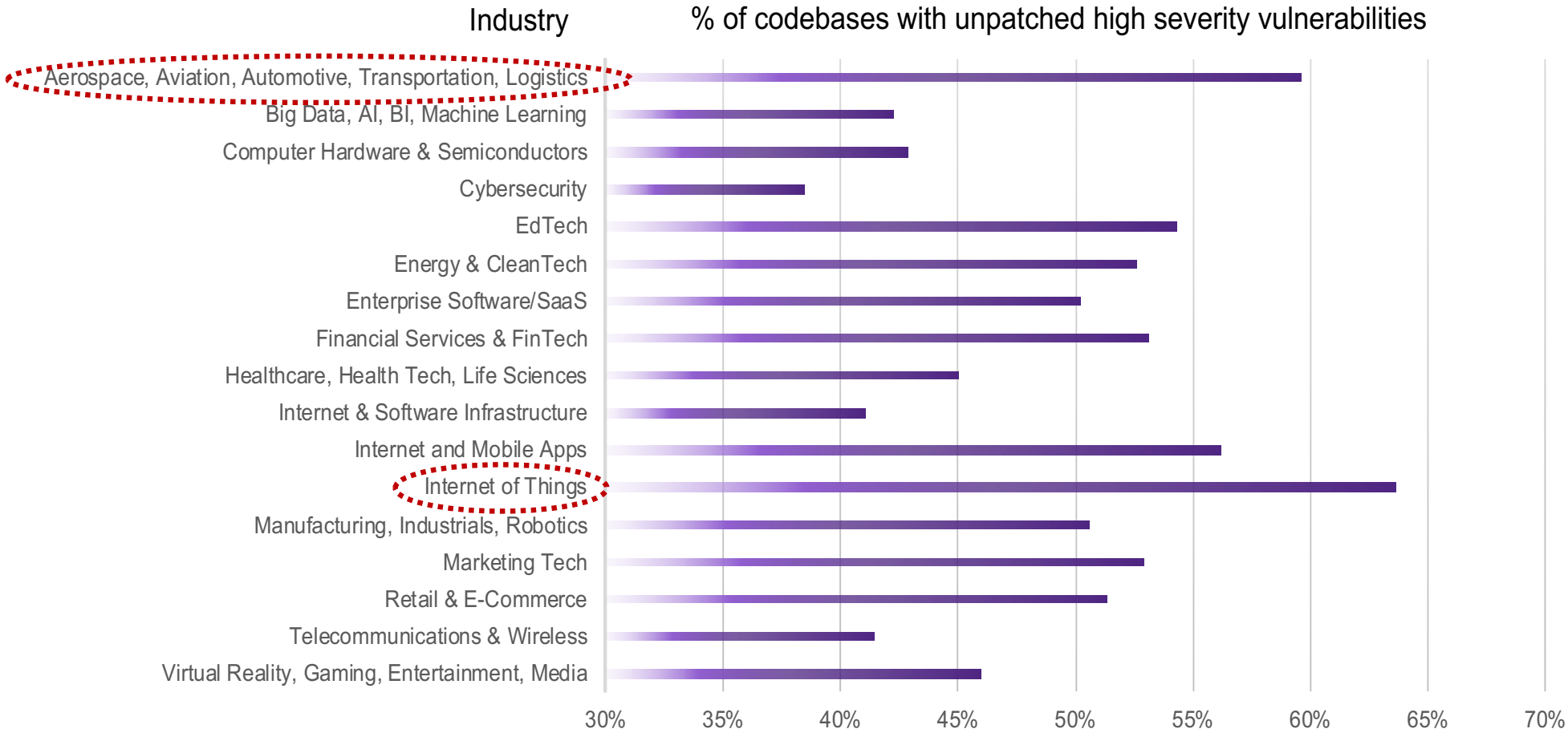
¿Qué cantidad de su base de código es SW *Open Source*?



Synopsys 2022 OSSRA Report

Synopsys OSS component survey

% with unpatched high severity bugs



Deuda Técnica (DT)

