Registro de Participación - Sprint 1

**Información General del Proyecto**

**Nombre del Proyecto:** Sistema de Web Scraping para análisis de competencia digital

**Cliente:** Restaurante "Rasca Cazuelas"

**Período del Sprint:** 11 de marzo - 8 de abril de 2025 (4 semanas)

**Modalidad de trabajo:** Mixta (presencial y virtual)

**Metodología:** Scrum

**Equipo de Proyecto**

|  |  |  |
| --- | --- | --- |
| **Nombre** | **Rol** | **ID** |
| Leonardo Valdes Palafox | Scrum Master / Desarrollador Principal | 18-011-0249 |
| Andrea Marlene Ortega Almendares | Diseñadora UI/UX / Desarrolladora Frontend | 17-003-1026 |
| C C | | |

**Product Owner:** Gerente de Marketing, Restaurante "Rasca Cazuelas"

**Supervisión Académica:** Profesora Alma Delia De Jesús Islao

**Curso:** Administración de Proyectos I (Ciclo escolar 2025-I)

***Índice***

***1.*** *Resumen Ejecutivo*

*2. Objetivos del Sprint*

***3.*** *Product Backlog Items (PBIs) Abordados*

***4.*** *Registros de Actividades del Sprint*

***4.1.*** *Eventos Scrum*

***4.1.1.*** *Sprint Planning*

***4.1.2****. Daily Scrums*

***4.1.3.*** *Sprint Review*

***4.1.4.*** *Sprint Retrospective*

***4.2.*** *Registro Detallado por Semana*

***5.*** *Entregables del Sprint*

***6.*** *Análisis de Riesgos y Mitigación*

***7.*** *Métricas y KPIs*

***7.1****. Métricas de Proceso*

***7.2****. Métricas de Producto*

***8****. Lecciones Aprendidas*

***8.1.*** *Aspectos Positivos*

***8.2.*** *Áreas de Mejora*

***9.*** *Plan para el Siguiente Sprint*

***9.1.*** *Objetivos Propuestos*

***9.2.*** *Mejoras de Proceso Propuestas*

1. **Resumen Ejecutivo**

El Sprint 1 se enfocó en el desarrollo de un sistema de Web Scraping para analizar la presencia digital de competidores del restaurante "Rasca Cazuelas" en la red social X (antes Twitter). Durante las cuatro semanas del sprint, el equipo logró desarrollar un sistema funcional que extrae y analiza tweets de los principales competidores, visualiza los datos en un dashboard interactivo y genera informes automatizados para el cliente.

# Logros Principales:

 Implementación de un sistema de web scraping con estrategias anti-detección

 Dashboard interactivo para visualización de datos

 Sistema de informes automatizados

 Análisis comparativo de competidores

# Métricas del Sprint:

 **Velocidad del equipo:** 28 puntos de historia

 **Historias de usuario completadas:** 10/12 (83.3%)

 **Tareas técnicas completadas:** 18/20 (90%)

 **Incidencias resueltas:** 8/8 (100%)

 **Impedimentos superados:** 4/4 (100%)

1. **Objetivos del Sprint**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Objetivo** | **Criterio de Éxito** | **Estado** |
| OBJ-  01 | Desarrollar un sistema de web scraping  funcional para X (Twitter) | Extracción exitosa de tweets de 3  competidores | ✅  Completado |
| OBJ-  02 | Crear un dashboard para visualización de  datos | Dashboard interactivo con al menos 5  métricas clave | ✅  Completado |
| OBJ-  03 | Implementar análisis comparativo de  competidores | Informes con métricas comparativas entre  los 3 competidores | ✅  Completado |
| OBJ-  04 | Automatizar la generación de informes | Sistema de generación automática de  informes semanales | ✅  Completado |
| C C | | | |

1. **Product Backlog Items (PBIs) Abordados**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Historia de Usuario / Tarea** | **Puntos** | **Estado** | **Responsable** |
| US-  01 | Como gerente de marketing, quiero extraer tweets de mis  competidores para analizar su presencia digital | 8 | ✅  Completado | Leonardo |
| US-  02 | Como gerente de marketing, quiero visualizar métricas de  engagement para entender la efectividad de la competencia | 5 | ✅  Completado | Andrea |
| US-  03 | Como gerente de marketing, quiero detectar contenido  multimedia para analizar estrategias visuales | 3 | ✅  Completado | Leonardo |
| US-  04 | Como gerente de marketing, quiero informes automatizados para  revisar periódicamente sin esfuerzo | 5 | ✅  Completado | Leonardo |
| US-  05 | Como gerente de marketing, quiero filtros temporales para  analizar tendencias a lo largo del tiempo | 3 | ✅  Completado | Andrea |
| US-  07 | Como gerente de marketing, quiero un dashboard responsivo  para consultar desde cualquier dispositivo | 3 | ✅  Completado | Andrea |
| TS-  01 | Implementar estrategias anti-detección para el scraper | 2 | ✅  Completado | Leonardo |
| C C | | | | |

