



# UAGro

UNIVERSIDAD AUTÓNOMA DE GUERRERO

Unidad Académica de Matemáticas

Maestría en Ciencias: Área Matemática Educativa

## Diseño de Actividades de Estudio e Investigación para una formación de Ingenieros en Sistemas Computacionales: El caso de PageRank

Tesis que presenta

**Julián Huitzi Patricio Martínez**

Para obtener el grado de

Maestría en Ciencias **Área Matemática Educativa**

Directores de tesis:

**Dra. Avenilde Romo Vázquez**  
**M.C. Miguel Díaz Cárdenas**

Debe ser primero el profesor Miguel Díaz  
Cárdenas, del Posgrado

**Chilpancingo, Guerrero, México.**





# Contenido

RESUMEN .....	6
INTRODUCCIÓN.....	7
1 LA MODELACIÓN EN LA FORMACIÓN Y LA PRÁCTICA DEL INGENIERO .....	9
1.1 Introducción .....	9
1.2 Matemáticas en la formación de futuros ingenieros y su evolución.....	9
1.3 El papel de la modelación matemática en la industria.....	12
1.3.1 Roles de la modelación matemática entre matemáticas e industria.....	12
1.3.2 Formación para la profesión, un ejemplo en la formación española de ingenieros.....	15
1.4 Modelación matemática en la formación.....	18
1.4.1 Proceso de modelación, competencia de modelación y condiciones para su implementación en el aula .....	19
1.4.2 La modelación matemática en el área laboral y su posible impacto sobre el aula .....	22
1.4.3 Modelación matemática en Ingeniería.....	23
1.5 Conclusión .....	24
2 ELEMENTOS DE LA TEORÍA ANTROPOLÓGICA DE LO DIDÁCTICO .....	26
2.1 Introducción .....	26
2.2 Nociones básicas.....	27
2.2.1 Institución .....	27
2.2.2 Praxeología .....	28
2.2.3 Praxeologías matemáticas y mixtas.....	28
2.3 Modelo praxeológico extendido .....	29
2.4 Niveles de organización praxeológica, proceso de estudio y niveles de co-determinación .....	31
2.4.1 Proceso de estudio y sus momentos .....	31
2.4.2 Niveles de determinación .....	33
2.5 Conclusión .....	34
3 CONSIDERACIONES METODOLÓGICAS .....	35
3.1 Introducción .....	35
3.2 Instituciones presentes en la formación del ingeniero en sistemas computacionales .....	35

3.3	Metodología para el diseño de actividades de modelación para formación de futuros ingenieros .....	37
3.3.1	Primera fase: Elección de un contexto extra matemático.....	38
3.3.2	Segunda fase: Análisis praxeológico e identificación de un modelo matemático.....	39
3.3.3	Tercera fase: Análisis del modelo en $E(M)$ .....	39
3.3.4	Cuarta fase: Diseño de la actividad didáctica.....	40
3.3.5	Consideraciones para la implementación en el ITCH .....	40
3.4	Conclusión .....	41
4	ANÁLISIS PRAXEOLÓGICO DE LA MODELACIÓN MATEMÁTICA EN EL MÉTODO PAGERANK .....	42
4.1	Introducción .....	42
4.2	Análisis del método PageRank .....	42
4.2.1	La problemática asociada a los buscadores Web.....	42
4.2.2	Construcción de la técnica (método) PageRank en D1 .....	45
4.2.3	Construcción de la técnica (método) PageRank en D2 .....	48
4.2.4	Validación e Implementación del Modelo de Importancias en D1 .....	56
4.2.5	Validación e Implementación del Modelo de importancias en D2 .....	60
4.3	Del análisis praxeológico de los buscadores web hacia el diseño de una actividad didáctica .....	65
4.3.1	Arborescencia praxeológica .....	65
4.3.2	Elección de una praxeología para el diseño de la actividad didáctica y su vinculación con otras praxeologías de modelación .....	73
4.4	Otros contextos de uso .....	75
4.5	Conclusión .....	76
5	LA ORGANIZACIÓN DIDÁCTICA basada en “PageRank” .....	78
5.1	Introducción .....	78
5.2	El diseño de la organización didáctica, “La problemática de los buscadores Web” .....	78
5.3	El entorno tecnológico teórico base del diseño de este proceso de estudio.....	79
5.4	Conclusión .....	91
6	RESULTADOS .....	93
6.1	Introducción .....	93
6.2	Condiciones de la implementación y reorganización de la organización didáctica.....	93

6.3	Análisis de la actividad 1 .....	94
6.3.1	Planteamiento de la problemática.....	95
6.3.2	Análisis de la problemática.....	96
6.3.3	Simulación de la situación: ordenación empírica vs criterio de ordenación .....	101
6.3.4	Representaciones de la estructura de la Web: grafo y matriz.....	105
6.3.5	Análisis de la estructura de la Web .....	107
6.4	Segunda Actividad, reorganización del proceso de estudio.....	109
6.4.1	Análisis de la estructura de la Web .....	109
6.4.2	Aproximaciones a los criterios de importancias.....	112
6.5	CONCLUSIÓN.....	114
7	CONCLUSIONES.....	117
8	REFERENCIAS .....	122

## RESUMEN

El presente trabajo de investigación tiene como objetivo general vincular la formación matemática de futuros ingenieros con la práctica profesional, a partir del diseño de una actividad de estudio y de investigación basada en modelación matemática.