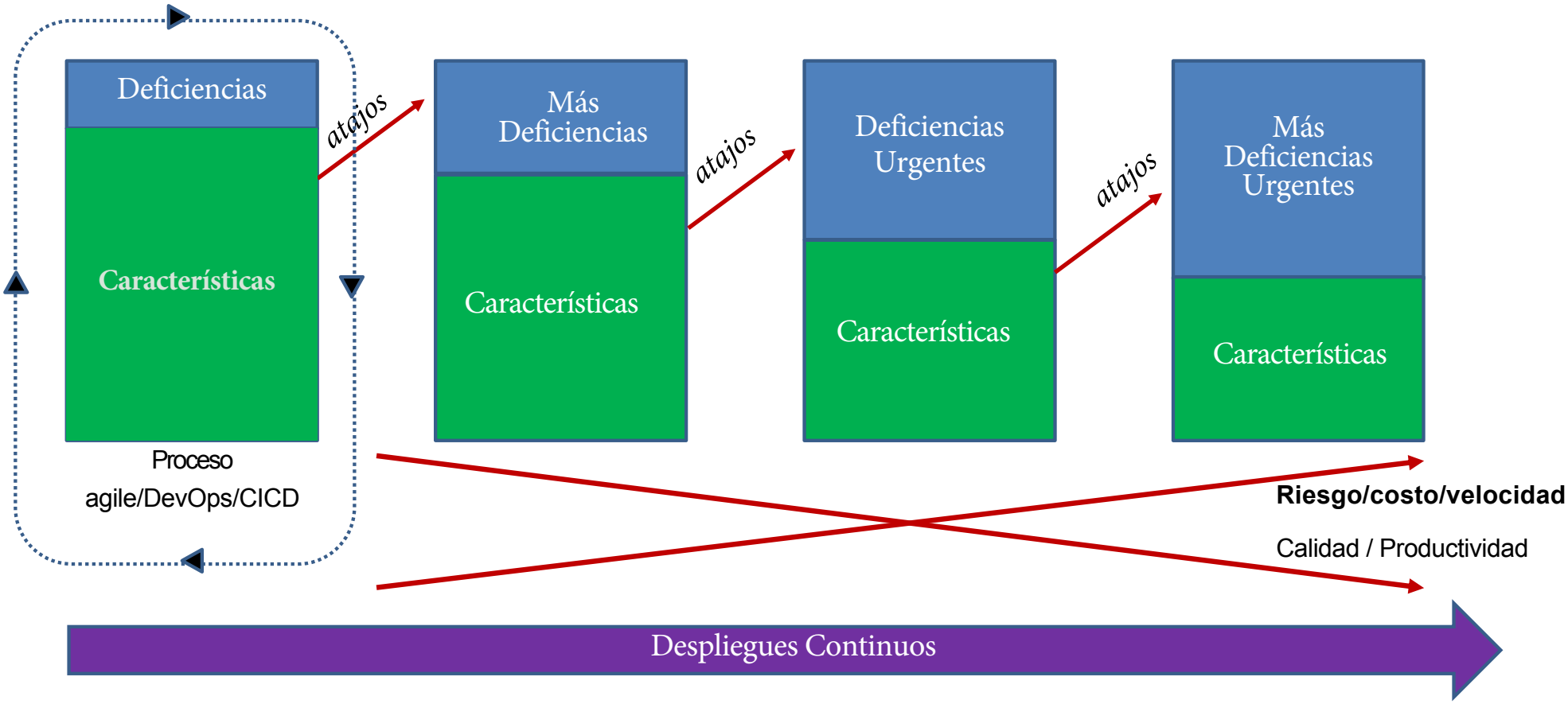
Es una medida de los costos de una organización por sistemas de TI deficientes, inflexibles y obsoletos. Es un problema de gran magnitud (1,520 millones de US$ y crece aceleradamente):

- El total de DT es entre un 20% y un 40% del valor total de de tecnología de una organización antes de depreciación.
- El 78% de las organizaciones ha acumulado más DT en el año pasado que en los anteriores.
  - pero solamente el 42% de las empresas sienten que pueden evaluar toda su DT.
- El 87% de los CIOs del mundo admitieron que la complejidad de sus sitemas de TI existentes no les permite invertir en servicios de la próxima generación **(¿que están haciendo al respecto?)**
- El promedio de horas que un desarrollador tiene que invertir en temas de DT es 13.5 de un total de 41.1, o sea el 33% de su tiempo.
- Investigaciones indican que el 70% de los equipos de desarrollo siempre o muy frecuentemente se saltan los pasos de seguridad debido a las presiones de tiempo cuando completan projectos (Mend).
- La DT será mucho peor si no se hace nada al respecto.

Southwest Airlines cancela 17,000 vuelos a finales de diciembre debido a DT en su viejo sistema de horarios de la tripulación. Más de 1M de pasajeros quedaron varados**, 1,100 millones de US$ en costos directos + $4 mil millones en pérdidas de mercado ese mes**

Invertirán más de $1,300 millones en actualizaciones y mantenimiento de sus sistemas de TI en 2023.

¿Como se acumula la DT en el tiempo?



Como se acumula la DT en el tiempo

Un estudio de los beneficios de refactoring en Microsoft arroja los siguientes porcentajes de ganancias

- Mejora en legibilidad (43%)
- Mejora en mantenimiento (30%)
- Mejora en flexibilidad y facilidad para agregar nuevas funcionalidades (27%)
- Menos errores de programación (27%)
- Modularidad mejorada (19%)
- Reducción en código duplicado (18%)
- Desempeño mejorado (12%)
- Reducción del tamaño del código (12%)
- Mejora en las pruebas (12%)
- Reducción del tiempo en lanzamiento (5%)

Stepsize: *-Las empresas que administran la DT estarán desplegando al menos un 50% más rápido-*

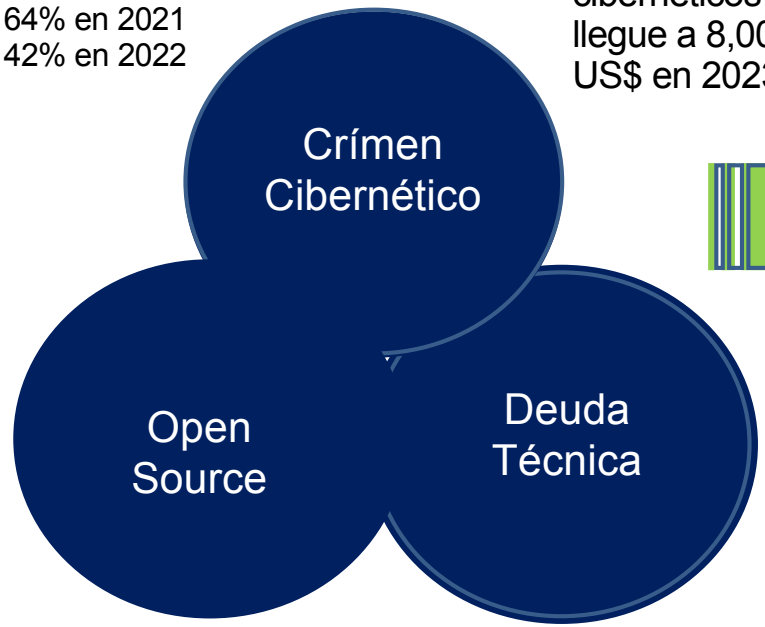
Puntos Importantes del Reporte 2022

Areas de problema

Costos incrementales debido a vulnerabilidad de SW:

- 64% en 2021
- 42% en 2022

El costo total de los crímenes cibernéticos se espera que llegue a 8,000 millones de US$ en 2023.