1. **Registros de Actividades del Sprint**

# Eventos Scrum

## Sprint Planning

**Fecha:** 11 de marzo, 2025

**Participantes:** Leonardo, Andrea, Gerente de Marketing (Product Owner)

**Duración:** 2 horas

## Resultados:

 Sprint Backlog definido y priorizado  Definición de objetivos del sprint

 Estimación de PBIs

 Identificación de 3 competidores principales para análisis

## Decisiones clave:

 Enfoque exclusivo en la red social X (Twitter) para el Sprint 1  Priorización de extracción de datos y visualización básica

## Daily Scrums

El equipo realizó Daily Scrums tres veces por semana (martes, miércoles y jueves) de forma presencial, con una duración de 15 minutos cada uno. Se mantuvieron comunicaciones adicionales vía WhatsApp los días sin reuniones presenciales.

## Sprint Review

**Fecha:** 8 de abril, 2025

**Participantes:** Leonardo, Andrea, Gerente de Marketing y Gerente General de "Rasca Cazuelas"

**Duración:** 1 hora

## Resultados:

 Demostración del sistema completo

 Presentación del análisis de los tres competidores principales

 Explicación de los informes automatizados

## Feedback del cliente:

 Alta satisfacción con los resultados obtenidos

 Valoración positiva de los informes comparativos

 Interés en expandir el sistema a otras redes sociales en sprints futuros

## Sprint Retrospective

**Fecha:** 8 de abril, 2025

**Participantes:** Leonardo, Andrea

**Duración:** 45 minutos

## Lo que funcionó bien:

 Comunicación frecuente y efectiva entre el equipo

 Adaptación rápida a los cambios solicitados por el cliente

 División clara de responsabilidades

## Áreas de mejora:

 Estimación más precisa para tareas complejas

 Documentación más sistemática durante el desarrollo

 Automatización de pruebas unitarias

## Acciones para el próximo sprint:

 Implementar una herramienta de seguimiento del tiempo para mejorar estimaciones

 Establecer estándares de documentación y revisarlos semanalmente

 Reservar tiempo específico para desarrollo de pruebas automatizadas

* 1. **Registro Detallado por Semana Semana 1 (11 - 15 de marzo, 2025) Daily Scrum (12 de marzo)**

*Progreso:*

 Leonardo: Estructura inicial del proyecto y configuración de entorno ✅

 Andrea: Primeros mockups del dashboard ✅

*Impedimentos identificados:*

 IMP-001: Restricciones significativas de la API oficial de Twitter

 **Resolución:** Decisión de proceder con web scraping directo

*Plan del día:*

 Leonardo: Desarrollo de la clase base para el scraper con Selenium

 Andrea: Continuar el diseño de la interfaz y definir métricas a visualizar

## Daily Scrum (13 de marzo)

*Progreso:*

 Leonardo: Implementación del primer prototipo de la clase TwitterScraper ✅

 Andrea: Finalización del diseño de la primera versión del dashboard ✅

*Incidencias identificadas:*

 INC-003: Mecanismos anti-bot en Twitter/X que podrían bloquear el scraper

 **Acción:** Leonardo propuso implementar estrategias anti-detección

*Plan del día:*

 Leonardo: Implementación de métodos básicos de extracción  Andrea: Implementación del dashboard con datos de ejemplo

## Comunicación Virtual (14 de marzo - WhatsApp)

 Leonardo compartió avances del scraper básico

 Andrea solicitó ejemplos de datos para ajustar visualización

## Semana 2 (18 - 22 de marzo, 2025)

**Daily Scrum + Demo al Cliente (18 de marzo)**

*Progreso y Demostración:*

 Presentación del primer prototipo funcional al cliente

 Leonardo: Implementación de métodos básicos de extracción y navegación ✅

 Andrea: Avance en la implementación del dashboard con datos simulados ✅

*Feedback del cliente:*

 Solicitud de incluir detección de contenido multimedia

 Necesidad de filtrar tweets promocionados

*Incidencias identificadas:*

 INC-004: Solicitud de detección de contenido multimedia

 INC-005: Tweets promocionados aparecen en los resultados

*Plan del día:*

 Leonardo: Implementación del método has\_media y filtros para tweets promocionados