Para situar esta problemática de investigación se analizan algunas contribuciones hechas a los estudios ICMI 3, ICMI 14 e ICMI 20, que analizan los roles de la modelación matemática tanto en la enseñanza de las matemáticas como en la industria. Asimismo se analizan trabajos de Bisell (2002) y de Romo-Vázquez (2014) que muestran particularidades de la actividad de modelación matemática en la ingeniería.

El trabajo se enmarca en la Teoría Antropológica de lo Didáctico que brinda herramientas, como las de institución y praxeología, para analizar la formación del ingeniero en sistemas computacionales, reconociendo las instituciones que la conforman, las relaciones entre éstas y las actividades de modelización matemática. Asimismo, estas herramientas se utilizan para analizar un contexto profesional, la programación Web y en particular la ordenación de una búsqueda en Internet, a partir de un documento representante de la institución profesional (Fernández, 2004) y de otro que involucra además de la práctica profesional, el álgebra lineal y su enseñanza (D'Andrea, 2012). El análisis praxeológico de estos documentos evidenció modelos matemáticos como grafos y matrices, que conforman el método utilizado por Google para determinar que una búsqueda en internet sea eficiente y que es conocido como PageRank. Este análisis sustenta la estructuración y el diseño de una Actividad de Estudio e Investigación (AEI), que busca generar una relación entre la formación matemática y la práctica profesional.

La AEI diseñada fue implementada en un grupo del ITCH, con un subsecuente análisis que permite mostrar sus potencialidades y límites, así como el planteamiento de posibles refinamientos.

Esta investigación presenta de manera general una propuesta teórico-metodológica para el diseño de actividades didácticas (y en particular AEI), para vincular las necesidades de la práctica profesional con la enseñanza de las matemáticas.

## INTRODUCCIÓN

La formación matemática del futuro ingeniero, en sus diferentes especialidades, tiene el propósito de brindarle herramientas que lo preparen para su futura vida profesional. Sin embargo, se ha observado la existencia de una desconexión entre dicha formación y las necesidades matemáticas actuales de las prácticas profesionales, que han evolucionado vertiginosamente en efecto al imponente desarrollo tecnológico. ¿Cómo adaptar la formación matemática a estas nuevas necesidades?

El análisis de diferentes trabajos (Li, 2013; Albertí, Amat, Busquier, Romero y Tejada, 2013; Geiger, 2013 y Ye, 2013) nos permiten identificar a la modelación matemática como “puente” entre la matemática y sus aplicaciones en el ámbito profesional. Esto nos lleva a cuestionar de manera más particular, ¿qué se requiere para lograr que modelación matemática que tiene lugar en la práctica profesional esté en el aula de matemáticas del futuro ingeniero? El análisis de trabajos (Artaud, 2007; Wake, 2007) nos permite considerar la necesidad de elegir contextos relevantes de la práctica profesional del futuro ingeniero, así como el análisis de la actividad de modelación que se realiza en éstos, como una vía para integrar la modelación matemática a la formación matemática del futuro ingeniero. El análisis de estos documentos, se presenta en el Capítulo 1, y permite mostrar la problemática de investigación, ¿cómo generar actividades de modelación matemática lo más “cercanas” a la práctica profesional? ¿cómo considerar las necesidades matemáticas básicas y avanzadas de la práctica en la formación? ¿Cómo reconocer las necesidades matemáticas en la práctica profesional? ¿La aparición de programas computacionales cada vez más potentes que encapulan el trabajo matemático obligan a una nueva enseñanza de las matemáticas? Para abordar esta problemática, se considera la formación del ingeniero en sistemas computacionales, del Instituto Tecnológico de Chilpancingo.

La Teoría Antropológica de lo Didáctico ofrece un modelo epistemológico para el análisis de la actividad humana, por lo que se elige para analizar tanto la actividad de modelación profesional del ingeniero como la de su formación. Consideramos principalmente las nociones de institución y praxeología, el modelo praxeológico extendido, que posibilita el análisis fino del componente práctico de la actividad en su contexto profesional; las Actividades de Estudio y de Investigación (AEI, en adelante) así como los momentos didácticos, que nos permitirán estructurar actividades didácticas de modelación u organizaciones didácticas. Estos elementos teóricos son presentados en el Capítulo 2.

