Microsoft presentó un nuevo sistema impulsado por inteligencia artificial (IA) multimodelo llamado MDASH para facilitar la detección y corrección de vulnerabilidades a gran escala. La compañía indicó además que actualmente está siendo evaluado por algunos clientes dentro de una vista previa privada limitada.

MDASH, acrónimo de *multi-model agentic scanning harness*, fue concebido como un sistema independiente del modelo utilizado, capaz de emplear agentes de IA especializados para distintas clases de vulnerabilidades, con el objetivo de descubrir, validar y demostrar fallos explotables de manera autónoma en bases de código complejas como Windows.

«A diferencia de los enfoques basados en un solo modelo, esta plataforma coordina más de 100 agentes de IA especializados mediante una combinación de modelos de frontera y modelos destilados para descubrir, debatir y demostrar vulnerabilidades explotables de principio a fin», explicó Taesoo Kim, vicepresidente de seguridad agéntica en Microsoft.

MDASH funciona como una *"tubería estructurada"* que recibe una base de código y genera hallazgos validados y comprobados mediante distintas etapas automatizadas.

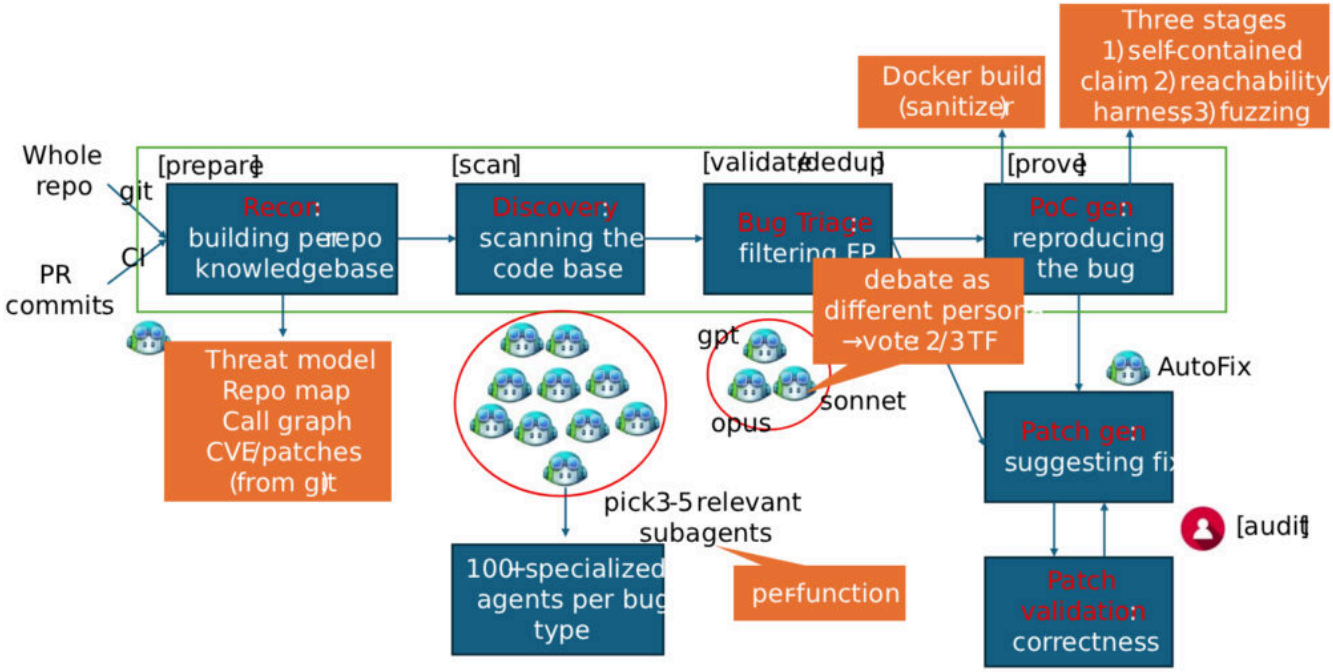
El proceso comienza con el análisis del código fuente para construir un modelo de amenazas y definir la superficie de ataque. Después, agentes especializados llamados *"auditores"* inspeccionan rutas de código potencialmente riesgosas para detectar problemas sospechosos. Posteriormente, otro conjunto de agentes, denominados *"debatidores"*, revisa y valida los hallazgos encontrados. Finalmente, el sistema agrupa resultados equivalentes y demuestra la existencia real de las vulnerabilidades.

La plataforma utiliza un panel configurable de modelos de IA: modelos de última generación (*state-of-the-art* o SOTA) para tareas de razonamiento complejo, modelos destilados para validaciones masivas y un segundo modelo SOTA independiente encargado de ofrecer una evaluación crítica adicional.

«El desacuerdo entre modelos también funciona como una señal: si un auditor identifica algo sospechoso y el debatidor no logra refutarlo, la credibilidad posterior de ese hallazgo

aumenta», señaló Microsoft. *«Un auditor no razona igual que un debatidor, y este tampoco opera igual que un demostrador. Cada etapa del proceso posee su propio rol, conjunto de instrucciones, herramientas y criterios de detención.»*

La empresa de Redmond indicó además que estos agentes especializados fueron entrenados tomando como referencia vulnerabilidades y exposiciones comunes (CVEs) detectadas previamente, junto con sus respectivos parches. También afirmó que la arquitectura permite adaptarse fácilmente a futuras generaciones de modelos.



MDASH ya ha sido probado en escenarios reales y logró identificar 16 de las vulnerabilidades corregidas en la edición de este mes de Patch Tuesday. Entre los fallos detectados se encuentran problemas relacionados con las capas de red y autenticación de Windows, incluidos dos errores críticos que podrían permitir ejecución remota de código:

- CVE-2026-33824 (CVSS: 9.8): una vulnerabilidad de tipo *double-free* en *"ikeext.dll"* que podría permitir a un atacante no autenticado enviar paquetes especialmente diseñados a un sistema Windows con Internet Key Exchange version 2 habilitado, provocando ejecución remota de código.
- CVE-2026-33827 (CVSS: 8.1): una condición de carrera en el componente TCP/IP de Windows (*"tcpip.sys"*) que permitiría a un atacante no autorizado enviar un paquete IPv6 manipulado a un equipo Windows con IPSec habilitado, derivando también en ejecución remota de código.

El anuncio de MDASH surge poco después de la presentación de Anthropic con su iniciativa Project Glasswing y de OpenAI con Daybreak, ambas enfocadas en acelerar mediante IA la detección, validación y mitigación de vulnerabilidades antes de que puedan ser explotadas por actores maliciosos.

«La implicación estratégica es evidente: el descubrimiento de vulnerabilidades mediante IA dejó de ser una curiosidad de investigación y ya opera como una defensa de nivel empresarial en producción. La ventaja sostenible no reside en un único modelo, sino en el sistema agéntico que lo rodea», concluyó Kim.