Las fallas debido a deficiencias en el *open source* subieron 650% en 2021

Es hoy en día el mayor obstáculo para el cambio de las bases de código existentes

Areas de Solución

AI no generativo

Para el 2025, los ataques a la cadena de suministro de software se triplicarán con respecto al 2021.

Estándares de calidad de Software

Si existiera una medida simple para un "buen" software, todos lo usaríamos y solicitaríamos esa calidad.

El marco más útil y común para ayudar en cada proyecto para definir las metas de calidad del software con precisión es el ISO/IEC 25000

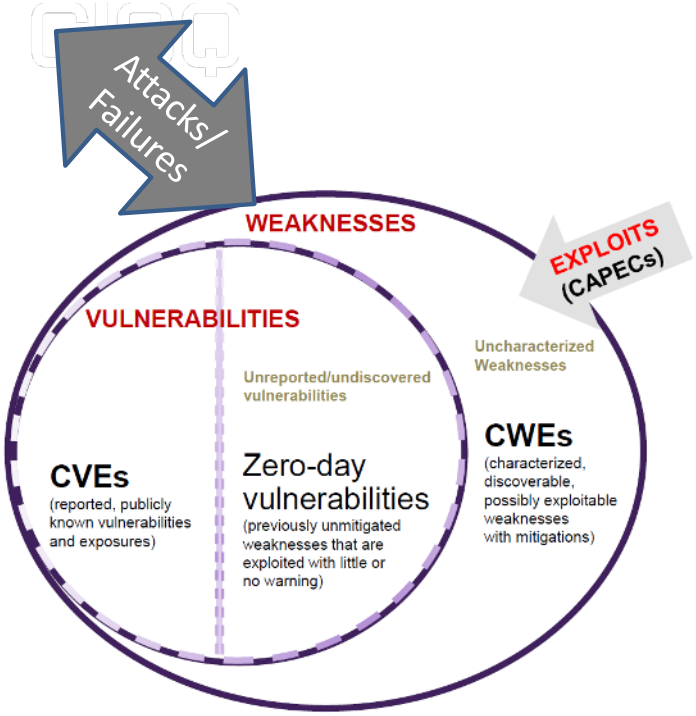
- ISO/IEC 25010 Define 8 características de calidad de software, o - dades , por ej: securidad, confiabilidad y mantenibilidad.
- ISO/IEC 25023 Describe como aplicar esas características para medir la calidad del producto de software a nivel de su comportamiento.

Por tanto se necesitaba una medida para el logro de la meta de calidad del software a nivel del código fuente.

- ISO/IEC 5055:2021* Define medidas de nivel de código fuente para 4 de las 8 características de calidad - confiabilidad, desempeño, eficiencia, seguridad y mantenibilidad.
- Cada característica de calidad de código ISO 5055 se compone de un grupo de debilidades seleccionadas de la taxonomía de CWEs de Common Weakness Enumeration. El CWE codifica más de 800 debilidades de software conocidas. Cada CWE es un patrón de código conocido que es una potencial falla en muchos sistemas existentes.
- Estos ahora se están automatizando en el desarrollo de nuevas herramientas de análisis de calidad de código

* por CISQ

La Taxonomía de CWE



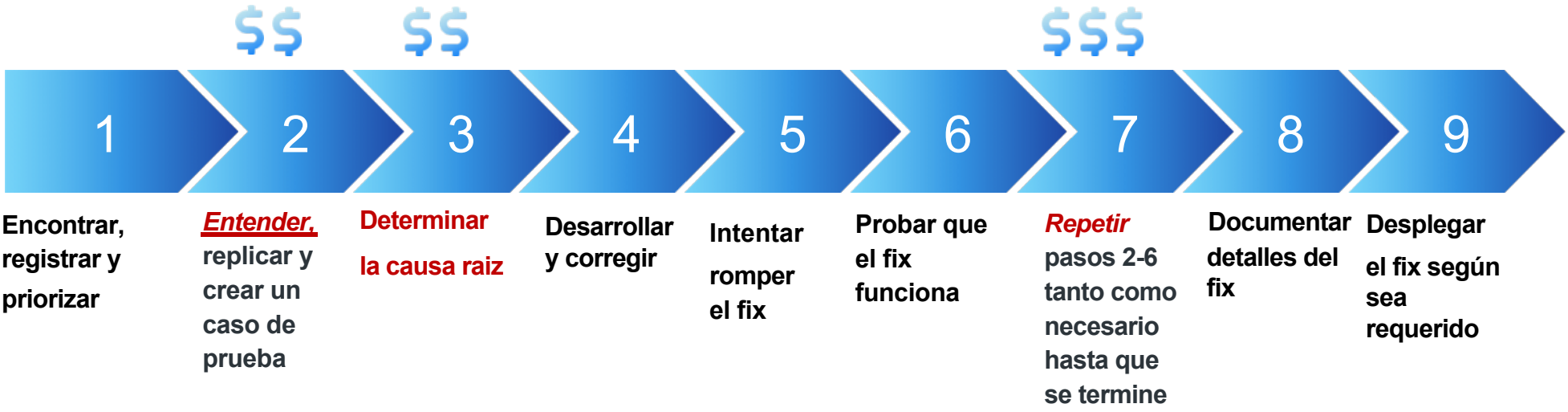
Sólo el 29% de los 12,000 desarrolladores entrevistados creen que la práctica activa de escribir código libre de vulnerabilidades debería priorizarse (¿Por qué?)

