

Towards Sentiment Analysis in Agile Development Environments

María Guadalupe Gramajo, Luciana Ballejos and Mariel Ale

CIDISI, Centro de Investigación y Desarrollo en Ingeniería en Sistemas de Información
CONICET - UTN

Lavaisse 610, Santa Fe, Argentina
{mgramajo, lballejo, male}@frsf.utn.edu.ar

Abstract. In agile software development projects, user stories are written in natural language in order to describe the functionalities and features requested by interested parties. Given the ambiguous nature of natural language, user stories may reflect feelings intrinsically expressed by the product owner during requirements education tasks. While their visualization and identification are complex, word processing and sentiment analysis techniques can be applied to extract useful and valuable information for the development team. For that reason, this paper proposes a methodology to analyze the affective dimension of user stories, considering the feelings and emotions that affect the communication process among team members and identify how these factors impact the software development process, specifically in activities related to the definition and specification of requirements.

1 Introducción

Las últimas décadas han sido testigo de una tendencia constante e irreversible hacia la globalización de los negocios y, en particular, de las empresas dedicadas al desarrollo de productos de software [1]. La influencia de este fenómeno ha generado escenarios de desarrollo global de software, donde los stakeholders se encuentran geográficamente distribuidos.

Gestionar proyectos de software con estas características permite aprovechar la escasa disponibilidad de recursos humanos altamente calificados que se encuentran distribuidos de manera global, y aumentar la productividad en función de las zonas horarias de los miembros del equipo de desarrollo [2]. Sin embargo, estos nuevos contextos de trabajo impactan de manera directa en las fases del ciclo de vida del desarrollo de software, donde las exigencias de comunicación y colaboración son mayores entre las partes interesadas, especialmente en las fases de educación y especificación de requisitos.

En relación a esta nueva tendencia de desarrollo, las prácticas ágiles han demostrado tener éxito en diversos proyectos de entornos distribuidos [3, 4], dado que abogan por una planificación adaptativa y desarrollo evolutivo. En los últimos años, las prácticas

ágiles han ganado una atención significativa, debido a su enfoque flexible para administrar la volatilidad de los requisitos y su énfasis en la amplia colaboración y comunicación entre clientes y desarrolladores, beneficiándose de las tecnologías y herramientas disponibles en la web [5].

Las prácticas ágiles representan un conjunto de procesos de desarrollo en los que, tanto los requisitos como el producto entregado, evolucionan de forma incremental a través de una serie de iteraciones cortas. En este contexto, los requisitos son especificados en historias de usuario. Las historias de usuario son descripciones breves que capturan elementos esenciales de un requisito y responden a la tripleta (As a/an, I want to, so that) para expresar el rol del usuario, la funcionalidad o comportamiento que se espera del sistema, y una breve descripción que justifica la funcionalidad deseada.

Las actividades de especificación de historias de usuario en los entornos ágiles comienzan con aproximaciones simples y cognoscibles, que luego a través de procesos de retroalimentación se refinan hasta derivar en su versión final [6]. Sin embargo, a menudo estas especificaciones presentan expresiones con carga emocional, una característica propia del lenguaje natural utilizado en la redacción de las historias de usuario. Esta situación es de esperar, debido a que el desarrollo de un producto de software requiere de sinergia y esfuerzo humano, lo cual implica interacción entre los miembros del equipo para resolver problemas complejos, estando sujetos a procesos afectivos y cognitivos [7].

Las expresiones con contenido emocional pueden generar malentendidos, insatisfacción y esfuerzo adicional, por lo que es necesario tomar acciones para su refinamiento, a los fines de detectar errores y corregirlos. Considerando esta situación y, con la intención de contribuir en las actividades que abordan las tareas de educación y refinamiento en los entornos ágiles, este estudio propone un enfoque para detectar y analizar los sentimientos intrínsecos expresados en las historias de usuario.

El contenido de este trabajo se encuentra organizado de la siguiente manera. La Sección 2 describe brevemente los conceptos preliminares de las técnicas de análisis de sentimientos y algunos antecedentes que reflejan la aplicación de estas técnicas en proyectos de desarrollo de software, en virtud de establecer un marco teórico introductorio. La Sección 3 enfatiza la necesidad de identificar y analizar sentimientos inmersos en las historias de usuario. La Sección 4 detalla la metodología aplicada y el flujo de trabajo de la herramienta propuesta. La Sección 5 describe una prueba de concepto aplicada sobre historias de usuario, a los fines de examinar la utilidad de la herramienta de análisis. La Sección 6 sintetiza las conclusiones y futuras líneas de investigación.

2 Marco Introductorio

En esta sección se describirán brevemente los principales conceptos de las técnicas de análisis de sentimientos. Posteriormente, se mencionarán algunos trabajos relevantes que indagan sobre su aplicación en la Ingeniería de Software, a los fines de identificar antecedentes en el área.

En los últimos años, el análisis de sentimientos se ha convertido en una de las áreas de investigación de mayor crecimiento [8]. El auge de las técnicas de análisis de sentimientos comenzó a mediados del año 2000, y se centró principalmente en el análisis de las opiniones de productos disponibles en la Web. Desde entonces, el uso del análisis de sentimientos ha alcanzado muchas otras áreas, entre ellas, la Ingeniería de Software. La aplicación de estas técnicas sobre expresiones en lenguaje natural permite analizar y aprovechar la riqueza semántica presente en redes sociales, blogs, foros y recursos textuales para generar información de valor.

