SMALL TALKS #15

Squad Agents

Cuando la IA deja de ser un asistente y empieza a trabajar en equipo



TESIS DEL EPISODIO

De usar IA como herramienta, a usarla como agente, a diseñar organizaciones digitales. Agent Squad no vende bots: propone un modelo operativo (roles, coordinación, handoffs, gobierno).

Invitado **Roberto Aguirre**

Guion de apoyo · respuestas mejoradas en las notas de cada slide

1

SMALL TALKS #15

¿Qué es un Squad de Agentes?

- Un equipo digital organizado: roles, entregables y una capa de coordinación, no bots en paralelo.
- El valor no está en cuántos agentes hay, sino en la calidad de los handoffs.
- Nova cumple la función de management: convierte intención en ejecución.
- Primero se modela el trabajo; los roles salen del proceso, no al revés.
- Un chatbot conversa; un squad sostiene un proceso completo.



> *"Un chatbot responde; un squad ejecuta.*

2

Del chat a una línea de producción

- El usuario declara el objetivo; el manager arma el plan y activa a los especialistas.
- Cada etapa entrega un artefacto con evidencia y criterios, no solo texto.
- Comunicación directa pero gobernada, con contratos claros de handoff.
- El contexto es infraestructura: conocimiento, estado y trazabilidad viajan con la tarea.
- Parte de la consistencia la da el software, no solo el modelo.

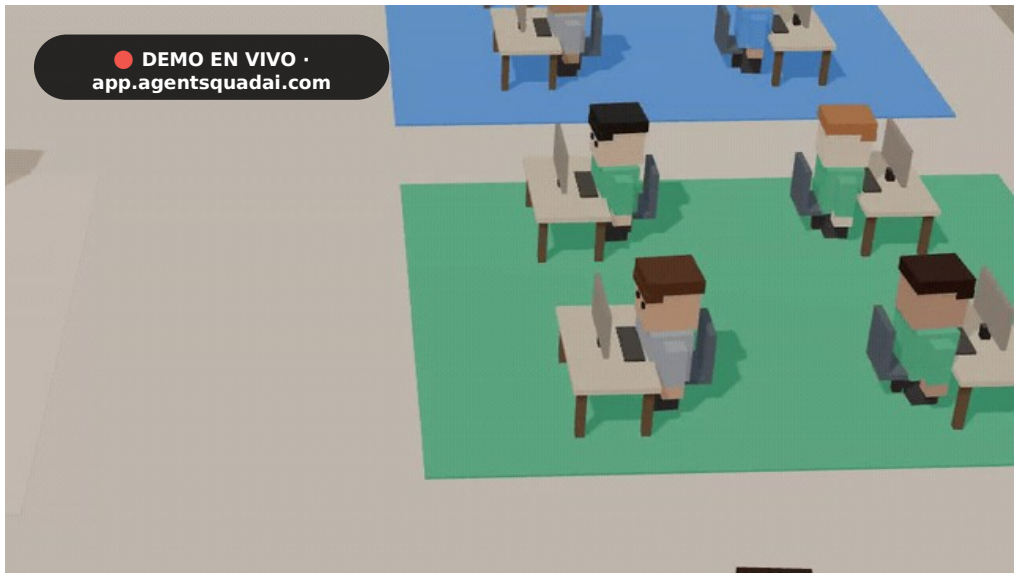


> *"El usuario gestiona el objetivo; Agent Squad gestiona la ejecución.*

3

El nuevo rol del humano: de ejecutor a supervisor

- Human in the loop no es aprobar cada paso: es autonomía graduada por riesgo.
- Orquestar (objetivos, reglas, límites) es distinto de aprobar (intervención puntual).
- Bajan búsqueda y coordinación; suben criterio, políticas y excepciones.
- El rechazo deja huella: aprendizaje operativo, no reentrenar el modelo.
- Aparece una disciplina nueva: gestionar capacidad digital.

