



SISTEMA PARA IDENTIFICAR, EXTRAER Y VISUALIZAR CONOCIMIENTO EN CORREOS ELECTRÓNICOS

Juan David Cruz Gómez

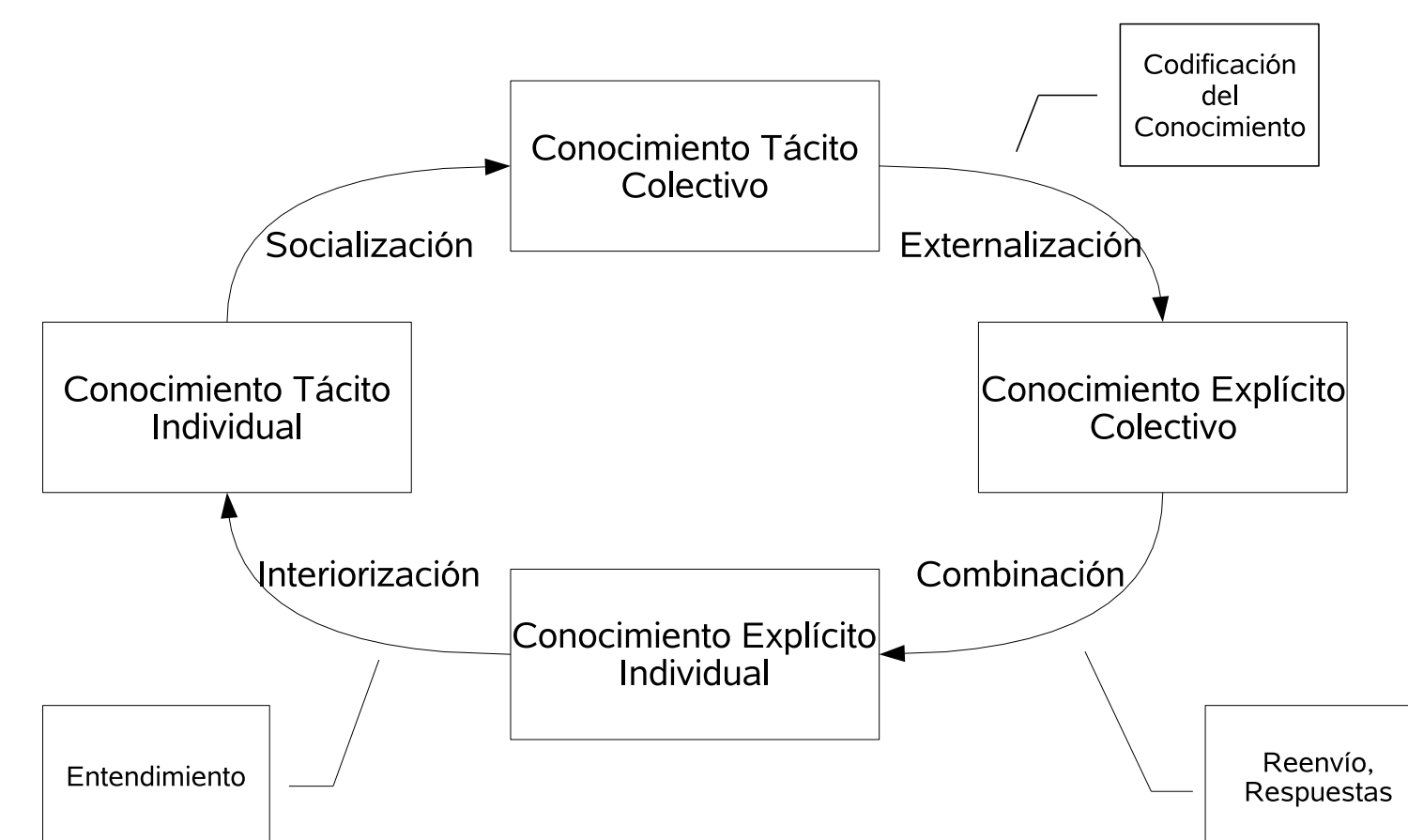
Universidad Nacional de Colombia, Maestría en Ingeniería – Ingeniería de Sistemas y Computación
Laboratorio de Investigación en Sistemas Inteligentes



Introducción

El conocimiento debe ser tratado como un activo más en las compañías, y como tal, debe ser gestionado de modo que agregue un gran valor a la misma. Este conocimiento es producido, transformado y transmitido por personas. Cada una de estas personas desempeñan diferentes roles dentro de la organización, por lo que se comunican con diferentes personas, dentro y fuera de ella sobre diferentes temas; esto indica un movimiento de conocimiento. Este trabajo presenta el avance hacia un modelo de gestión de conocimiento que parte de la información de los correos electrónicos personales.