El análisis de sentimientos (AS), es una técnica de minería de opinión utilizada para determinar la polaridad de las emociones identificadas en recursos textuales. En términos generales, el AS clasifica el contenido de un texto como positivo, neutral o negativo en función de indicadores contextuales, es decir, su propósito es identificar el sentimiento general del autor en un texto de opinión [7].

Una de las propuestas más citadas en la literatura con respecto a la aplicación del AS y el uso de técnicas de minería de datos es el estudio presentado por Hu y Liu [9], donde en el contexto de venta de productos online (e-commerce), proponen la extracción de las características de los productos vendidos que han sido evaluadas por los clientes expresando su opinión. Luego, aplicando estas técnicas, clasifican los comentarios en positivos o negativos. Su objetivo es proporcionar un reporte útil para proveedores y vendedores con respecto a la opinión de los clientes sobre un determinado producto.

Posteriormente, las técnicas de AS comenzaron a aplicarse en el contexto de los proyectos de software. Algunos estudios recientes revelan el uso de las técnicas de AS para detectar y analizar cómo se sienten las personas involucradas en los procesos de desarrollo de software. Jurado y Rodríguez [10] señalan la existencia de factores que pueden afectar la productividad, la calidad de las tareas y la satisfacción laboral, siendo uno de estos factores, las emociones. Es por esta razón que proponen el AS para monitorear las características emocionales en el proceso de desarrollo de software. Para llevar a cabo esta propuesta, recopilaron problemas de algunos de los proyectos alojados en GitHub, considerando diferentes entornos de desarrollo y lenguajes de programación. Posteriormente, realizaron el AS sobre cada issue, a través de un enfoque léxico. Este enfoque proporciona coincidencias entre una palabra específica y su sentimiento relacionado. Los resultados mostraron que los desarrolladores dejan sentimientos subyacentes en el texto, y esa información podría usarse para analizar el proceso de desarrollo, con lo cual los autores en esta propuesta evidencian la viabilidad de la aplicación de las técnicas de AS en el ciclo de vida del desarrollo del software.

Patwardhan [11], en su trabajo propone implementar la identificación de los sentimientos y emociones presentes en los equipos de desarrollo geográficamente distribuidos, durante el ciclo de lanzamiento del software. Para ello, analiza múltiples fuentes de información tales como, correos electrónicos, discusiones en foros web, revisiones de códigos, opiniones, encuestas y mensajes relacionados con el producto de software en desarrollo. El autor también evaluó la existencia de una posible relación entre la polaridad de los sentimientos expresados por los equipos y los factores técnicos y sociales a los que están expuestos, entre ellos, la ubicación geográfica, el género, el nivel de experiencia, la duración de la liberación del producto de software, el tamaño del equipo, recuento de problemas del código y los comentarios de revisión del código.

Esta investigación, además, propone el desarrollo de un panel de control de emociones utilizando las técnicas de AS para medir el éxito del proyecto y la salud emocional del equipo de desarrollo.

Los resultados demostraron la existencia de información semántica valiosa en lo que puede parecer una charla corporativa o una sobrecarga de correo electrónico. Este trabajo concluye enfatizando que el proceso de identificación de sentimientos puede ser adoptado como una herramienta valiosa para que los gerentes evalúen el progreso del proyecto y analicen cómo la dinámica del equipo afecta el sentimiento general respecto a la entrega del producto de software.

Wrobel [12], propone en su investigación revelar cuáles son los estados emocionales presentes en el proceso de desarrollo de software, con qué frecuencia ocurren y cuál es su impacto en la productividad. Para ello, realiza una encuesta para la cual los datos fueron recopilados durante cuatro semanas, analizándose cincuenta y seis respuestas. Los resultados revelaron que, entre los estados emocionales negativos, la mayoría de los encuestados indicaron que con más frecuencia experimentaban frustración y enfado. Estos estados emocionales son identificados como los más riesgosos para la productividad, ya que afectan el desempeño de los programadores.

Marshall y otros [7] proponen el análisis del contenido emocional de las publicaciones realizadas por los estudiantes en la herramienta de gestión de proyectos denominada SEREBRO, una propuesta presentada por los autores en una investigación anterior [13]. El objetivo de la investigación es realizar un seguimiento de los esfuerzos de desarrollo y el comportamiento emotivo de los equipos de trabajo a medida que desarrollan un producto de software, como parte de sus proyectos finales de la materia Ingeniería de Software. El conjunto de datos consta de más de mil trece publicaciones en el foro, producidas por cinco equipos en tres sprints. Los equipos de estudiantes siguen un proceso de desarrollo de software ágil, utilizando la metodología Scrum con sprints adaptados a un entorno educativo. Las técnicas de AS se aplicaron para clasificar cada publicación como positiva, negativa, conflictiva o sin emociones. Cada publicación en el foro ha sido evaluada manualmente, utilizando técnicas de AS, incluida la localización de palabras clave, el análisis léxico y los enfoques basados en conceptos. El método de análisis propuesto en esta investigación consiste en buscar correlaciones significativas entre las métricas de actividad emocional y otras medidas de rendimiento obtenidas durante el desarrollo del producto de software, para lo cual los autores utilizan el coeficiente de correlación de Pearson.