Para diseñar de actividades de modelación para la formación de futuros ingenieros, que se encuentren sustentadas por el análisis de contextos profesionales, recurrimos a la propuesta metodológica inicialmente propuesta en Macías (2012). Esta propuesta nos

Revisar  
edacción

permite estructurar la investigación a través de las fases que en ella se plantean: *Elección de un contexto extra-matemático, Análisis praxeológico e identificación de un modelo matemático, Análisis del modelo en  $E(M)$  y Diseño de la actividad didáctica*. Estas fases son refinadas de acuerdo a las necesidades y a las condiciones propias de la presente investigación y se presentan en el Capítulo 3.

¿importa enfatizar que es  
"a profundidad?"

Debido a la importancia fundamental del *análisis praxeológico* del contexto como sustento del diseño de actividades, se presenta a **profundidad** en el Capítulo 4, así como el *análisis del modelo en  $E(M)$* , la arborescencia praxeológica resultante y un breve análisis de contextos en otras áreas que consideramos de potencial utilidad para el *diseño de la actividad didáctica*. Este diseño, que utiliza los momentos didácticos para estructurar una (AEI) se presenta y detalla en el Capítulo 5.

La implementación de la organización didáctica diseñada, se llevó a cabo en el Instituto Tecnológico de Chilpancingo, las condiciones y los resultados de esta implementación son presentados en el Capítulo 6 del presente trabajo.

Finalmente, se presentan en el Capítulo 7 las conclusiones de esta investigación.

¿Romo-Vázquez se refiere a Modelación? porque usa más bien modelización, es entendido.

# 1 LA MODELACIÓN EN LA FORMACIÓN Y LA PRÁCTICA DEL INGENIERO

## 1.1 Introducción

Los estudios sobre los procesos de enseñanza de las matemáticas en la formación de futuros profesionistas, y especialmente en la formación de ingenieros, se consideran investigaciones realizadas en Matemática Educativa, (Pollak, 1988; Howson, Kahane, Lauginie y Turckheim 1988 y Romo-Vázquez, 2014).

Los que estudian tales procesos han otorgado a la modelación matemática, un papel preponderante en la formación de dichos profesionistas. Para conocer cómo es que su incorporación al curriculum de matemáticas en tales formaciones se ha venido dando y modificando en efecto, estudiaremos algunos puntos de vista actuales sobre modelación: 1) la modelación matemática vista como interfaz entre la industria y el proceso formativo de profesionales, para destacar su función en dicho proceso; para tal estudio se consideran diferentes contribuciones del estudio ICMI 20, dedicado a este tema, (Li, 2013; Albertí, Amat, Busquier, Romero y Tejada, 2013; Geiger, 2013 y Ye, 2013), que se interesan por las necesidades que la industria plantea a la formación universitaria, en aras de formar profesionistas competentes para el área laboral. 2) También analizamos contribuciones hechas al estudio ICMI 14, dedicado a la modelación y las aplicaciones matemáticas (Niss, Blum y Galbraith, 2007; Henning y Keune, 2007; Artaud, 2007; Wake, 2007), quienes describen el proceso de modelación matemática, las competencias que debe desarrollar el estudiante, algunas consideraciones para la implementación de la modelación en el aula, la utilidad de analizar los procesos de modelación desarrollados en el área laboral, sus características y cómo puede esto ser considerado para y en el aula. 3) Para complementar este análisis, se considera el trabajo de Bisell y Dillon (2000), en el que se estudian las características de las actividades de modelación que realizan los ingenieros, así como las diferencias que existen con respecto a otras perspectivas de modelación.

## 1.2 Matemáticas en la formación de futuros ingenieros y su evolución

El papel de la matemática en la formación de ingenieros se ha ido modificando a lo largo de la historia. Esto puede reconocerse a través del análisis **delos** primeros modelos-propuestos por *Monge*, *Laplace* y *Le Verrier*, respectivamente, entre 1794 y 1860- de formación de la Escuela Politécnica Francesa, una de las primeras escuelas de ingenieros (Romo-Vázquez, 2009). El ideal enciclopedista del primer modelo de *Monge*, que concebía una posible alianza entre conocimientos teóricos y aplicaciones, fue sustituido por otro donde el desarrollo de la teoría pura era privilegiado (modelo de *Laplace*) y posteriormente por uno distinto donde las aplicaciones ocupaban el el lugar preponderante (modelo de *Le Verrier*). Los rasgos principales de estos modelos, muy brevemente descritos, muestran que concebir un equilibrio entre teoría matemática y aplicaciones, resulta sumamente complejo.