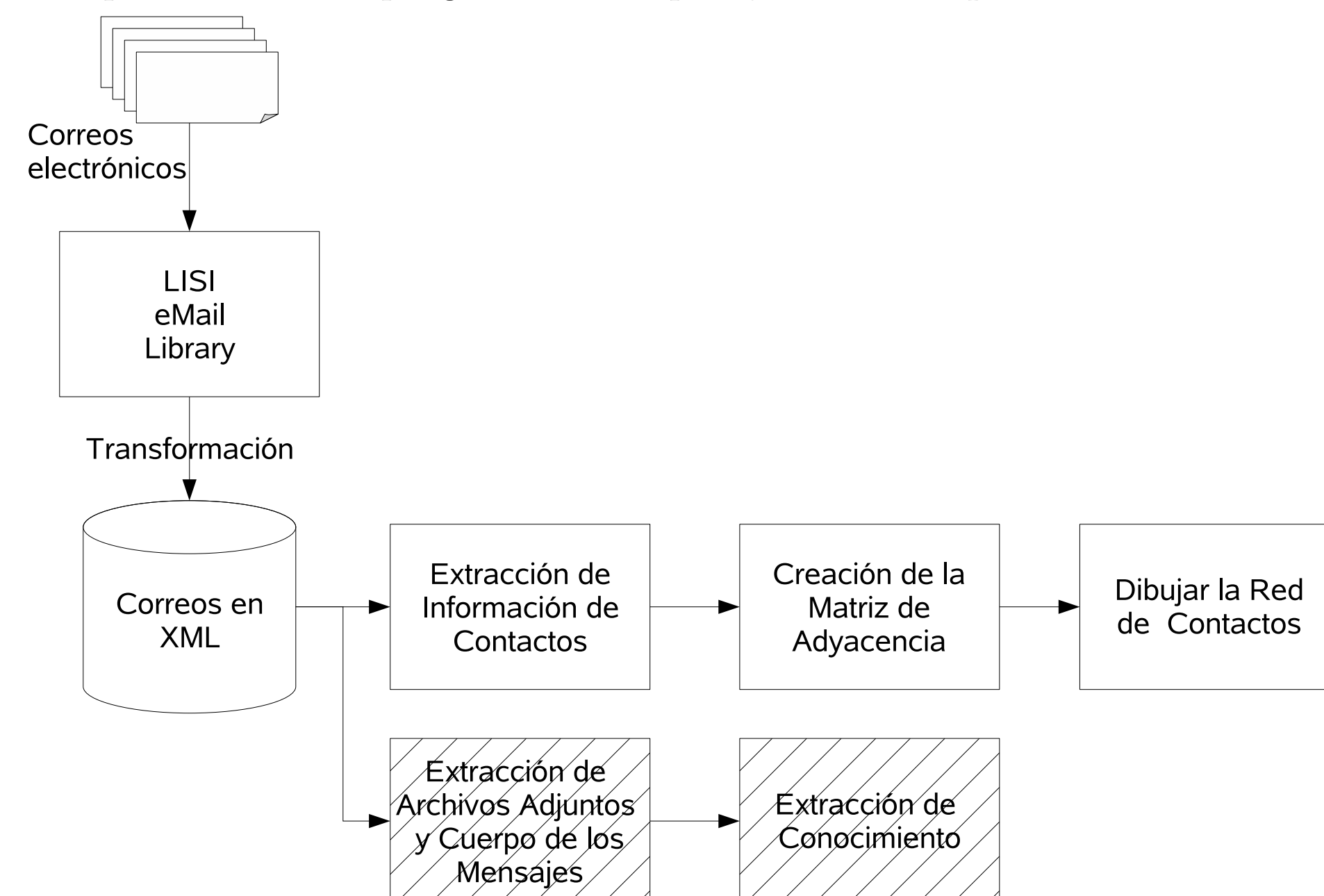
El conocimiento tácito es el que es generado a través de la experiencia y es muy difícil de codificar, por lo tanto es muy difícil de extraer y gestionar.



En la figura, se ve el ciclo del conocimiento propuesto por Nonaka [NT95], el cual indica el proceso de transformación que existe entre el conocimiento tácito y el explícito.