Por otro lado, Islam y Zibran [14], proponen un estudio comparativo de cuatro diccionarios de sentimientos de propósito general, es decir, diccionarios independientes del dominio de aplicación. En particular, examinan si estos diccionarios, creados con diferentes métodos, muestran resultados diferentes en la detección de sentimientos en textos de Ingeniería de Software. Los diccionarios son evaluados sobre tres conjuntos de datos, correspondientes a comentarios de revisiones de código, desbordamiento de pila y comentarios de problemas extraídos del sistema de seguimiento JIRA. En su investigación destacan que los estudios que utilizan diccionarios de propósito general para detectar los sentimientos, obtienen resultados de baja precisión, debido a que éstos deben operar sobre un dominio técnico para lo cual no fueron inicialmente ideados. Considerando esta situación, los autores direccionaron sus esfuerzos e investigaciones

en determinar el diccionario que obtiene mayor precisión respecto a la detección de sentimientos sobre los conjuntos de datos analizados, e identificar en qué medida los diferentes diccionarios coinciden al detectar la polaridad de los sentimientos. Los resultados de esta investigación reflejan que el diccionario que presenta un mayor rendimiento respecto a la detección de sentimientos sobre los conjuntos de datos analizados es Sentistrength.

A pesar de que existen diversas propuestas en la literatura que indagan sobre la aplicación de las técnicas de AS en el área de la Ingeniería de Software, no se hallaron investigaciones que hagan uso de ellas sobre historias de usuario, artefacto utilizado en las prácticas ágiles para capturar los requisitos. Por otro lado, es necesario destacar que si bien en su investigación Marshall y otros [7], proponen la inserción del análisis afectivo en el contexto de desarrollo de software ágil, el mismo sólo se reduce a un análisis en el entorno educativo, a diferencia de la herramienta de análisis de sentimientos propuesta en esta investigación, ideada para brindar soporte a las tareas de educación y especificación en proyectos de desarrollo empresarial.

3 Sentimientos en las Prácticas Ágiles

En esta sección se realiza una breve descripción del escenario que motiva la identificación y análisis de los sentimientos inmersos en las historias de usuario.

Las prácticas ágiles surgen como una reacción de la industria de desarrollo de software frente a las metodologías clásicas, en la cuales los proyectos eran planificados en fases lineales y secuenciales de tipo cascada, gestionados con métodos como Gantt, PERT, entre otros [15]. En su investigación, Demmer y otros complementan este pensamiento sosteniendo que estas prácticas surgieron de la creencia que un mayor acercamiento con la realidad del desarrollo de productos basados en el aprendizaje, innovación y cambio daría mejores resultados [16]. Los principios ágiles ponen énfasis en construir software que funcione y que pueda ser usado rápidamente, y en reducir la extensa documentación relacionada a requisitos en las metodologías tradicionales [17]. Las prácticas ágiles se centran en equipos multifuncionales con capacidad para decidir por ellos mismos y apuntan a iteraciones rápidas, con el cliente dando su opinión continuamente.

En los entornos ágiles, las necesidades del dueño del producto (*product owner*, en inglés) se documentan en historias de usuario (*user stories*, en inglés). Éstas, al ser redactadas en lenguaje natural, pueden introducir incertidumbre y ambigüedad en sus especificaciones. Considerando que la ambigüedad es una característica intrínseca del lenguaje natural ésta puede conducir a la inserción de palabras y frases que expresan sentimientos. Este escenario convierte a las historias de usuario en un artefacto candidato para la aplicación de las técnicas de análisis de sentimientos.

Una especificación expresada en lenguaje natural puede tener dos dimensiones perceptibles: polaridad sentimental (*sentiment polarity*, en inglés) e intensidad sentimental (*sentimental intensity*, en inglés). La polaridad sentimental indica la positividad, negatividad o neutralidad de la expresión, mientras que la intensidad sentimental captura la

fuerza de la expresión emocional. Consecuentemente, la presencia de carga emocional conduce a especificaciones ambiguas e influye en las métricas asociadas a las historias de usuario utilizadas para la evaluación y aseguramiento de calidad [18]. La Figura 1 sintetiza las relaciones que se establecen entre las entidades involucradas en el proceso de análisis de sentimientos.

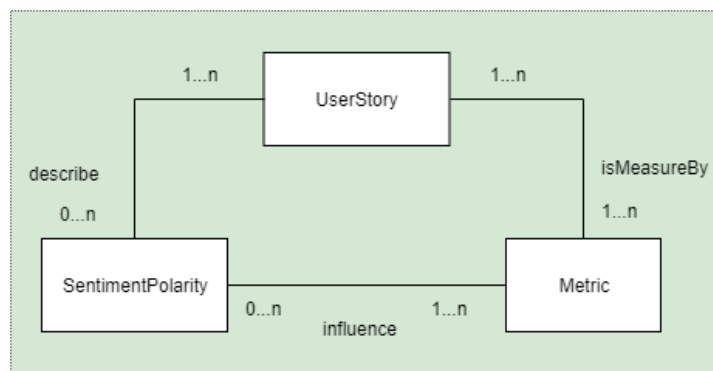


Fig. 1. Representación de las entidades relacionadas en el análisis de sentimientos.