Common Weakness Enumeration (CWE)
Common Vulnerability Enumeration (CVE)

Las 25 CWEs más importantes de 2022 (MITRE)

	ID	Name	Score	KEV Count (CVEs)	Rank Change vs. 2021
1	CWE-787	Out-of-bounds Write	64.20	62	0
2	CWE-79	Improper Neutralization of Input During Web Page Generation ('Cross-site Scripting')	45.97	2	0
3	CWE-89	Improper Neutralization of Special Elements used in an SQL Command ('SQL Injection')	22.11	7	+3
4	CWE-20	Improper Input Validation	20.63	20	0
5	CWE-125	Out-of-bounds Read	17.67	1	-2
6	CWE-78	Improper Neutralization of Special Elements used in an OS Command ('OS Command Injection')	17.53	32	-1
7	CWE-416	Use After Free	15.50	28	0
8	CWE-22	Improper Limitation of a Pathname to a Restricted Directory ('Path Traversal')	14.08	19	0
9	CWE-352	Cross-Site Request Forgery (CSRF)	11.53	1	0
10	CWE-434	Unrestricted Upload of File with Dangerous Type	9.56	6	0
11	CWE-476	NULL Pointer Dereference	7.15	0	+4
12	CWE-502	Deserialization of Untrusted Data	6.68	7	+1
13	CWE-190	Integer Overflow or Wraparound	6.53	2	-1
14	CWE-287	Improper Authentication	6.35	4	0
15	CWE-798	Use of Hard-coded Credentials	5.66	0	+1
16	CWE-862	Missing Authorization	5.53	1	+2
17	CWE-77	Improper Neutralization of Special Elements used in a Command ('Command Injection')	5.42	5	+8
18	CWE-306	Missing Authentication for Critical Function	5.15	6	-7
19	CWE-119	Improper Restriction of Operations within the Bounds of a Memory Buffer	4.85	6	-2
20	CWE-276	Incorrect Default Permissions	4.84	0	-1
21	CWE-918	Server-Side Request Forgery (SSRF)	4.27	8	+3
22	CWE-362	Concurrent Execution using Shared Resource with Improper Synchronization ('Race Condition')	3.57	6	+11
23	CWE-400	Uncontrolled Resource Consumption	3.56	2	+4
24	CWE-611	Improper Restriction of XML External Entity Reference	3.38	0	-1
25	CWE-94	Improper Control of Generation of Code ('Code Injection')	3.32	4	+3

Entender, encontrar y corregir fallas



Un porcentaje significativo (50%+) de los costos de proyectos de sofware hoy en día, no se gasta en la actividad creativa de construcción de software, sino más bien en la actividad de depurar y corregir errores.

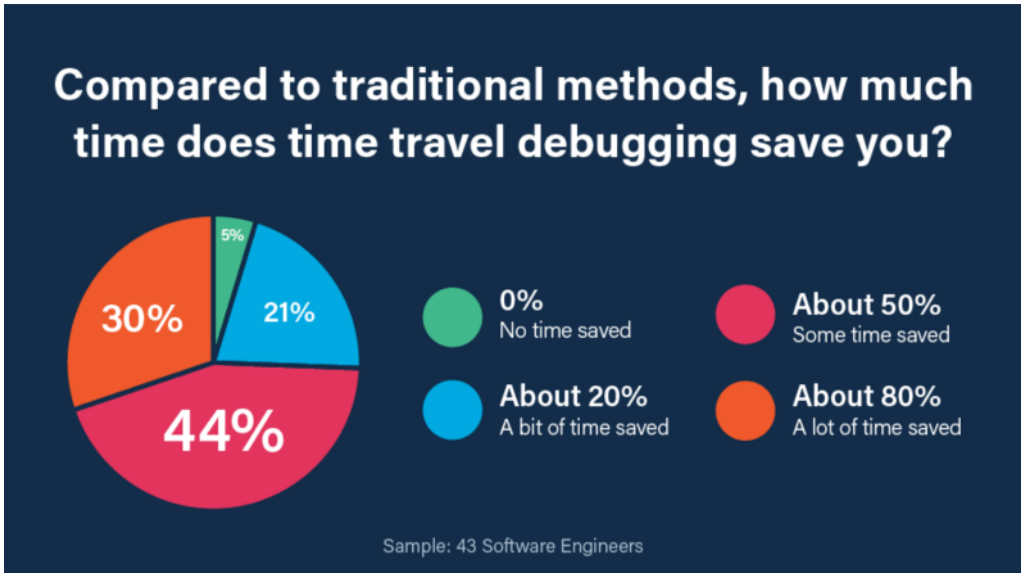
Donde buscar

Recomendaciones prácticas para encontrar deficiencias, vulnerabilidades, debilidades, errores y oportunidades de refactoreo:

- A nivel de código algunos lenguajes son más sensibles a errores que otros (e: JavaScript - fácil de aprender y fácil de introducir errores con éste).
- A nivel de diseño y arquitectura, entre más complejos los componentes y sus interdependencias, más probable que se generen errores complicados y confusos.
- Interfaces de usuario de alta interacción es más probable que contengan errores.
- Librerías de terceros son más propensas a contener errores. Con el SW *Open Source se* puede acceder al código. Si es código cerrado, el fabricante es responsable.
- La primera versión de algo nuevo se garantiza que tiene muchos errores.
- Los errores que representan situaciones inusuales y no anticipadas que dependen del contexto de uso son de los más difíciles de encontrar.

Como buscar

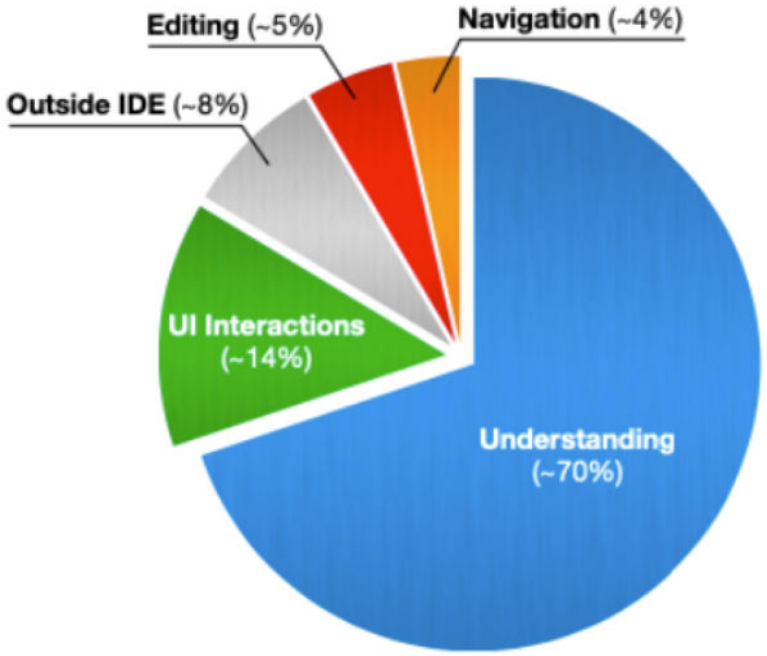
- Programas de cazarecompensas de errores - popularizados
  - En 2021, un investigador de seguridad se ganó una recompensa de $107,500 por identificar problemas de seguridad en los parlantes inteligentes de Google Home que podrían ser explotados para instalar *backdoors* y para que se convirtieran en dispositivos de grabación.
- Herramientas de análisis de código nuevo (ej: CAST, etc)
- Herramientas de reparación de programas automáticas
- Depuradores interactivos sensitivos-contexto



undo™

Las herramientas Inteligencia Artificial (AI)/ Machine Learning (ML) ya están disponibles

¿Como invierten su tiempo los desarrolladores de código?



Herramientas de ejemplo:

- GitHub Copilot
- MS BugLab
- Amazon CodeWhisperer
- Getafix
- Kite
- VS Intellicode
- Etc.

Lo que falta son las herramientas de prevención de deficiencias basadas en AI que se entrenan para conocer patrones (ej: CWEs) para detectarlas en tiempo real cuando se está escribiendo el código.

Beneficios de usar AI en automatizaciones de pruebas:

- 70% reducción en costos de mantenimiento
- 3X incremento en velocidad de pruebas de desarrollo
- 30% incremento en cobertura de pruebas

MicroFocus - 2021

Recomendaciones generales

Líderes/Ejecutivos

- Establezca la calidad como un *ciudadano de primera clase* - adopte DevQualOps
- Haga mejores preguntas: interna y externamente
- Mida la calidad de software y CPSQ en su organización

Empresas:

- **Atienda proactivamente los defectos de software y vulnerabilidades que generan a fallas**
- **Reduzca siginifcativamente la carga de DT existente**