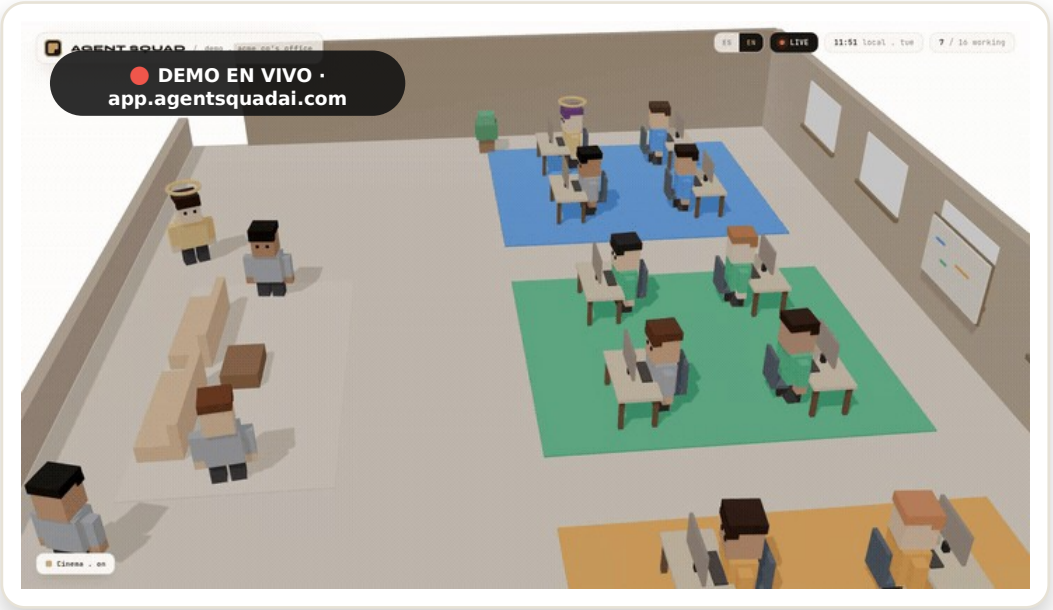


> *"La autonomía no se maximiza: se gobierna.*

4

¿Cómo se entrena un Squad? (conocimiento)

- No es "entrenarlo": es conectarlo a una arquitectura de conocimiento.
- 150 páginas de documentación no son 150 páginas de conocimiento ejecutable.
- Bases vectoriales y chunking son recuperación, no la solución completa.
- Se desacopla conocimiento, reglas y generación; después se valida.
- De "el modelo tiene info" a "la organización gobierna lo que el modelo usa".



> *"La organización gobierna lo que el modelo puede usar.*

5

Casos reales: bajar la idea a tierra

- Software Factory: se modela la cadena de valor, no se copian cargos humanos.
- Startup en Colombia: el cuello de botella era el contexto, no programar.
- YouTube: demanda, ángulo, guion, producción, QA y publicación conectadas.
- Salud: reducir fricción sin invadir el criterio clínico.
- Sirve a cualquier operación con procesos repetibles y handoffs.



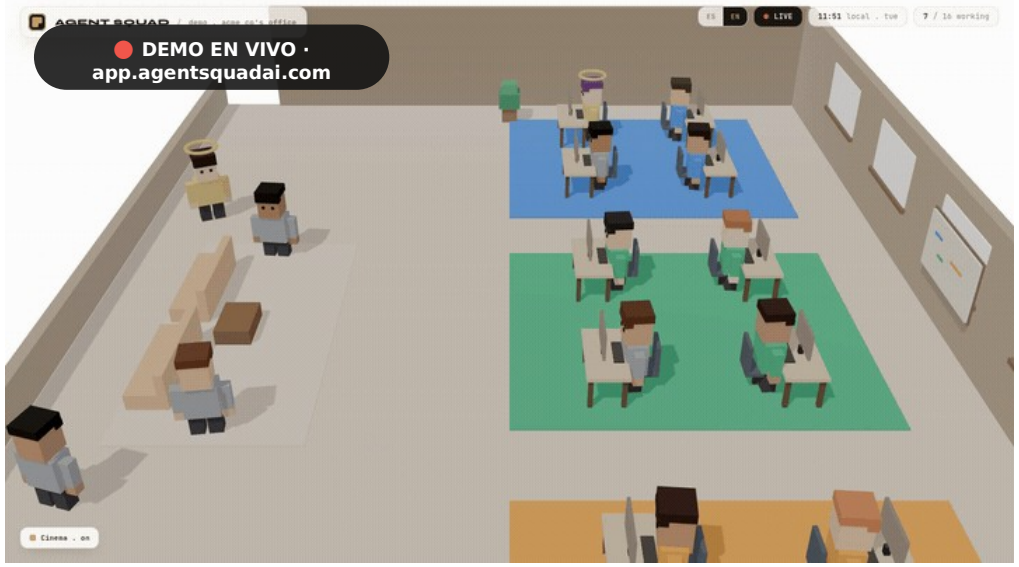
> *"La conversación se convierte en trabajo trazable.*

6

SMALL TALKS #15

Un modelo agnóstico a la industria

- Modelamos patrones de trabajo, no una industria específica.
- Estándar: orquestación, observabilidad, memoria, escalación, costos y gobierno.
- Específico: conocimiento, herramientas, reglas y métricas del negocio.
- Lo reutilizable es el sistema operativo; lo específico es el dominio.
- Una empresa podrá diseñar su propia organización digital de agentes.



> *"El runtime puede evolucionar. El producto es cómo organizamos el trabajo.*

7

SMALL TALKS #15

Costos, modelos y arquitectura

- Usar siempre el modelo más potente suele ser mala arquitectura económica.
- Matriz: complejidad, riesgo, latencia, costo y calidad. Y después medir.
- Modelos pequeños para clasificar, extraer, rutear y transformar.
- El modelo fuerte se reserva para la última milla creativa.
- Medir costo por resultado, no solo tokens totales.