Si bien la construcción de las historias de usuario dispone del formato léxico sintáctico representado en la Figura 2, a menudo, el uso de esta plantilla no garantiza alcanzar especificaciones completas, consistentes y correctas ya que, al igual que muchas otras buenas prácticas de desarrollo de software, presenta variaciones al ser aplicado en entornos reales de desarrollo, quedando meramente definidas en un marco teórico.

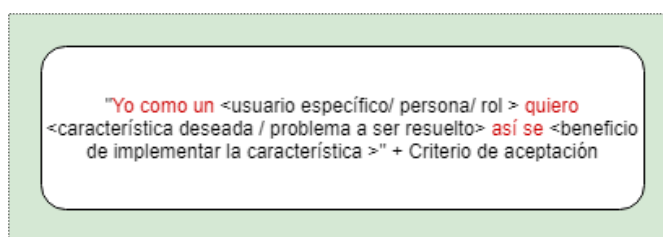


Fig.2. Formato léxico sintáctico de una historia de usuario.

Lucassen y otros [19] mencionan en su trabajo que, a pesar del uso de plantillas para educir los requisitos en los entornos ágiles, algunas secciones que conforman una historia de usuario pueden ser afectadas por la inserción de subjetividad y carga emocional de quien que se encuentre a cargo de su redacción. Esta situación ocurre especialmente en las secciones donde se expresan las características deseadas, el valor y los beneficios de implementar una determinada funcionalidad, debido a la variedad de construcciones

textuales posibles que brinda el lenguaje natural. La Figura 3 señala las secciones de las historias de usuario en las que podrían presentarse inserciones de sentimientos.

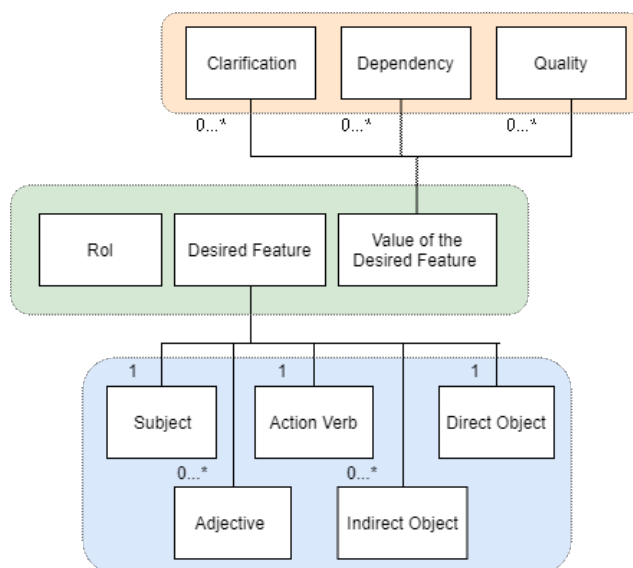


Fig. 3. Secciones de una historia de usuario con posibles inserciones de sentimientos [17].

La Figura 3 muestra las secciones de una historia de usuario en la que es posible la inserción de expresiones sentimentales. Una de ellas es la sección correspondiente a la etiqueta “*I want to*” de la tripleta mencionada en la Sección 1, donde se especifican las características y funcionalidades deseadas. Dicha sección debería estar compuesta de al menos tres elementos: sujeto, verbo y objeto directo. De manera adicional también puede contener un adjetivo y objeto indirecto, estando estos últimos sujetos a una o múltiples inserciones dentro de la redacción de la historia de usuario, dando lugar a más de una posible construcción textual.

Otra sección sujeta a inserciones de emociones y sentimientos es la correspondiente a la etiqueta “*So that*”, utilizada para justificar la implementación de la funcionalidad deseada. La descripción de por qué es necesario implementar una determinada funcionalidad, puede conducir a la incorporación de dependencias con otras funcionalidades, o la definición de ciertas propiedades de calidad deseadas para el requisito a implementar, o incluso generar una justificación desmedida de la necesidad del requisito. Por ejemplo: “Como usuario me encantaría poder consultar de forma segura mi cuenta bancaria para visualizar las transferencias realizadas de un modo fácil e intuitivo”. Lo cual implica la existencia de un registro de transferencias, y la necesidad de propiedades de calidad que se deberían considerar. Dada la variedad de las posibles construcciones textuales mencionadas, esta situación puede conducir a la incorporación de expresiones sentimentales durante las tareas de especificación.

Por otro lado, la presencia de carga emocional en estos artefactos también puede influir en la priorización de las historias de usuario del backlog. El product owner es el responsable de identificar las funcionalidades requeridas y ordenar éstas según su prioridad, también decide cuáles deberían ir al principio del backlog para el siguiente sprint, y asigna un orden para cada componente y su resolución a lo largo del ciclo de vida de desarrollo, en donde el backlog puede seguir siendo refinado continuamente [16]. Si las historias de usuario presentan carga emocional, entonces éstas pueden influir en el direccionamiento de los esfuerzos del equipo de trabajo.

Por todo lo expresado anteriormente, es de esperar que surjan dificultades durante el proceso de especificación en las historias de usuario. Por este motivo, este trabajo propone el desarrollo de una herramienta que brinde soporte a las actividades de especificación y refinamiento, aplicando técnicas de análisis de sentimientos.