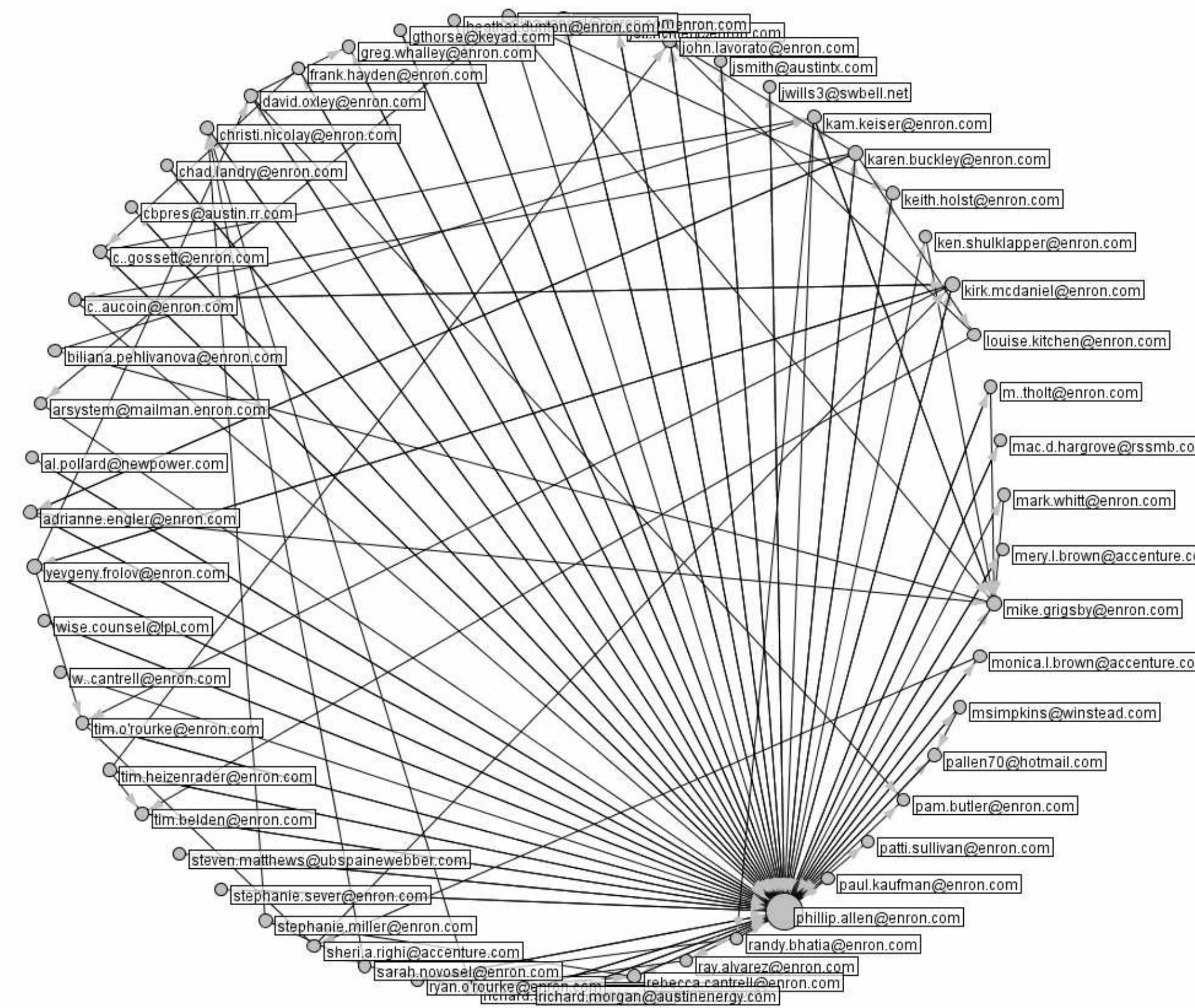