Siguiendo con la historia, pero un siglo más tarde, consideramos el estudio ICMI 3 realizado por la Comisión Internacional de la Enseñanza de las Matemáticas (ICMI, por sus siglas en inglés), titulado “*Mathematics as a service subject*”, el cual busca posicionar a la matemática como disciplina de servicio. En la presentación y primera contribución del estudio, los editores Howson, Kahane, Lauginie y Turckheim (1988), desarrollan ampliamente lo que “debe” entenderse por disciplina de servicio y muestran que reconocer la utilidad de la matemática, lejos de subestimar el valor de esta disciplina lo enriquece. Para ello, recurren a ejemplos de dos premios Nobel de química y de economía. El trabajo premiado en el área de química es el de la determinación de estructuras cristalinas, que involucra un método basado en análisis de Fourier y de probabilidad. Aunque no se precisa, podemos suponer que dentro de éste hay una fuerte presencia de la modelación matemática. Estos ejemplos y otros elementos sirven de base a los editores para cuestionar el rol de la enseñanza de las matemáticas en la formación de futuros profesionistas no matemáticos, arquitectos, doctores, administradores, etc. Cuestionamiento que puede ser abordado, reconociendo en primer lugar las necesidades matemáticas de cada formación y luego tratando de atenderlas. Esto supone, la creación de nuevos planes de estudio y una novedosa configuración de las asignaturas de matemáticas. Para la elección de los contenidos, se proponen dos criterios:

- Elegir los tópicos que uno supone como los más útiles para la vida laboral del estudiante.
- Enseñar lo que necesitan inmediatamente para el aprendizaje de la disciplina principal.

De éstos, el segundo criterio fue demandado fuertemente, no solo por profesores, sino por estudiantes, que buscaban encontrar coherencia entre la matemática que se les enseñaba, y aquellos cursos en los que la necesitaban. Los editores resaltan que esto conlleva un problema fundamental, ciertos conceptos no pueden ser estudiados tan pronto como se solicitan. Por ejemplo, el químico o el físico solicitan estudiar funciones de varias variables, antes de que el matemático pueda impartirlas. Se afirma entonces, que es el primer criterio el que debe ser considerado. Esto implica, tener en cuenta que en un mundo en el que la ciencia avanza a pasos agigantados, identificar las necesidades del profesionista se convierte en un proceso continuo. Aun así, los editores sostienen que la necesidad de cumplir con este criterio, es mucho mayor que la del segundo, siendo necesario trabajar en conjunto con los especialistas de las disciplinas correspondientes. Asimismo, se resalta que sin importar el tema enseñado, debe dedicarse un tiempo para mostrar a los estudiantes la belleza de las matemáticas.

Una de las aportaciones más importante, a este estudio, es la contribución de Pollak (1988), investigador que durante 35 años se desempeñó como matemático en la industria, laborando en los Laboratorios Telefónicos Bell, empresa particular debido a que realiza actividades de desarrollo científico. Por lo tanto, la empresa tenía la necesidad de que sus ingenieros poseyeran ciertos conocimientos, que consideraban necesarios. Para garantizarlo, se fundó un programa educativo de formación para los

ingenieros contratados, impartido a aquellos que solo poseyeran título universitario. Dentro de este programa, con una duración de tres años, se les enseñaba álgebra lineal, variables complejas, series de Fourier, transformada de Laplace, teoría de probabilidad, estadística, física de semi-conductores y gran cantidad de temas, que en la época no formaban parte de la formación universitaria regular de los ingenieros eléctricos y mecánicos. Pero, que para la empresa y las actividades que dentro de ella se desarrollaban, resultaban necesarias. La empresa decide también contratar matemáticos, que pudieran convertirse en informáticos expertos, y también se encarga de formarlos. Sin embargo, para esto se consideró de vital importancia que tales estudiantes fueran capaces de aplicar la matemática al mundo real, considerando que era imposible ser un estudiante de ingeniería destacado, sin la capacidad o instinto para abordar problemas reales. Aunque, si era posible ser un excelente estudiante de matemáticas y no tener experiencia, o talento para aplicar las matemáticas a otros campos. Por lo tanto resultó muy importante buscar en ellos cualquier indicio de la habilidad de aplicar la matemática.

Su aportación resulta de gran importancia debido a que permite observar que las necesidades de la industria, del ambiente laboral, deben ser tomadas en cuenta para la formación de universitarios competentes. A su vez, deja observar la perspectiva del empleador, al que más que el conocimiento por sí mismo, le interesa que el profesionista sea capaz de aplicar la matemática al mundo real. Se requiere que sean capaces de analizar y comprender el problema, de imaginar el posible tipo de respuesta, así como la capacidad de obtenerla por cualquier medio (cálculo mental, calculadora, lápiz y papel, computadora, etc.). Más importante aún, se resalta la necesidad de comprender la matemática, evitando el estereotipo del ingeniero que solo memoriza y aplica fórmulas conocidas, sin saber por qué y cómo funcionan: “La comprensión es vital para aplicar la matemática correctamente”. Finalmente, se espera que los empleados estén listos para enfrentar situaciones en las que el medio de solución no sea claro, y sea necesario cuestionarse, ¿cuál podría ser el método para resolverla? El autor expresa que los cursos de construcción de modelos son particularmente eficientes para preparar al estudiante para este tipo de actividad y que para lograr esto se debe pensar no sólo en el conocimiento que se pretende adquieran los estudiantes, sino también en las formas de pensamiento asociadas a estos conocimientos. Esto es plasmado de la siguiente manera:

Antes que todo, necesitamos tener conocimiento del hecho de que el pensamiento matemático, el pensamiento analítico, estructural, cuantitativo, sistemático, puede ser aplicado al mundo real y generar observaciones de gran valor; en otros términos, que la modelización matemática es posible y puede ser eficaz. (Pollak, 1988, p. 32)

Finalmente, el autor menciona que la matemática universitaria tiene la responsabilidad de preparar al estudiante para su profesión, resaltando que los cambios y avances de la ciencia y la tecnología provocan que las aplicaciones de la matemática cambien también, por lo que se debe comprender que la enseñanza de la matemática debe seguir y atender estos cambios. Derivándose de esto, que las instituciones educativas tendrán

siempre la responsabilidad de seguir los avances del campo laboral, identificar las nuevas necesidades que surgen, para considerarlas en la formación matemática del futuro profesionalista.

Más recientemente se ha generado un nuevo estudio ICMI, el estudio ICMI 20, sobre el papel de la matemática en la industria, y la manera en que esta relación puede ser útil para la enseñanza. Su análisis, nos permitirá observar la evolución del rol asignado a la modelación matemática.

### 1.3 El papel de la modelación matemática en la industria

Para analizar el papel de la modelación matemática en la industria se consideró el estudio ICMI 20, cuyo título es, Interfaces educativas entre la matemática y la industria. Este estudio fue elaborado por la Comisión Internacional de la Enseñanza (ICMI) en colaboración con el Consejo Internacional para la Matemática Industrial y Aplicada ICIAM (por sus siglas en inglés), teniendo como propósito mostrar la relación entre educación matemática y utilización de la matemática en la industria. Se dirige a un público amplio, matemáticos en general, matemáticos aplicados en particular, autoridades responsables del desarrollo educativo y económico de los países participantes así como educadores. Se hace énfasis en que el estudio pretende mostrar los desafíos de una enseñanza que debe dotar a los estudiantes de herramientas matemáticas que les permitan enfrentar problemáticas “reales” y particularmente profesionales. En especial, de aquellos que trabajarán como “matemáticos en la industria”, así como “analistas”, “científicos”, “especialistas”, “ingenieros”, así como de quienes se encuentren en trabajos que no tienen relación directa con las matemáticas. Para lograrlo, se plantea que la modelación matemática sea un interfaz entre la matemática y la industria (aplicaciones), permitiendo llevar las herramientas matemáticas al terreno de lo práctico. Esto, incluso si la formación es predominantemente matemática o no, ya que se trata de una competencia necesaria para el campo laboral.

ICIAM en el  
paréntesis ...  
¿cierto?

Para este análisis, se consideraron las contribuciones propuestas por Li, Albertí, Amat, Busquier, Romero y Tejada; Geiger y de Ye, pues en ellas se discute el papel de la modelación matemática en el área laboral, particularmente en la ingeniería. En ellas se muestra que las necesidades de la industria han motivado la preocupación sobre la integración de la modelación en la enseñanza del futuro profesionalista así como algunas de las dificultades que deben ser superadas en esta integración.

APA Style

#### 1.3.1 Roles de la modelación matemática entre matemáticas e industria

En la contribución hecha por Li (2013) al estudio ICMI se señala que la modelación matemática es la interfaz más importante entre las matemáticas y la industria, siendo el camino “inevitable” que lleva hacia la aplicación de las matemáticas. Éste es considerado como el primero de tres roles que la modelación juega, el segundo es relacionar temas con sus aplicaciones y el tercero impulsar el desarrollo de las matemáticas aplicadas modernas. A continuación se describen estos tres roles.

Control de "viudas"

1) *Modelación, camino inevitable a las aplicaciones en la industria*

Para ilustrar el primer rol, la modelación es definida como un proceso en el que el contexto es visto como un ente general. Aunque se precisan algunos contextos, ingeniería, economía, finanzas y sociología, todos son vistos como entes en los que existen problemas que pueden ser enfrentados o resueltos a partir de un tratamiento matemático. Lo que supone que todo problema es modelable matemáticamente:

Para resolver un problema práctico a través de una aproximación matemática, sin importar si viene del campo de la ingeniería, economía, finanzas o sociología, necesitamos establecer una conexión entre el problema y la matemática. Para ser exactos, tenemos que transformar en primer lugar el problema práctico en un problema matemático relevante, entonces realizar el análisis y cálculos del problema matemático y finalmente aplicar la solución a la práctica, para comprobar si es que puede resolver el problema planteado. Este proceso, especialmente el primer paso, es llamado modelación matemática, eso es, configurar un modelo matemático para el problema práctico explorado. (Li, 2013, p.51)

La fase experimental de este proceso de modelación es considerada como igual de importante que su fase inicial, donde se establece la relación entre el problema y el modelo matemático. Esto debido a que es la fase que determina si el problema fue resuelto exitosamente o no y si debe realizarse nuevamente el proceso.