4 Metodología

Esta sección describe la metodología aplicada a las historias de usuario, a los fines de abordar la detección de los sentimientos inmersos durante su especificación.

Lo que propone este trabajo es el desarrollo de una herramienta de análisis de sentimientos que actúe de soporte a la especificación y revisión de historias de usuario. La misma fue ideada con el objetivo de detectar las posibles inserciones de sentimientos y carga emocional al redactar las historias de usuario.

El proceso inicia con la entrevista al cliente con la finalidad de capturar necesidades y restricciones del sistema a desarrollar. A continuación, el miembro del equipo de trabajo que tiene asignada la tarea de especificar los requisitos genera las historias de usuario iniciales. Posteriormente, las historias de usuario se someten al analizador de sentimientos propuesto y se observan los resultados procesados por la herramienta, antes de proceder a la etapa de grooming y planificación de tareas. Si se detectaron anomalías en las historias de usuario, el responsable de esta tarea debe decidir si las historias deben ser refinadas y mejoradas en relación a su redacción, o si se continúa con la siguiente etapa a partir los resultados obtenidos.

Cabe destacar que la herramienta de análisis propuesta tiene por objeto detectar y visualizar los sentimientos inyectados durante la redacción de las historias de usuario. Sin embargo, mejorar su redacción y eliminar subjetividades es responsabilidad del equipo de trabajo. La Figura 4 muestra la metodología propuesta.

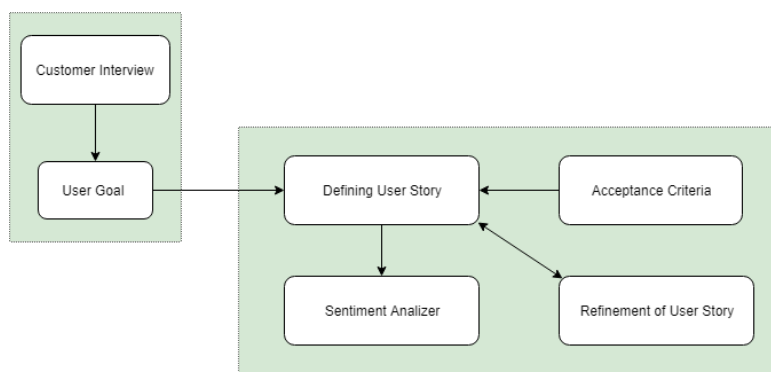


Fig. 4. Metodología propuesta.

En la siguiente subsección se detalla el flujo de trabajo de la herramienta de AS.

4.1 Flujo de Trabajo de la Herramienta de Análisis de Sentimientos

Este estudio aborda el análisis de las historias de usuario, a los fines de identificar las expresiones que presentan carga emocional. Para detectar la inserción de sentimientos, las historias de usuario son sometidas a un proceso de inspección a través de la herramienta de análisis propuesta. En caso de detectar expresiones sentimentales, la herramienta indica al usuario las secciones que presentan carga emocional y provee un mensaje comunicando la necesidad de revisión y refinamiento de la historia de usuario. El objetivo de esta propuesta es disminuir la subjetividad de la expresión y contribuir con el aseguramiento de la calidad. La Figura 5 sintetiza el procedimiento lógico del analizador de sentimientos propuesto en este trabajo.

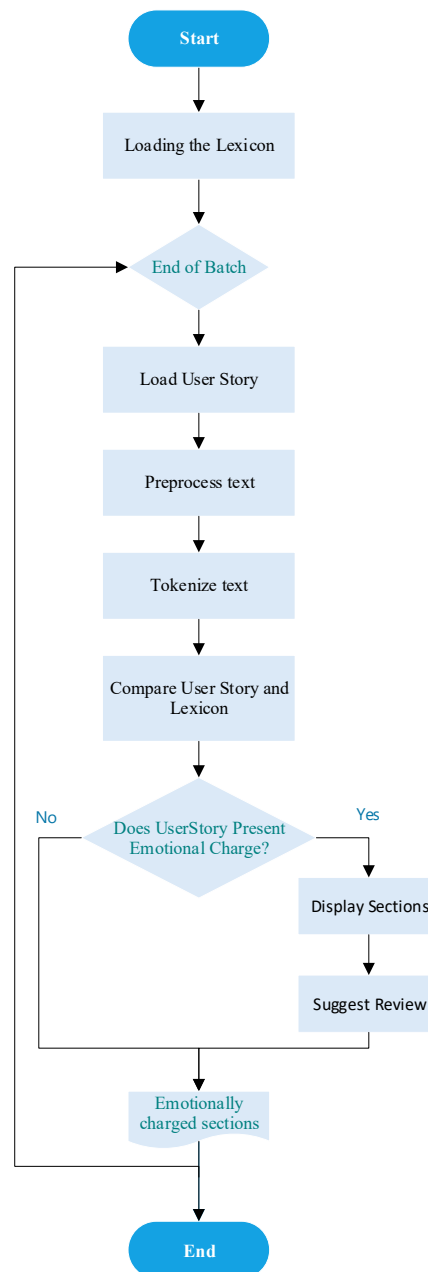


Fig. 5. Procedimiento lógico del analizador de sentimientos.