 Andrea: Integración de datos reales en el dashboard

## Daily Scrum (19 de marzo)

*Progreso:*

 Leonardo: Método has\_media completado y filtros para tweets promocionados implementados ✅

 Andrea: Dashboard con datos reales preliminares ✅

*Impedimentos identificados:*

 IMP-002: Ralentización del scraper después de extraer ~30 tweets

 **Acción:** Leonardo investigará optimizaciones

*Plan del día:*

 Leonardo: Optimización del proceso de scrolling y exportación a CSV

 Andrea: Adición de gráficos comparativos entre restaurantes

## Daily Scrum (20 de marzo)

*Progreso:*

 Leonardo: Optimización del proceso de scrolling y exportación a CSV completados ✅

 Andrea: Gráficos de barras comparativos para visualizar engagement implementados ✅

*Incidencias identificadas:*

 INC-006: Inconsistencias en la extracción de fechas en diferentes formatos

 **Acción:** Leonardo propuso solución para normalizar formatos

*Plan del día:*

 Leonardo: Implementar mecanismos de reintento para manejar fallos temporales

 Andrea: Implementar filtros temporales en el dashboard

## Comunicación Virtual (21 de marzo - WhatsApp)

 Leonardo compartió una actualización del scraper con mejor manejo de errores

 Andrea mostró capturas del dashboard con filtros temporales implementados

## Semana 3 (25 - 29 de marzo, 2025)

**Daily Scrum + Demo al Cliente (25 de marzo)**

*Demostración al cliente:*

 Presentación de versión funcional del scraper que extraía datos de tres cuentas

 Mostración del dashboard preliminar con datos reales de la última semana

*Feedback del cliente:*

 Alta satisfacción con la velocidad de desarrollo

 Solicitud de análisis comparativo más detallado de tipos de contenido

 Petición de informes semanales automatizados

*Progreso:*

 Leonardo: Finalización de mecanismos de reintento y manejo de errores ✅

 Andrea: Implementación de filtros temporales en el dashboard ✅

*Incidencias identificadas:*

 INC-007: Solicitud de informes semanales automatizados

 INC-008: Necesidad de análisis más granular de tipos de contenido

*Plan del día:*

 Leonardo: Iniciar implementación de generación de informes automáticos y clasificación de contenido

 Andrea: Ajustar dashboard según feedback del cliente

## Daily Scrum (26 de marzo)

*Progreso:*

 Leonardo: Inicio del desarrollo del sistema de generación de informes automáticos y clasificación preliminar de tipos de contenido ✅

 Andrea: Ajustes al dashboard según feedback del cliente ✅

*Impedimentos identificados:*

 IMP-003: Cambio en la estructura DOM de Twitter/X

 **Acción:** Leonardo actualizó selectores CSS afectados

*Plan del día:*

 Leonardo: Finalizar primera versión del sistema de informes y refinar clasificación de contenido  Andrea: Implementar visualización de tipos de contenido en el dashboard

## Daily Scrum (27 de marzo)

*Progreso:*

 Leonardo: Primera versión del sistema de informes automáticos y clasificación refinada de contenido

✅

 Andrea: Visualización de tipos de contenido en el dashboard implementada ✅

*Incidencias identificadas:*

 INC-009: Problemas con el procesamiento de tweets con emojis

 **Acción:** Leonardo se encargará de solucionarlo

*Plan del día:*

 Leonardo: Corregir el procesamiento de tweets con emojis e iniciar documentación técnica  Andrea: Optimizar la interfaz para dispositivos móviles

## Comunicación Virtual (28 de marzo - WhatsApp)

 Leonardo informó que el problema con los emojis estaba resuelto  Andrea compartió capturas del dashboard en versión móvil

## Semana 4 (1 - 8 de abril, 2025) Daily Scrum (1 de abril)

*Progreso:*

 Leonardo: Implementación de estrategias avanzadas anti-detección y avance en documentación técnica ✅

 Andrea: Finalización de la optimización para dispositivos móviles ✅

*Impedimentos identificados:*

 IMP-004: Degradación del rendimiento en cuentas con muchos tweets

 **Acción:** Leonardo propuso optimizaciones adicionales

*Plan del día:*

 Leonardo: Implementar optimizaciones de rendimiento para grandes volúmenes de datos  Andrea: Desarrollar características de exportación de informes desde el dashboard

## Daily Scrum (2 de abril)

*Progreso:*

 Leonardo: Implementación de optimizaciones con mejora del rendimiento en un 40% y adición de sección sobre legalidad a la documentación ✅