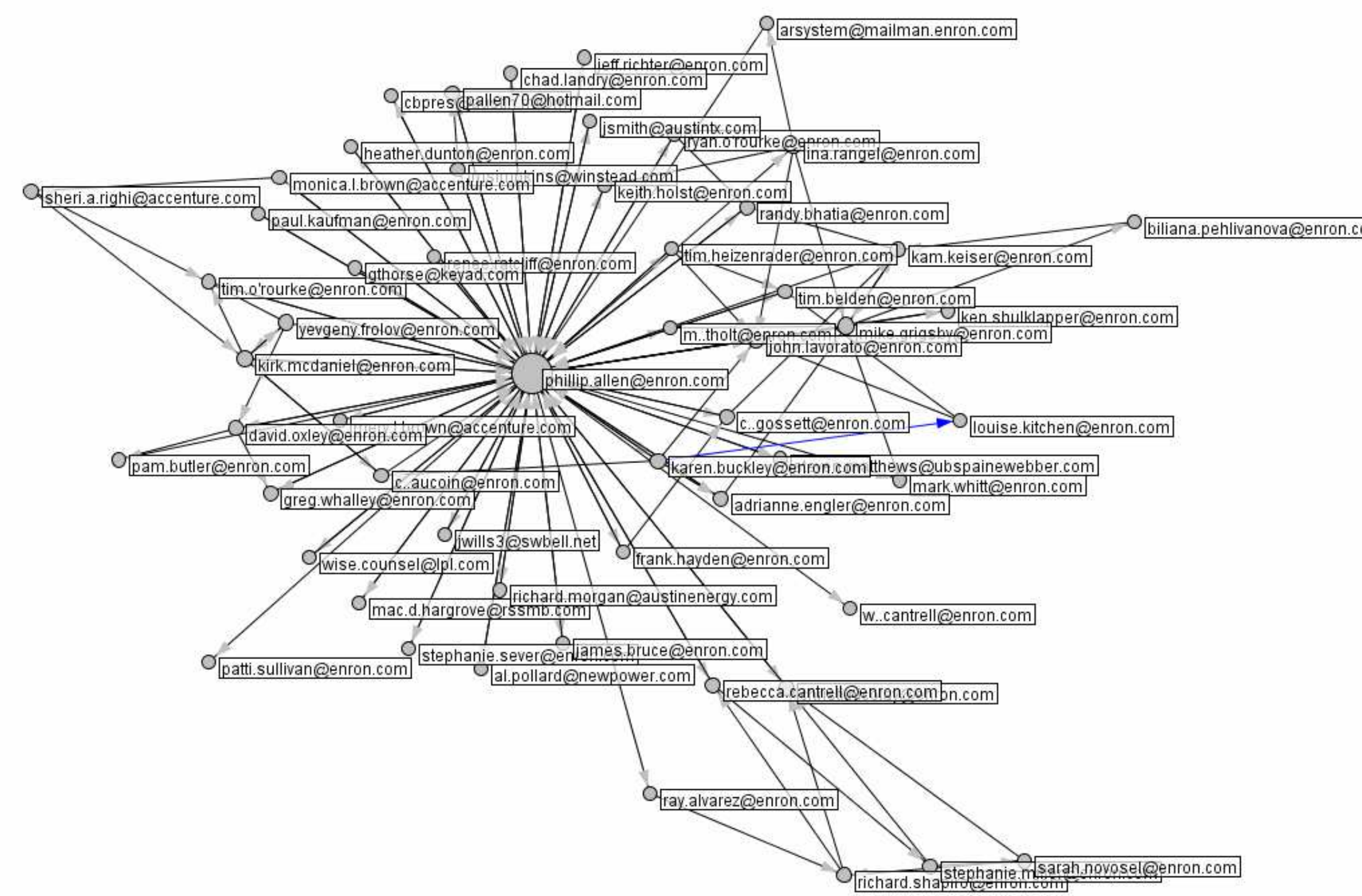
Planteamiento del Problema