El flujo de trabajo aplicado para identificar los sentimientos mediante la herramienta analizadora consiste en: (i) cargar el léxico de sentimientos; (ii) cargar la historia

de usuario que se desea analizar; (iii) realizar el preprocesamiento del texto para eliminar cualquier signo o caracter no significativo (signos de puntuación, signos de exclamación e interrogación, etc.); (iv) tokenizar el texto; (v) comparar historia de usuario con el léxico de sentimientos; (vi) identificar los sentimientos inyectados; (vii) visualizar los resultados; (viii) sugerir revisión.

Al finalizar el proceso, los sentimientos resultan identificados y se muestran a través de una interfaz gráfica, para proporcionar una visualización de los resultados obtenidos a los miembros responsables de la actividad de educación y refinamiento de historias de usuario. Las historias de usuario que presenten carga emocional son retornadas indicando la sección que requiere revisión.

Para desarrollar la herramienta de análisis de sentimientos se codificó un script utilizando el lenguaje de programación R versión 3.4.4. Esta propuesta utiliza R Studio, un entorno de software libre ampliamente utilizado para realizar minería de texto y análisis de datos [20]. También se utilizaron dos librerías disponibles para el lenguaje R: (1) *Tidyverse*, que permite tokenizar el texto. Un token es una unidad significativa de texto, como una palabra u oración, en la que estamos interesados usar para su análisis [21]. La librería *Tidyverse* se utilizó para tokenizar las historias de usuario en palabras; por otro lado, (2) *Tidytext* permite importar el léxico de sentimientos [22]. La librería *Tidytext* contiene tres léxicos de sentimientos en su conjunto de datos denominado *sentiments*. Ellos son: el léxico AFINN [23] que asigna a las palabras un puntaje que varía entre -5 y 5, los puntajes negativos indican sentimientos negativos y los puntajes positivos indican sentimientos positivos, el léxico NRC [24] que clasifica las palabras de una manera binaria (“si” / “no”) en las categorías positivo, negativo, ira, anticipación, miedo alegría, tristeza, sorpresa y confianza y, por último, el léxico Bing [25] que clasifica las palabras en categorías positivas y negativas.

Cabe destacar que la librería *Tidytext* además, permite importar el léxico con el cual se desea trabajar mediante la función `get_sentiments()`, con lo que es posible obtener léxicos de sentimientos específicos. Para la herramienta de análisis de sentimientos propuesta se importó el léxico Bing, porque el mismo dispone sólo de dos categorías posibles para clasificar las palabras, a diferencia de los otros léxicos que proponen un rango de valoración, lo que implica la incorporación de tareas adicionales en el proceso de análisis. En esta propuesta el interés reside en analizar sólo la polaridad de los sentimientos, es por ello que resulta útil la clasificación de las palabras en las dos categorías (positivas y negativas) tal como ofrece el léxico Bing.

En la siguiente sección se detalla una prueba de concepto, a los fines de examinar la utilidad de la herramienta de análisis propuesta.

5 Prueba de Concepto

En esta sección se propone una prueba de concepto para analizar y evaluar los fines prácticos de la herramienta. Se utilizó la plataforma colaborativa Trello¹ para extraer historias de usuario y someterlas a un proceso de inspección y análisis.

Las historias de usuario extraídas fueron sometidas al analizador con la intención de detectar posibles sentimientos inyectados durante las tareas de especificación. En caso de detectar inserciones sentimentales, es posible visualizar las secciones donde fueron identificadas, a través la interfaz gráfica diseñada. La misma, además, proporciona un mensaje indicando al usuario la necesidad de revisión y refinamiento a los fines de mejorar su objetividad, tarea bajo responsabilidad del equipo de trabajo involucrado en las actividades de educación y refinamiento.

Cabe destacar que la interfaz propuesta proporciona el perfil del usuario y la historia de usuario asignada, lo que facilita el registro de las modificaciones realizadas por un usuario determinado, contribuyendo esto al control de versiones.

Por otro lado, también se proporciona un menú principal, desde donde es posible seleccionar la historia de usuario del proyecto que se desea analizar. La Figura 6 presenta la interfaz gráfica de la herramienta propuesta donde es posible visualizar las historias de usuario seleccionadas para su análisis, siguiendo el procesamiento detallado en la Sección 4.

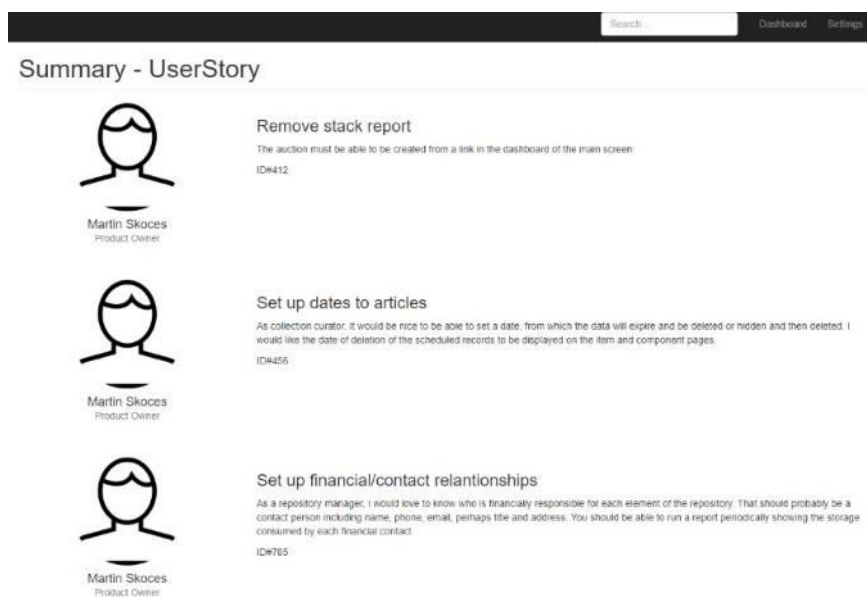


Fig.6. Interfaz gráfica de historias de usuario seleccionadas para analizar.