Equipos/Proyectos

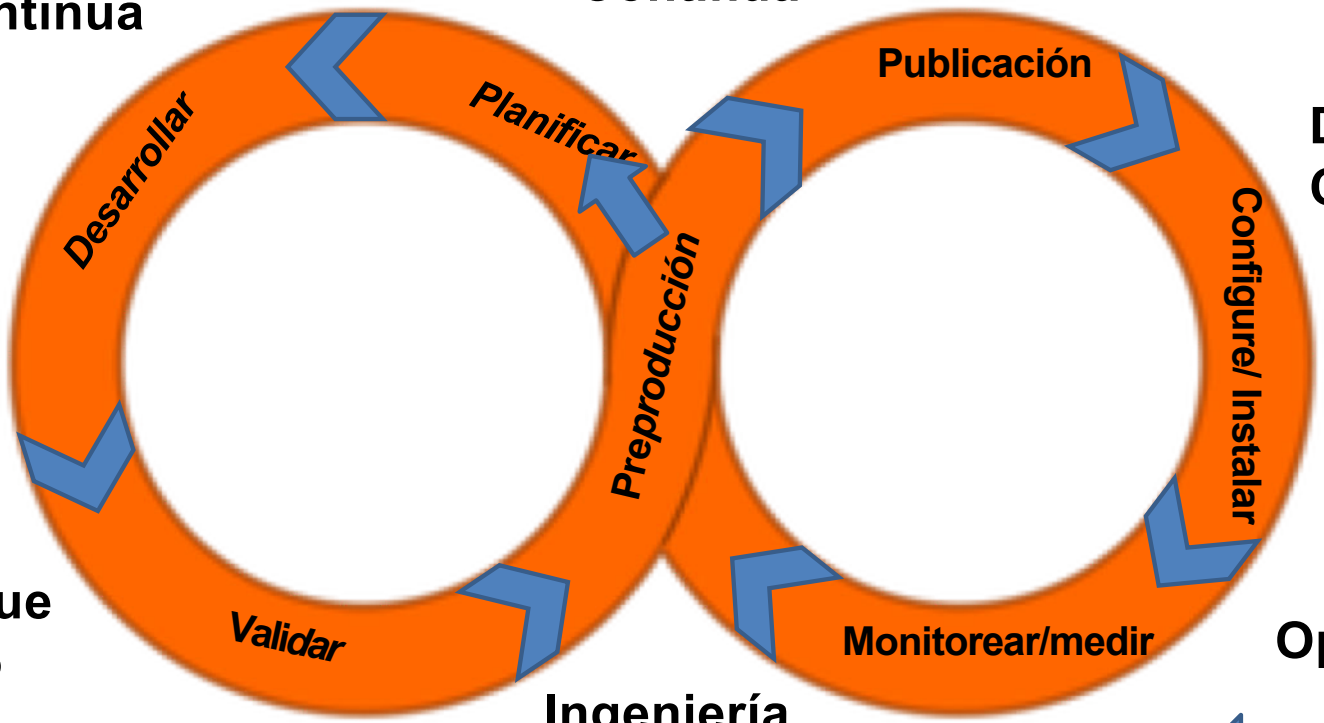
- Insista en alto desempeño
- Utilice las mejores prácticas y herramientas.
- Defina y de seguimiento a los objetivos de calidad - mida de acuerdo con estos objetivos
- Evite los compromisos y restricciones que sean arbitrarios y poco realistas.

Personas

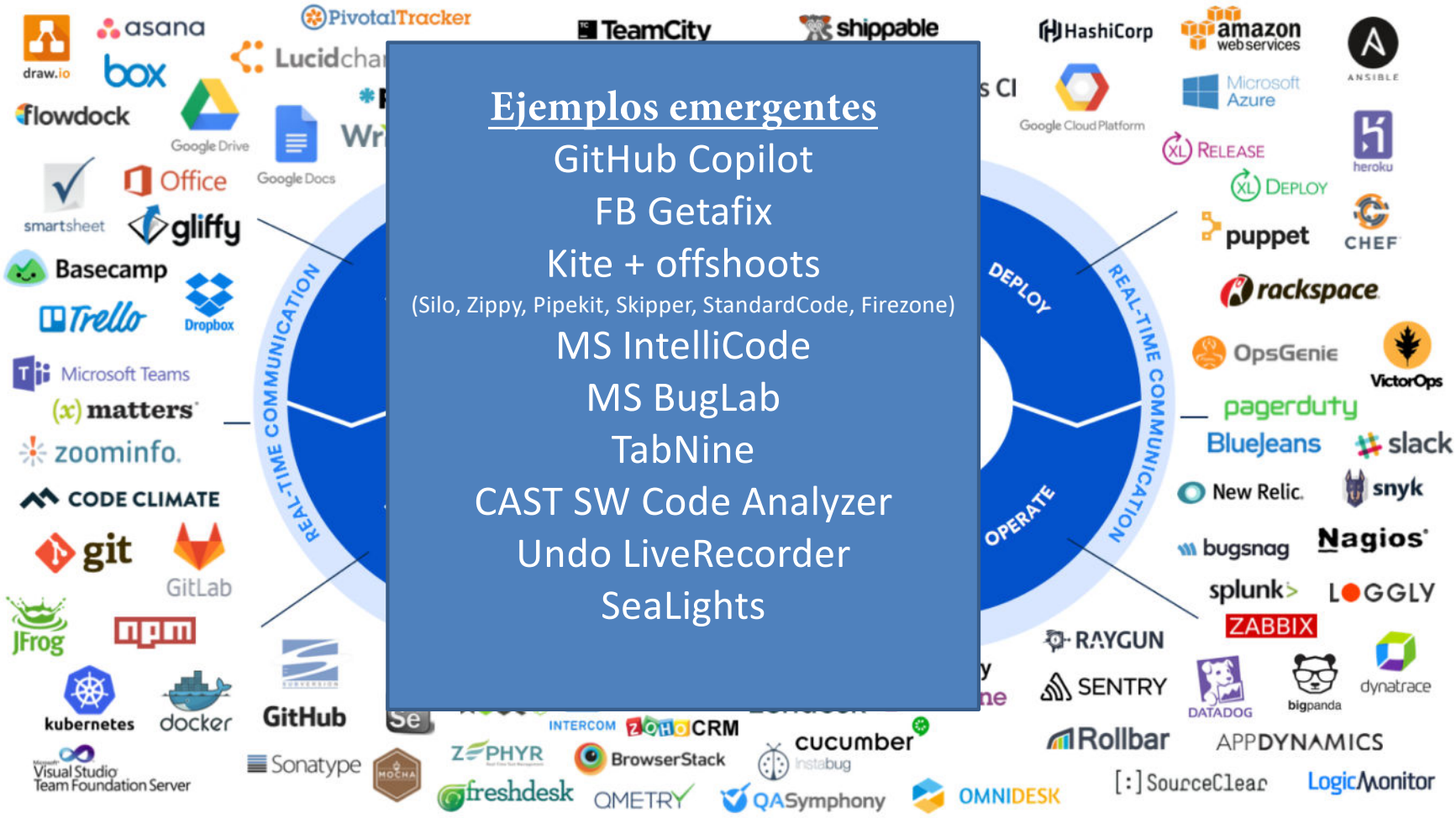
- Aprenda y desarrolle un enfoque disciplinado
  No le tema a la calidad/metricas de CPSQ
  Utilice el conocimiento existente de las fuentes de patrones de fallas y defectos de calidad estructural (nuevas herramientas IA)