Se debe desarrollar un sistema de análisis de correos electrónicos que permita extraer las características relevantes, como información de contactos y temas, entre otros, de los mensajes y de los documentos adjuntos de modo que se puedan asociar con las relaciones identificadas en las redes sociales extraídas de los mensajes de correo electrónico. Con esta asociación se busca darle un sentido más amplio a las relaciones identificadas entre personas, permitiendo que el usuario pueda resolver preguntas del tipo *¿Quién sabe qué?*.



En la gráfica se ve la arquitectura preliminar del sistema.

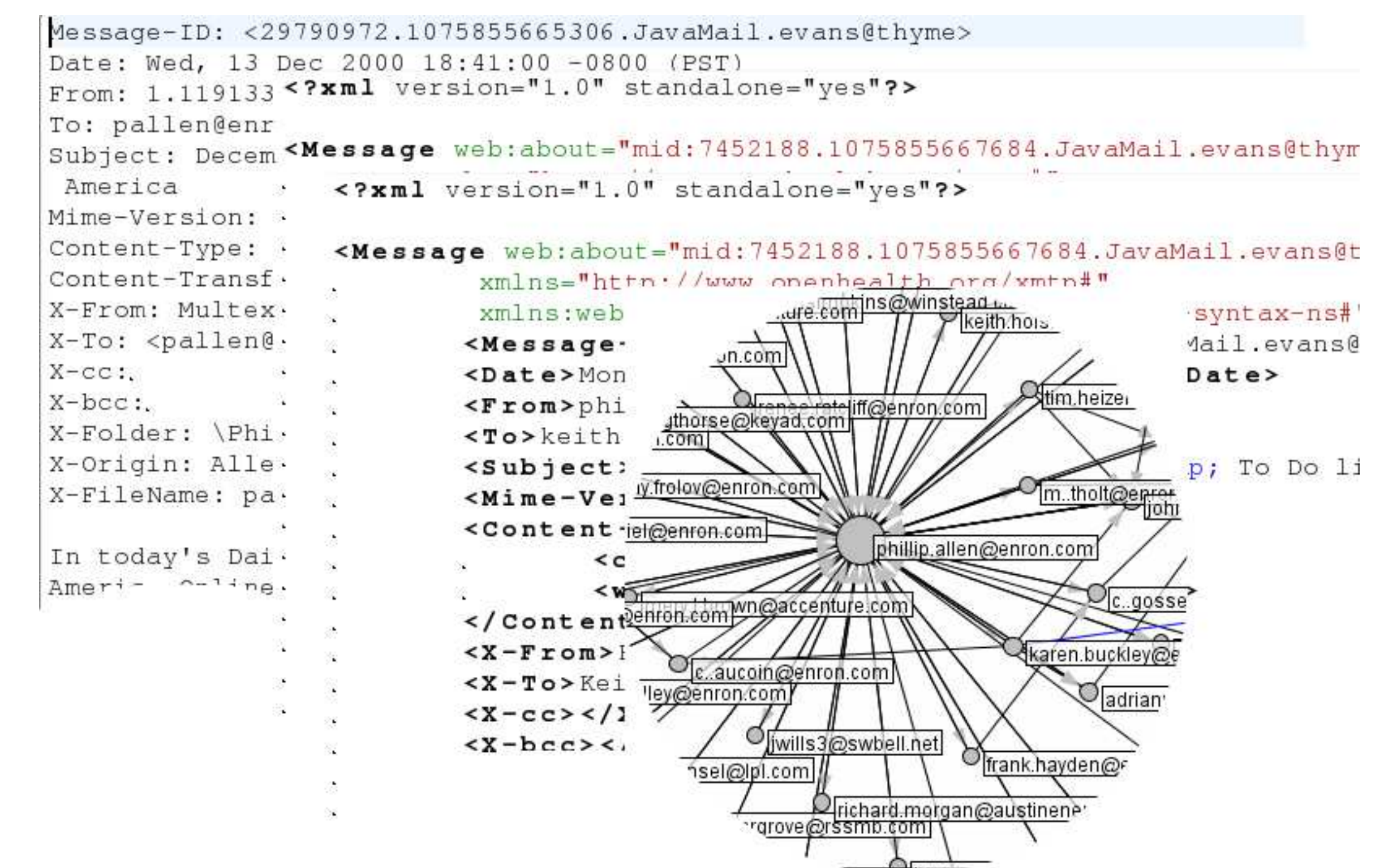
Red de Ejemplo



Las redes anteriores pertenecen al mismo dueño, la diferencia es el método de pintado que presentan. En la primera, es evidente que se trata de una red personal (egocéntrica).

El Sistema

El primer paso es convertir los mensajes de correo en XML, y luego, se genera una matriz de adyacencias, con la que la red será construida.



En este caso, se no se utiliza una matriz de 1 y 0, sino listas de nodos, donde el primero, corresponde a origen, y el resto de la lista, al destino.

Trabajo Futuro

- Analizar el texto de los correos y de los archivos adjuntos.
- Extraer la información extraída y relacionarla con la red social
- Revisar otros algoritmos de organización de redes para visualizar al mismo tiempo categorías de texto y redes sociales.
- Identificar el conocimiento tácito de los documentos y del flujo de la información.

References

[NT95] I Nonaka and H Takeuchi. *The knowledge creating company*. Oxford University Press, 1995.