A continuación, se describen dos historias de usuario extraídas para su análisis de la plataforma colaborativa Trello. Las mismas se encuentran redactadas en inglés. Por lo tanto, para especificar el proceso de análisis realizado en este trabajo, se usará su traducción al idioma español.

El primer ejemplo enuncia: *“Como administrador de la colección. Sería bueno poder fijar una fecha, a partir de la cual los datos expirarán y se borrarán u ocultarán y*

luego se borrarán. Deseo que la fecha de eliminación de los registros programados se muestre en las páginas del artículo y de los componentes". En esta historia de usuario, se detectaron palabras y expresiones que expresan sentimientos. Ellas son: "sería bueno", "expirarán" y "deseo". Aquellas que el analizador detectó y que se tuvieron en cuenta para reemplazar en el texto son "sería bueno" y "deseo", mientras que "expirarán" en realidad no aporta subjetividad al texto introducido.

Luego de ser identificadas y visualizadas en la interfaz, el miembro a cargo de las tareas de especificación y refinamiento puede realizar las modificaciones pertinentes, a fin de obtener historias de usuario objetivas, sin expresiones sentimentales ni grados de preferencias sobre las funcionalidades solicitadas. A modo de ejemplo, se presenta la redacción alternativa que se propone para esta historia de usuario, tarea bajo responsabilidad del equipo de trabajo a cargo su redacción: *"Como administrador de la colección. Quiero poder establecer una fecha después de la cual los datos caducarán y se borrarán u ocultarán y luego se borrarán. Debo poder ver la fecha de eliminación de los registros programados que se muestre en las páginas del artículo y de los componentes"*. La Figura 7 muestra la historia de usuario sometida al analizador de sentimientos y su redacción alternativa.



Fig. 7. Historia de usuario con inyección de sentimientos y su redacción alternativa.

La Figura 7 muestra las secciones en las que el analizador identificó expresiones sentimentales. La interfaz proporciona en el margen inferior un cuadro de texto para insertar la historia de usuario modificada de acuerdo a los criterios del miembro responsable de su refinamiento.

En un segundo ejemplo, la historia de usuario original describe lo siguiente: *"Como gestor de repositorios, me encantaría saber quién es financieramente responsable de cada elemento del repositorio. Esa debe ser probablemente una persona de contacto y*

debe incluir nombre, teléfono, correo electrónico, tal vez título y dirección. Debería poder ejecutar un informe periódicamente que muestre el almacenamiento consumido por cada contacto financiero". En este caso, sólo se detectó que la expresión "me encantaría" presenta sentimientos, por lo que se propone su reemplazo para obtener una descripción neutral. La redacción alternativa que se propone en este segundo ejemplo es: *"Como administrador de repositorios, quiero una forma de rastrear quién es financieramente responsable de cada ítem en el repositorio. Probablemente una persona de contacto y debe incluir nombre, teléfono, correo electrónico, título y dirección. Debo poder ejecutar un informe periódicamente que muestre el almacenamiento consumido por cada contacto financiero"*.

Al extraer las palabras que insertan sentimientos, es posible observar un texto con menor grado de subjetividad y mayor neutralidad en la expresión. Por otro lado, esta herramienta ayuda a detectar las posturas y puntos de vista del product owner con respecto a la funcionalidad o necesidad especificada en la historia de usuario y, además, facilita la obtención de una especificación sin grados de preferencia por las funcionalidades que se solicitan. Esto contribuye con las tareas de priorización de historias de usuario en el backlog, ya que esta tarea debe ser llevada a cabo de acuerdo a los criterios técnicos, sin ser influenciada por los sentimientos expresados en las historias de usuario.

6 Conclusiones y Trabajos Futuros

En este trabajo se presentó el prototipo de una herramienta para realizar análisis de sentimientos sobre historias de usuario, con la intención de mejorar su objetividad durante la especificación de requisitos en los proyectos de desarrollo que implementan metodologías ágiles. El prototipo del analizador desarrollado toma como entrada historias de usuario redactadas en lenguaje natural, junto a un léxico de sentimientos denominado Bing. Las historias de usuario son sometidas a un proceso de inspección y análisis. En caso de detectar expresiones sentimentales, la herramienta advierte al usuario esta situación y devuelve una sugerencia de revisión sobre aquellas secciones de las historias de usuario en las que se detectaron sentimientos.

La ejecución de la herramienta de análisis propuesta sobre las historias de usuario ha demostrado ser útil para reducir la ambigüedad presente en estas especificaciones que conduce a la inserción de palabras y frases que expresan sentimientos. Es por ello, que puede ser utilizada como herramienta de soporte en las tareas de educación y refinamiento de las historias de usuario. Actualmente, la intención de los autores es analizar el nivel de aceptación de la propuesta por parte de los miembros responsables de las tareas de educación de requisitos, y la viabilidad de su aplicación en contextos ágiles.