Modelo DevQualOps

Existen prácticas y herramientas que apoyan este modelo

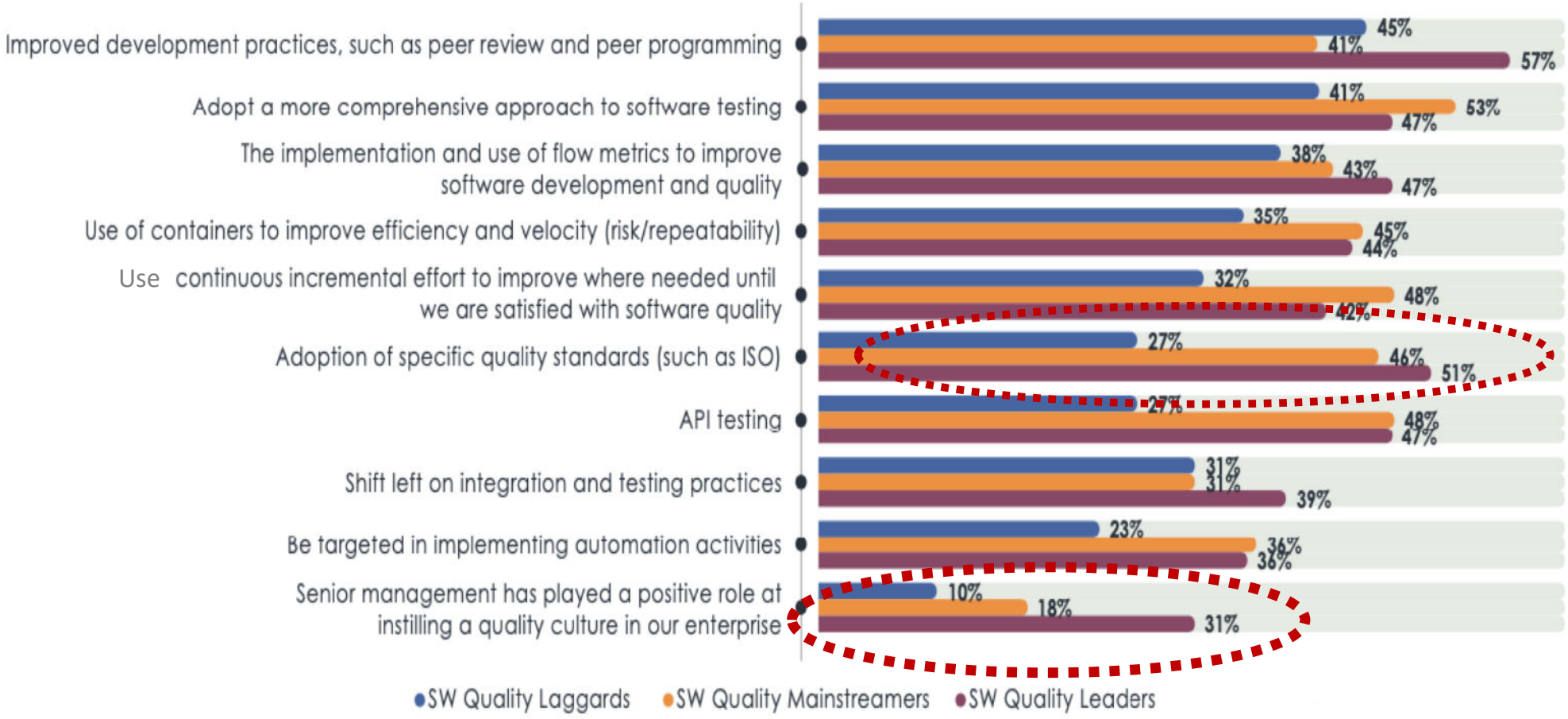


Muchas herramientas para ayudar



Continue to Use Best SQE Practices

2021 Micro Focus State of Software Quality Survey



Use the DevQualOps model

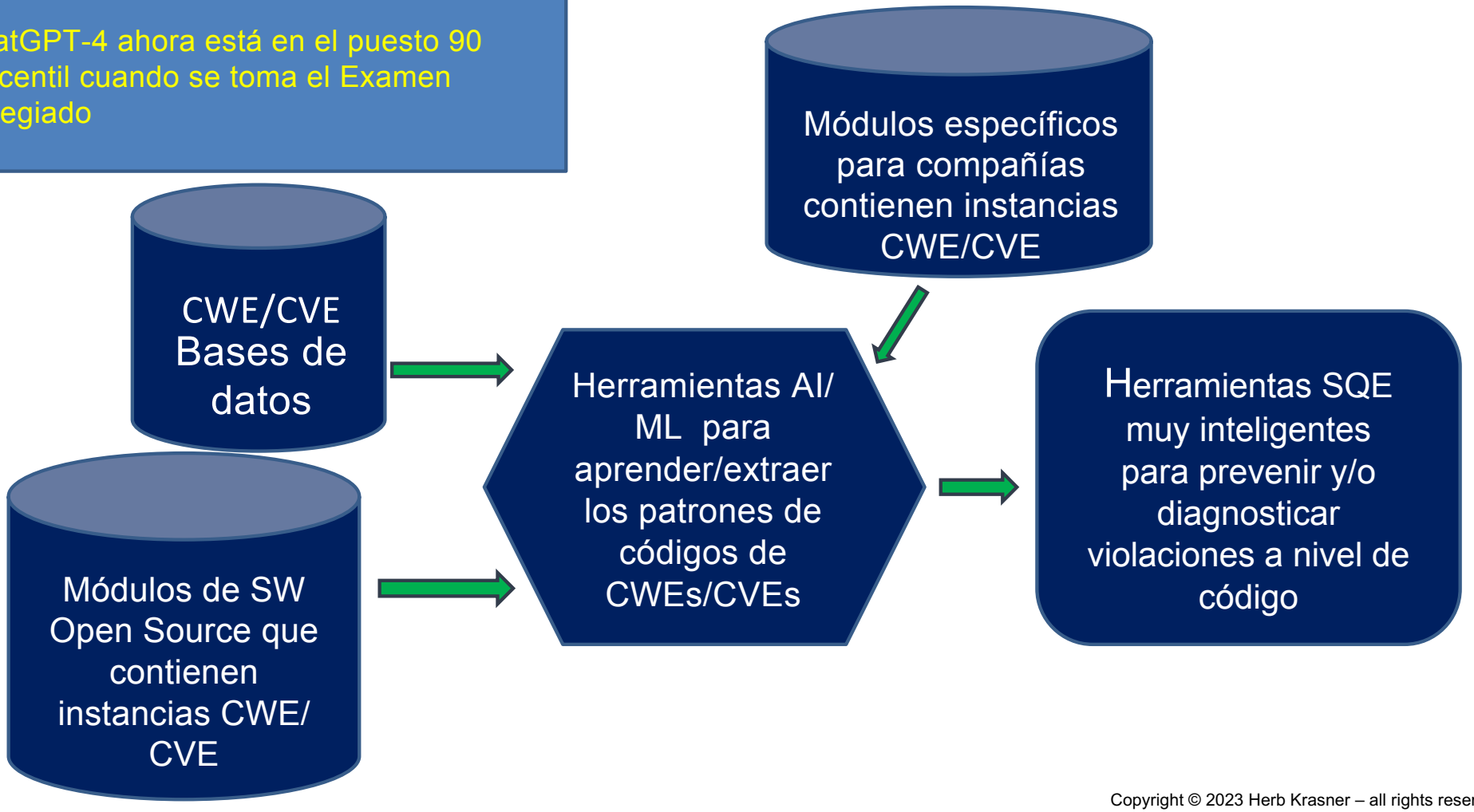
Recomendaciones de Desarrollo 2023

Recomendaciones específicas para desarrolladores de software y organizaciones de TI:

- Utilice los estándares de calidad de software, métricas relacionadas y herramientas emergentes.
- Analice y evalúe la calidad de todos los componentes de terceros / SW Open Source a ser incluidos en un sistema. Monitoree de cerca en operación. Aplique los parches oportunamente
- Evite los modelos DevOps y CI/CD que no incluye las mejores prácticas y herramientas de ingeniería de calidad continua
- Incluya la remediación de DT continua en su SDLC
- Invierta en el profesionalismo, conocimiento y herramientas de sus ingenieros de software
  - Considere la posibilidad de que sus desarrolladores se certifiquen en ISO/IEC 5055 para el conocimiento del código crítico y debilidades de arquitectura

Tecnología SQE Emergente

ChatGPT-4 ahora está en el puesto 90 percentil cuando se toma el Examen Colegiado



Predicciones para 2023-24