## 2) Relacionar temas y aplicaciones

El segundo rol de la modelación, lo hemos traducido del inglés: *Second, mathematical modeling is the key to related subjects and application* como, relacionar temas y aplicaciones, sin embargo nos parece que esta traducción no ilustra el rol que el autor propone. Pues para su ejemplificación presenta seis ejemplos que muestran cómo el estudio avanzado de ciertos fenómenos, permitió sentar las bases de modelos matemáticos “teóricos”, muy importantes en el desarrollo de la ciencia. Inicialmente, considera que la geometría euclidiana constituye un modelo matemático básico, que actualmente, permite comprender el espacio tridimensional ( $R^3$ ) en el que se encuentra el mundo real. El segundo ejemplo es el de las tres leyes del movimiento planetario propuestas por Kepler, que resultan del profundo trabajo de observación astronómica de Tycho y que posteriormente fueron “estrictamente” probadas por Newton, utilizando los principios de la mecánica y de la ley de gravitación universal. Los siguientes cuatro ejemplos son más recientes y se relacionan directamente con la ingeniería, como lo ilustra la siguiente cita:

Algunas ecuaciones diferenciales importantes en mecánica y física, como la ecuación de Maxwell en electrodinámica, la ecuación de Navier-Stokes y la ecuación de Euler en hidrodinámica, la ecuación de Schrödinger en mecánica cuántica, etc., son todos modelos matemáticos que reflejan la esencia de estos temas y crean los marcos teóricos fundamentales de las disciplinas relacionadas. Al mismo tiempo contienen todos los resultados importantes y potencial de

aplicación, estos modelos juegan sin duda un papel esencial y ocupan posiciones clave en sus campos respectivos. (Li, 2013, p. 52)

Nótese que este segundo rol consiste en relacionar al mundo o algunos de sus fenómenos con modelos matemáticos teóricos. Los ejemplos elegidos por el autor muestran modelos bastante generales, estables y fundamentales, algunos sustentan desarrollos matemáticos disciplinares, como es el caso de la geometría y otros el avance de disciplinas de especialidad, como es el de la mecánica o la hidráulica.

### 3) *Avances de las matemáticas aplicadas*

El tercer y último rol está dado por el avance de las matemáticas aplicadas debido a la evolución científica sin precedente, que ha ampliado la cantidad de áreas en las que la matemática encuentra aplicación. Esto va desde áreas “tradicionales” como la mecánica y la física, a áreas como son: la química, la biología, la economía, las finanzas, las ciencias de la información, la ciencia de materiales, la ciencia ambiental, la ciencia energética e incluso las ciencias sociales. Se explica que este proceso que se mantiene en desarrollo, requiere de la modelación matemática, que debe en muchos casos superar grandes dificultades y retos para avanzar sobre campos nuevos. Esto brindará más oportunidades tanto para el avance de las matemáticas como de las matemáticas aplicadas.

Dado que las leyes de muchos nuevos campos aún están siendo exploradas, la construcción de modelos matemáticos relevantes no es fácil de conseguir, o aún enfrenta dificultades sustanciales que representan un gran desafío para nosotros hoy en día. Por lo tanto, la importancia de la modelación matemática es mucho más eminente que antes, y se ha convertido en una parte vital de la matemática aplicada moderna, la cual puede y proveerá más oportunidades y prospectos más amplios para la matemática aplicada así como para la disciplina de la ciencia matemática al completo. (Li, 2013, p. 52)

Finalmente, el autor plantea que la educación matemática debería buscar integrarse con otras ciencias y en un sentido más amplio, con el “mundo externo”, para permitir al estudiante alcanzar una verdadera comprensión del desarrollo de la matemática y de su utilidad para solucionar problemas “reales”.

Considerando las características y funciones de la matemática, la educación matemática no debería ser separada de las otras ciencias y el mundo externo, limitándose a sí misma y girando solamente alrededor de sus conceptos matemáticos, métodos y teorías. De hacerse esto será un obstáculo en los estudiantes, para la comprensión sobre el comienzo y el subsecuente desarrollo de los conceptos matemáticos, métodos y teorías, así como para que los estudiantes se inspiren a utilizar las herramientas matemáticas voluntaria y activamente, para resolver todo tipo de problemas del mundo real, y para la mejora de los logros matemáticos del estudiante. (Li, 2013, p. 53)

conviene revisar "Enseñar" procesos. UN proceso puede asociarse a ACCIONES QUE SE LLEVAN A CABO, ETAPAS, ACTOS...