Durante la ejecución del prototipo sobre las historias de usuario seleccionadas, se ha detectado la ausencia de ciertas palabras que expresan sentimientos en el léxico utilizado, lo que induce a error en la identificación. Es por esto, que planeamos abordar en próximas investigaciones el uso de otros léxicos y diccionarios propios del dominio de la Ingeniería de Software.

Como trabajos futuros, se pretende, por un lado, ejecutar la herramienta sobre historias de usuario en contextos de desarrollo real, a fin de validar la propuesta. Por otro lado, la intención de los autores es mejorar la herramienta a través de la inserción de métricas que permitan medir la objetividad alcanzada luego de la detección de los sentimientos inyectados en las historias de usuario y su posterior refinamiento realizado por el equipo de trabajo. Conjuntamente, se pretende generar pruebas sobre muestras en conjuntos de datos públicos de mayor tamaño y realizar comparaciones sobre los resultados con el objetivo de incorporar mejoras a la herramienta.

Reconocimientos

Los autores agradecen el apoyo brindado por la Universidad Tecnológica Nacional (SIUTIFE0004923TC).

Referencias

1. Moitra, J., Herbsleb, D.: Global software development. *IEEE Softw.* 18, 16–20 (2001).
2. Garrison, G., Wakefield, R.L., Xu, X., Kim, S.H.: Globally distributed teams: The effect of diversity on trust, cohesion and individual performance. *ACM SIGMIS Database DATABASE Adv. Inf. Syst.* 41, 27–48 (2010).
3. Farmer, M., "DecisionSpace infrastructure: agile development in a large, distributed team," Agile Development Conference, Salt Lake City, UT, pp. 95-99 (2004).
4. Fowler, M.: Using an Agile Software Process with Offshore Development. 1–8 (2003).
5. Abrahamsson, P., Salo, O., Ronkainen, J., Warsta, J.: Agile software development methods: Review and analysis. *arXiv Prepr. arXiv1709.08439* (2017).
6. Rees, M.: A feasible user story tool for agile software development? *Proc. - Asia-Pacific Softw. Eng. Conf. APSEC.* 22–30 (2002).
7. Marshall, A., Gamble, R.F., Hale, M.L.: Outcomes of Emotional Content from Agile Team Forum Posts. 2016 IEEE/ACM 1st Int. Work. Emot. Aware. *Softw. Eng.* 6–11 (2016).
8. Mäntylä, M. V., Graziotin, D., Kuutila, M.: The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers. 27, 16–32 (2018).
9. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.* pp. 168–177 (2004).
10. Jurado, F., Rodriguez, P.: The Journal of Systems and Software Sentiment Analysis in monitoring software development processes : An exploratory case study on GitHub ' s project issues. *J. Syst. Softw.* 104, 82–89 (2015).
11. Patwardhan, A.S.: Sentiment Identification for Collaborative, Geographically Dispersed, Cross-functional Software Development Teams. 20–26 (2017).
12. Wrobel, M.R.: Emotions in the software development process. In: *Human System Interaction (HSI). The 6th International Conference on.* pp. 518–523 (2013).
13. Jorgenson, N., Hale, M., Gamble, R., "SEREBRO: facilitating student project team collaboration". The 33rd International Conference on Software Engineering Waikiki (2011).
14. Islam, M.R., Zibran, M.F.: A comparison of software engineering domain specific sentiment analysis tools. In: *2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER).* pp. 487–491 (2018).
15. Agile Manifesto, Avail. at <http://agilemanifesto.org/>, (2001).
16. Deemer, P., Benfield, G., Larman, C., Vodde, B. *The scrum primer*, Avail. at

- <http://assets.scrumtraininginstitute.com/downloads/1/scrumpriemer121.pdf>. (2010).
17. Lai, S.-T.: A User Story Quality Measurement Model for Reducing Agile Software Development Risk. *Int. J. Softw. Eng. Appl.* 8, 75–86 (2017).
 18. Wnuk, K., Maddila, K.C.: Agile and lean metrics associated with requirements engineering. *Proc. 27th Int. Work. Softw. Meas. 12th Int. Conf. Softw. Process Prod. Meas. - IWSM Mensura '17.* 33–40 (2017).
 19. Lucassen, G., Dalpiaz, F., van der Werf, J.M.E.M., Brinkkemper, S.: Improving agile requirements: the quality user story framework and tool. *Requir. Eng.* 21, 383–403 (2016).
 20. The R Project for Statistical Computing. <https://www.r-project.org/>. Accessed 8 Feb. 2019.
 21. Hadley, W. “Tidyverse”. <https://www.rdocumentation.org/packages/tidyverse/versions/1.2.1>. Accessed 8 Feb. 2019.
 22. Silge, J. “Tidytex”. <https://www.rdocumentation.org/packages/tidytex/versions/0.2.0>. Accessed 8 Feb. 2019.
 23. Nielsen, F. A. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages* 718 in *CEUR Workshop Proceedings* 93-98 (2011).
 24. Kiritchenko, S., Zhu, X., & Mohammad, S. M. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50, 723-762 (2004).
 25. Liu, B.: Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* 5, 1–167 (2012).