CPSQ + DT Tendencias

- CPSQ continuará erosionando el PIB real de EE.UU. (ya es >10%)
- La DT continuará limitando la productividad general de SW
- Un agudo faltante de profesionales de TI, exacerbará lo anterior.
  - ❖ 3.4 millones de posiciones vacantes
- Los problemas de cibercrimen, deficiencias de SW Open Source y DT continuarán creciendo.

Problemas continuos con:

- Sistemas críticos de infraestructura
- Sistemas de salud
- Media tech no social
- Cadenas de abastecimiento software
- Ataques globales de ciberguerra

La tendencia podría allanarse si las organizaciones adoptan las recomendaciones que se han dado en nuestra serie de reportes

Mi lista de deseos:

- **DevQualOps** superará el uso de DevOps y los métodos de integración continua/desarrollo continuo
- Las herramientas de inteligencia que usan SWE dentro de las organizaciones llegará al mercado público
- Los estándares SQE evolucionarán para poder atacar los 4 subfactores de la calidad de SW
- Los MITRE CWEs/CVEs serán reconocidos ampliamiente y usados en todos los proyectos de desarrollo IT
- La AI generativa no será usada erróneamente para la generación de código

El último chart

Contácteme con preguntas de seguimiento o comentarios

- hkrasner@utexas.edu

Gracias!

Tiempo para preguntas y respuestas.