Conviene REVISAR "enseñar" conocimientos ... El conocimiento se adquiere, se construye, se comparte, ... y con base en ello, se desarrollan habilidades y capacidades...

Existe, por tanto, una gran necesidad de enseñar conocimientos matemáticos, pero sobre todo de enseñar los procesos creativos de la matemática. Asimismo, se recalca fuertemente la importancia de comprender la “esencia espiritual” y los “patrones de pensamiento” de la matemática, como medio para comprender el núcleo de la disciplina. Y también la creación de una cultura matemática, enriquecida y animada por actividades como concursos de modelación, que funcione como medio de fortalecimiento. Concluyendo que ambas cosas ayudarán en la transformación de la matemática, en una herramienta a disposición del estudiante.

Para resumir, la educación matemática no debería enfocarse a la introducción de conceptos matemáticos, métodos y conclusiones dados a cucharadas al estudiante, sino hacer un mayor énfasis en permitir que el estudiante comprenda la esencia espiritual y patrones de pensamiento de la matemática, que domine el núcleo de la disciplina, y que voluntariamente edifique una cultura matemática, de manera que haga de las matemáticas un arma versátil que pueda ser usada toda su vida. (Li, 2013, p. 54)

El autor insiste en que la matemática no puede verse como un conjunto de contenidos sino que es necesario reconocer su “esencia espiritual” y el pensamiento matemático, como lo señala Pollak (1988) en el ICMI 3. Sin embargo, esta propuesta deberá, desde nuestro punto de vista, acompañarse de cambios en la organización de la enseñanza, que puedan favorecer este paradigma en las aulas. Este cambio en la formación podría generar, como lo demanda la industria, egresados competentes, hábiles y capaces de resolver problemas “reales”, relacionando las matemáticas con otras disciplinas y con la “realidad”.

### **1.3.2 Formación para la profesión, un ejemplo en la formación española de ingenieros**

La contribución de Albertí, et al (2013), aborda el interés de las autoridades educativas españolas sobre la relación existente entre la actividad industrial y la enseñanza matemática. Enfocándose en la relación entre la matemática y la actividad profesional, particularmente de los ingenieros. Para ello, en primer lugar presentan un panorama general sobre la matemática y la manera en que el avance tecnológico ha modificado sus contextos de uso, lo que permite cuestionarse sobre las matemáticas que se utilizan actualmente en el área de trabajo.

Siempre se ha dicho que las matemáticas son usadas casi en todas partes, pero los contextos de su uso y la manera en que es utilizada podrían haber cambiado en los últimos años. Lo que hace varias décadas tenía que resolverse con lápiz y papel, fue resuelto más tarde con calculadora y puede ser resuelto ahora con computadora. El software a menudo parece estar incorporado a la maquinaria. ¿Qué clase de matemática está siendo utilizada hoy en día en el campo laboral? (Albertí, et al, 2013, p.185)

Si la matemática que se utiliza en el campo laboral ha cambiado, resulta natural para los autores preguntarse ¿es la matemática enseñada en la escuela la que se utiliza en la

práctica profesional? Es decir, ¿la manera en que los profesionales resuelven un problema es similar a la manera en que la resuelve un educador matemático? Cuestionando entonces si la enseñanza está formando realmente estudiantes laboralmente competentes.

Los autores indican que la matemática que se enseña usualmente, no tiene similitudes con la matemática que los estudiantes utilizarán al enfrentar problemas pertenecientes a su ambiente laboral. Por lo tanto, es posible asumir que los estudiantes no están siendo preparados adecuadamente para su vida profesional. Debido a esta situación, se cuestiona el papel del profesor y su capacidad de formar estudiantes matemáticamente competentes. “¿Cómo puede un profesor promover competencias matemáticas coherentes con un trabajo no-académico o ambientes industriales, si su educación matemática se desarrolló fuera de estos ambientes?” (Albertí, et al, 2013, p.187)

La formación académica del profesor se desarrolla de la manera “tradicional” en la que se privilegia el estudio de conceptos. En esta formación no se abordan contextos en los que la matemática pueda o sea utilizada para resolver problemas “reales” y tampoco la manera en que el conocimiento matemático es empleado dentro de ámbitos laborales. Esto causa que el profesor no considere estos aspectos al desarrollar sus actividades en el aula, lo que **afecta finalmente al estudiante.**

Se remarca entonces la necesidad urgente de educar **estudiantes matemáticamente competentes, motivada principalmente por el sector industrial y su necesidad de trabajadores suficientemente capacitados.** Se requiere entonces que los estudiantes no solo aprendan matemáticas, sino que cobren conciencia de que tal conocimiento será de gran utilidad para su futura vida laboral. “Los estudiantes deben aprender matemáticas y comprender que el conocimiento matemático que han adquirido es verdaderamente útil para responder o crear respuestas a los problemas que enfrentarán en sus vidas, así como en su actividad profesional, una vez concretada su educación secundaria.” (Palmer, Amat, Busquier, Romero, Tejada, 2013, p.187)