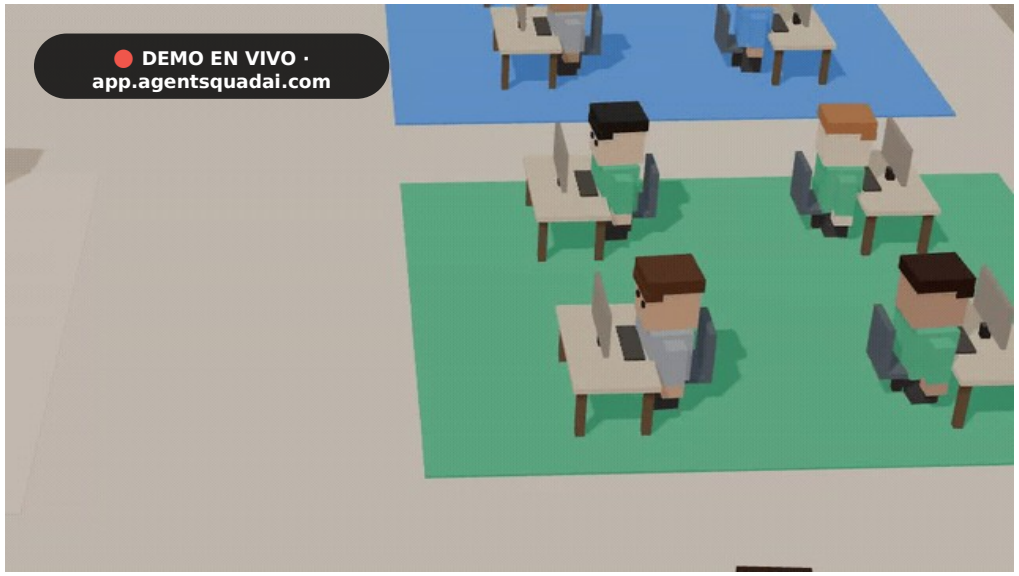
> *"No optimizamos tokens; optimizamos el costo de producir una buena decisión.*

7A

SMALL TALKS #15

Caso de uso: Salud en Colombia

- Problema: fragmentación e información dispersa; no reemplazar al médico.
- Funciones separadas con permisos; escala al profesional ante juicio clínico.
- Historia clínica: mínimo acceso necesario, cifrado y trazabilidad.
- Diagnóstico: apoyo informacional; la interpretación es humana.
- Onboarding más consistente con fuentes actualizadas.



> ***"En salud no buscamos autonomía máxima, sino autonomía apropiada al riesgo.***

¿Qué pasa con el trabajo humano?

- Pierden peso coordinación, búsqueda, preparación y seguimiento.
- Suben criterio, diseño de procesos y manejo de excepciones.
- Equipos pequeños con capacidad antes reservada a los grandes.
- Riesgo: supervisar demasiado; la supervisión también se diseña.
- La habilidad central: delegar bien a sistemas digitales.



> *"El próximo problema de la IA no es crear más agentes, sino aprender a gestionarlos.*

9

SMALL TALKS #15

¿Estamos preparados para una organización de agentes?

- Primero entender los procesos; la IA expone la operación real.
- El desafío es más organizacional que tecnológico.
- No automatizar del todo lo de alto impacto y baja reversibilidad.
- Autonomía = política explícita: acciones, datos, presupuesto, reversa.
- Control: observabilidad, límites de gasto, doble aprobación, poder detener.



> *"El progreso no es sacar personas del loop, sino reducir intervención sin perder control.*

10

Mirando hacia adelante

- Organigramas conviviendo con mapas de capacidades digitales.
- Hablaremos de roles y permisos, no de puestos calcados del humano.
- Una startup multiplica su capacidad, pero no reemplaza la estrategia.
- Ya es real: herramientas, conocimiento y workflows; falta autonomía robusta.
- Error mayor: confundir inteligencia con autonomía.



> *"No estamos automatizando tareas sueltas; estamos diseñando organización digital.*

FRASES ANCLA

Para repetir durante el episodio

- Un chatbot responde; un squad ejecuta.
- Agent Squad no es una colección de agentes: es un modelo operativo.
- El usuario gestiona el objetivo; Agent Squad gestiona la ejecución.
- No automatizamos tareas sueltas; diseñamos organización digital.
- El runtime puede evolucionar. El producto es cómo organizamos el trabajo.
- La autonomía no se maximiza: se gobierna.
- No optimizamos tokens; optimizamos el costo de una buena decisión.
- El próximo problema de la IA es aprender a gestionar agentes, no crear más.
- Pasamos años construyendo inteligencia. Ahora, la organización que la rodea.