



Departamento de Ciencias de la Computación  
UNIVERSIDAD DE CHILE

# Documento de revisión de Diseño

## Repositorio de papers NicLabs

<b>Grupo:</b>	4
<b>Administrador de proyecto:</b>	Sebastián González
<b>Analista:</b>	Francisco Carter
<b>Tester:</b>	José Ojeda
<b>Diseñadores-Implementadores:</b>	Juan José Alegría Matías Rodríguez Romina Romero
<b>Fecha entrega:</b>	Martes 05 de Mayo de 2015

## Personas que expusieron

- **Juan José Alegría** (Diseñador-Implementador)
- **Matías Rodríguez** (Diseñador-Implementador)

## Observaciones durante la presentación de diseño

- Equipo docente se señala que el ingreso de *papers* “luce” manual en los diagramas presentados. Grupo aclara que cliente solicita la creación de un *script* que llame automáticamente a aquella funcionalidad de ingreso de datos a sistema.
- En la revisión anterior, equipo docente sugiere investigar la existencia de CDs con conferencias. Grupo indica que preguntó a cliente, obteniendo una respuesta negativa.
- Respecto al modelo de datos, equipo docente señala que sería bueno ya tener una métrica para similitud entre *papers*, pero plantea dudas respecto de si es necesario implementarla para la 1° iteración (habiendo cosas más urgentes que hacer).
- Equipo docente aclara que un *paper* puede aparecer en un sólo *proceeding*, pues no sería adecuado que haya “autoplagio”. A nivel de base de datos, la cardinalidad de la relación entre tablas implicada, no está bien.
- Equipo docente nota que si la *data* a procesar viene errónea (respecto de caracteres ilegibles), el software en construcción debería poder ofrecer la posibilidad de editar textos relacionados con un *paper* (ejemplo: podrían haber problemas de codificación de caracteres cirílicos o de otro tipo, los que podrían mostrarse incorrectamente en el sistema, lo cual no es deseable).
- Respecto al punto anterior, grupo indica que ése proceso de edición está considerado (tal vez no como un formulario web), pero por ahora está la posibilidad de ir directamente a la base de datos para corregir posibles caracteres ilegibles.
- Equipo docente sugiere la posibilidad de escoger varios *papers*, para luego descargarlos.
- Equipo docente solicita hacer explícito el algoritmo para “parsear”, y qué se “parseará” exactamente.

## Observaciones que aparecieron durante el transcurso de la creación de diseño

- Identificación del tema de un *paper*: De lo contemplado hasta ahora, es bastante complejo identificarlo desde el archivo pdf, por lo que se ha pensado en dejarlo para la iteración 2.
- Similaridad de *keywords*: Por ejemplo, “big data” y “Big\_Data” corresponden fundamentalmente a la misma palabra clave (“Big Data”), pero sus caracteres son distintos. Se ha resuelto que **no** se hará algo al respecto (con el diseño planteado hasta el momento, ambas palabras clave se considerarán distintas), pues intentar resolverlo añadiría complejidad extra al proyecto (manejo de sinónimos, diccionarios, etc.).
- Buscar restricciones dentro de faceta: Podría darse el caso que una faceta tuviera asociadas muchísimas restricciones (por ejemplo, la faceta “Autor” podría tener cientos de restricciones), por lo que se ve la posibilidad de agregar un pequeño buscador dentro de las facetas (o bien, sólo para las facetas que poseen restricciones numerosas).

## Comentario del tester

El día **viernes 24 de Abril** hubo una reunión en la que estuvieron presentes: administrador, analista y todos los diseñadores-implementadores, en la cual se “aterrizaron” varios puntos clave del diseño presentado.

Lo anterior, junto al posterior trabajo de los diseñadores-implementadores, incidió en una gran claridad – notada por el equipo docente en el transcurso presentación del diseño – del documento de diseño generado, que de seguro permitirá avanzar a la etapa de implementación teniendo ya bastante claro qué hay que hacer y cómo.