 Andrea: Finalización de funciones de exportación de informes ✅

*Incidencias identificadas:*

 INC-010: Necesidad de instrucciones detalladas de instalación

 **Acción:** Leonardo se encargará de incluirlas en la documentación

*Plan del día:*

 Leonardo: Realizar pruebas exhaustivas del sistema y añadir instrucciones detalladas de instalación

 Andrea: Realizar pruebas de usabilidad del dashboard

## Daily Scrum (3 de abril)

*Progreso:*

 Leonardo: Finalización de pruebas del sistema e instrucciones de instalación y uso ✅

 Andrea: Finalización de pruebas de usabilidad y ajustes finales ✅

*Plan para cierre del sprint:*

 Leonardo y Andrea: Preparar la demostración final para el cliente

 Leonardo: Finalizar ajustes de última hora y revisar documentación completa  Andrea: Preparar ejemplos de casos de uso para la demostración

## Comunicación Virtual (4 de abril - WhatsApp)

 El equipo coordinó los detalles finales para la demostración  Leonardo compartió la versión final para pruebas

 Andrea preparó un guion para la demostración con casos de uso reales

1. **Entregables del Sprint**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Entregable** | **Estado** | **Responsable** |
| DEL-01 | Sistema de Web Scraping para X (Twitter) | ✅ Completado | Leonardo |
| DEL-02 | Dashboard de visualización | ✅ Completado | Andrea |
| DEL-03 | Sistema de informes automatizados | ✅ Completado | Leonardo |
| DEL-04 | Documentación técnica y manual de usuario | ✅ Completado | Leonardo |
| DEL-05 | Presentación de resultados para el cliente | ✅ Completado | Leonardo y Andrea |
| C C | | | |

1. **Análisis de Riesgos y Mitigación**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Riesgo** | **Probabilidad** | **Impacto** | **Estrategia de mitigación** | **Responsable** |
| RISK-  01 | Bloqueo del scraper por  mecanismos anti-bot | Alta | Alto | Implementación de estrategias  anti-detección y rotación de IPs | Leonardo |
| RISK-  02 | Cambios en la estructura  DOM de X/Twitter | Media | Alto | Monitoreo regular y actualización  rápida de selectores | Leonardo |
| RISK- 03 | Restricciones legales del web scraping | Baja | Alto | Documentación de  consideraciones legales y uso ético | Leonardo |
| RISK-  04 | Volumen de datos excesivo  que ralentiza el sistema | Media | Medio | Optimización de carga y  procesamiento por lotes | Leonardo |
| C C | | | | | |

1. **Métricas y KPIs**
   1. **Métricas de Proceso**

|  |  |  |  |
| --- | --- | --- | --- |
| **Métrica** | **Valor** | **Objetivo** | **Varianza** |
| Velocidad del equipo | 28 puntos | 30 puntos | -6.7% |
| Porcentaje de completitud del Sprint | 91.6% | 90% | +1.6% |
| Cambios en el alcance durante el Sprint | 2 | ≤3 | Dentro del objetivo |
| Tiempo promedio de resolución de impedimentos | 1.2 días | 2 días | +40% (mejor) |
| C C | | | |

* 1. **Métricas de Producto**

|  |  |  |  |
| --- | --- | --- | --- |
| **Métrica** | **Valor** | **Objetivo** | **Varianza** |
| Tiempo de carga del dashboard | 2.3 segundos | <3 segundos | +23.3% (mejor) |
| Precisión en extracción de datos | 97.8% | >95% | +2.8% |
| Satisfacción del cliente | 4.8/5 | >4/5 | +20% |
| Cobertura de funcionalidades solicitadas | 90% | >85% | +5% |
| C C | | | |

1. **Lecciones Aprendidas**

# Aspectos Positivos

 La comunicación directa y frecuente con el cliente permitió ajustar el producto rápidamente a sus necesidades

 La división clara de responsabilidades mejoró la eficiencia del equipo

 Las demostraciones tempranas generaron confianza y validaron la dirección del desarrollo

# Áreas de Mejora

 La estimación de tareas complejas debe considerar un margen de error mayor

 El proceso de documentación debe ser más sistemático y continuo

 Las pruebas automatizadas deben implementarse desde el inicio del proyecto

1. **Plan para el Siguientes Sprints**

# Objetivos Propuestos

 Extender el análisis a otras redes sociales (Instagram y Facebook)

 Implementar análisis de sentimiento en los comentarios

 Mejorar el sistema de reportes con visualizaciones avanzadas

# Mejoras de Proceso Propuestas

 Implementación de una herramienta de seguimiento del tiempo (Toggl)

 Establecimiento de estándares de documentación con revisiones semanales

 Reserva de tiempo específico para desarrollo de pruebas automatizadas