



Estructura para Propuestas de Tesis de Magister

Dr. John Atkinson

Departamento de Ingeniería Informática

Universidad de Concepción

2011





Estructura

- Resumen
- Formulación de la Propuesta
- Hipótesis
- Objetivos
- Metodología
- Plan de trabajo



Resumen

- Debe ser claro e informativo. Describe los principales puntos que se abordarán en el proyecto incluyendo objetivos, metodología y resultados esperados.
- Un resumen bien formulado facilita la adecuada descripción y comprensión del proyecto. (1 página)



Formulación de la Propuesta

- Esta sección debe contener la exposición general del problema, precisar los aspectos nuevos a desarrollar, señalando los enfoques actualmente en uso en el tema de investigación, así como los fundamentos teóricos y análisis bibliográfico que lo avalan.(8 pág)



Hipótesis

- Diferentes enfoques para plantear una hipótesis:
 - Proposición del estudio
 - Pregunta a responder
 - Problema a resolver
 - Presunciones a probar

Hipótesis de Colón: El mundo es redondo, navegando hacia el Oeste se retorna al Este.



Hipótesis

¿Qué se hace con una hipótesis?

- Se prueba (prueba analítica)
- Se entrega evidencia que apoya o refuta la hipótesis (empírica)
- Un plan de cómo probar la hipótesis :
 - Resultados deben ser *medibles*
 - Resultados debe ser *repetibles*
 - Resultados deberían ser *generalizables*
- Sugerencias para formular una hipótesis:
 - *A es mas rápido/más pequeño/menos complejo que B*
 - *A es equivalente a B*



Objetivos

- **Objetivos**
 - Los objetivos se refieren a metas específicas que se cumplen si la hipótesis es verdadera.
- **Objetivos específicos**
 - Plantear objetivos específicos que cuya resultados puedan ser medibles o informados con productos concretos.



Metodología

- Describe los métodos que se planea utilizar para abordar los objetivos. Una buena estrategia es usar cada uno de los objetivos como guía para describir el cómo se llevan a cabo. (3 pág)



Plan de Trabajo

- Relacione los objetivos planteados con etapas y actividades concretas, señalando el tiempo de comienzo y término de cada una (1 pag)



Estructura

- Resumen (1 pag)
- Formulación de la propuesta (8 pag)
- Hipótesis
- Objetivos
- Metodología (3 pag)
- Plan de Trabajo (1 pag)



Formulación: Motivación

- *“En un proceso de recuperación de información, el usuario se enfrenta generalmente a un conjunto de documentos que responden a una consulta. Cuando se encuentra en esta situación, el usuario debe decidir qué y en cuál orden recorrerá los documentos para obtener la información requerida. En el peor de los casos, será necesario recorrer todos los documentos para obtener toda la información disponible. Los principales inconvenientes de este procedimiento son los siguientes:*
 - *Los usuarios no están dispuestos a recorrer muchos documentos y la lista de resultados puede ser enorme.*
 - *Muchos documentos duplican o comparten mucha información. Por lo tanto, obtener toda la información no es igual a recorrer todos los documentos.”*
- *“Este trabajo propone definir un modelo que apoye la navegación de información. En el contexto de sistemas de información, navegación de información se asocia a la visualización de los resultados de un proceso de recuperación de información, lo cual.....”*



Formulación: Originalidad

- *“La idea de esta tesis es mejorar la entrega de resultados de un proceso de recuperación con un esquema de navegación. A pesar de que existen trabajos orientados a apoyar la visualización de resultados de consultas vía procesamiento de clustering, esta propuesta apunta a desarrollar un nuevo modelo que vaya más allá de indicar documentos similares y que, a partir de las similitudes y diferencias entre documentos, proponga un camino de recorrido, es decir, una secuencia de documentos a recorrer. La implementación de este modelo de navegación debe tomar en cuenta el costo computacional de manera de hacer posible su aplicación real.”*



Formulación: Trabajo Relacionado

- *Un conjunto de documentos temáticamente relacionados puede provenir de una biblioteca digital que organiza documentos en clusters o puede ser un conjunto de documentos que resultan de una búsqueda en una base de datos textual. En ambos casos, el problema involucra recorrer estos documentos. En recuperación de información, un enfoque exitoso se ha basado en una descripción o resumen del contenido del conjunto de documentos. Los enfoques más usuales incluyen [11]:*
 - *Jerarquía por categorías, donde se asignan etiquetas al documento de acuerdo a su categoría, y los documentos se ordenan jerárquicamente. Ejemplos de sistemas que utilizan jerarquías son ACM (Association for Computing Machinery) que ha desarrollado una jerarquía de aproximadamente 1200 etiquetas de categorías, o el sitio de búsqueda Yahoo [16] ...*



Formulación: Trabajo Relacionado

- *Un aspecto importante para cualquiera de los enfoques de generación de resúmenes y estrategias de visualización es la comparación de documentos. Para comparar documentos, dos tipos estrategias son: comparación por palabras claves y análisis contextual por frases. Varios de los métodos propuestos en base a estas dos estrategias usan en alguna medida el modelo vectorial de recuperación de información [2] [7] [3] [5]. Estos métodos utilizan generalmente la comparación de documentos para realizar agrupamiento de resultados (clustering) [3][5][10]. En [10] se entrega una completa revisión y evaluación de técnicas de agrupamiento de documentos.*



Formulación: Resultados Esperados

- *El producto de esta tesis será la definición de un modelo de representación de las relaciones de contenido de documentos, un algoritmo de navegación entre estos documentos y un prototipo que reciba un conjunto de documentos de entrada (posiblemente la salida de un buscador tradicional) y de una sugerencia de navegación que cambie dinámicamente a medida que el usuario efectivamente visite los documentos*



Hipótesis

- *“Este trabajo plantea como hipótesis que el recorrido de documentos de una lista de documentos ordenados, en base a un ranking clásico de recuperación de información, es menos eficiente que la utilización de una estrategia de recorrido que considere el contenido de información de los documentos recuperados”*



Objetivos

- *Objetivo:*
Definir una estrategia de navegación que permita recorrer eficientemente un conjunto de documentos de un espacio de información.
- *Objetivos Específicos:*
 - *Analizar estrategias de comparación de documentos, basados tanto en frases como en palabras claves.*
 - *Definir una estrategia eficiente de comparación de documentos, que permita obtener tanto la similitud como la diferencia entre ellos.*
 - *Modelar el espacio de información mediante una estructura que capture las similitudes y diferencias entre documentos.*
 - *Crear un prototipo para realizar las pruebas.*



Metodología

- *Para evaluar las estrategias de comparación de documentos, se considerará un conjunto de documentos relacionados semánticamente, los cuales serán caracterizados mediante modelos propuestos con anterioridad. Como avance a esta propuesta se han revisado ...*
- *Para modelar el espacio de información se pretende utilizar teoría de grafos. Se propone modelar el espacio de información como un grafo dirigido, donde cada nodo representa*
- *La implementación del prototipo considerará dominios de documentos Web, para el cual se analizaran dos casos:*
- *Como método de evaluación se considerará esencialmente la acumulación de ganancia de información en la secuencia de documentos recorridos, de manera que una ganancia mayor y más rápida sea la deseada.*