La actividad profesional del ingeniero, se desenvuelve característicamente entre el terreno de la matemática y el de las disciplinas que requieren de ella para resolver problemáticas pertenecientes a su área de interés particular. Es decir, que las labores del ingeniero tienen como parte de sus funciones **conectar la matemática con sus aplicaciones industriales.** Dado que la modelación ya ha sido reconocida como puente entre ambas, podemos establecer entonces que la actividad de modelación matemática tiene un carácter central para la práctica profesional del ingeniero. Es por esto que, durante su formación, resulta necesario que el ingeniero desarrolle las habilidades necesarias para llevar a cabo actividades de modelación exitosamente. “La conexión entre la matemática pura y sus aplicaciones en el progreso industrial es una tarea crucial. En particular, el ingeniero tiene que obtener las competencias necesarias para desarrollar esta colaboración.” (Palmer, Amat, Busquier, Romero, 2013, Tejada, p.190)

Conviene distinguir que tipo de ingeniero es el que marco es ese tipo de ingenieros...

tradicional DE BIEBET QUE ESTUDIO DE LA PROFESIÓN y por ende del perfil de egreso, PORQUE NO TODOS SE ORIENTAN HACIA LA INDUSTRIA ... ¿CÓMO SE DEBEN LAS REFERENCIAS? porque una cuestión que se nota en la redacción es que en algunos momentos parecen argumentos de Julián y otros los entiendo...

Totalmente de acuerdo... un proceso se comprende o se analiza o se explica cómo se lleva a cabo, más que "enseñarlo" o puede ser que se refieran a "teoría" o "evidencia" cómo se realiza un proceso? Aludo a ello por la interrogante que coloque al inicio de la página 15

Esta preocupación sobre la formación de futuros profesionistas, adecuadamente preparados para su actividad laboral, es retomada en otras contribuciones, por ejemplo, la realizada por Qixiao Ye. Específicamente, la aportación de Ye integra la modelación matemática a un curso de Cálculo, pues considera que al abordar actividades de modelación en el aula, el estudiante puede comprender de forma más efectiva el proceso. En general, su interés mayor es encontrar respuesta a la la cuestión: ¿cómo es que científicos e ingenieros resuelven problemas “reales” a través de la modelación matemática? Para lo cual considera que es necesario estudiar en primer lugar, cómo se lleva a cabo la actividad matemática en ambientes laborales y académicos; para luego llevar tales contextos al aula, a través del diseño de actividades de modelación. En su contribución se detecta un objetivo didáctico fundamental derivado de la necesidad de progreso laboral, hacer que el estudiante sea matemáticamente competente, el cual tiene su origen en el principio: *a mejor preparación, mayores son las posibilidades de ser exitoso laboralmente.*

La experiencia nos dice que mientras más sólida sea la base matemática que un estudiante universitario tenga, mayor será su probabilidad de éxito, sin importar en que clase de empleo de desenvuelva. La modelación matemática es un “puente” para el uso de la matemática en la resolución de problemas del mundo real, y a menudo la modelación tiene una importancia crítica en la resolución de ciertos problemas. (Ye, 2013, p. 385)

Esta necesidad de egresados matemáticamente competentes, con miras a un mayor éxito en el campo laboral, aparece también en la contribución hecha por Vince Geiger, quien señala que ya se ha evidenciado la existencia de una gran demanda de trabajadores matemáticamente capaces, misma que continúa en aumento. Esta demanda surge como respuesta al aumento en la complejidad de la actividad matemática que se realiza dentro de las industrias, la cual sigue progresando constantemente.

Las investigaciones sobre el uso de la matemática en el área de trabajo e industria han demostrado una demanda en aumento de una fuerza de trabajo matemáticamente competente. Además de esto, las demandas matemáticas de las actividades que los trabajadores realizan en un amplio rango de industrias se están volviendo cada vez más complejas. (Geiger, 2013, p. 271)

Esta tendencia de crecimiento y avance industrial ha puesto en evidencia que el conocimiento matemático por sí mismo no resulta suficiente para que el estudiante opere adecuadamente en el ambiente laboral. Resulta necesario desarrollar también habilidades externas a la matemática, tales como, el pensamiento adaptable, disposición a aprender nuevas maneras de resolver problemas, la capacidad de utilizar herramientas tecnológicas en la resolución de problemas, etc. Estas habilidades no solo son características de los procesos de modelación, sino que para el ámbito laboral tienen una magnitud de importancia en el mismo nivel que el conocimiento matemático.

...la capacidad de pensar de manera adaptable, una disposición a seguir aprendiendo nuevas maneras de resolver problemas al tiempo que estos surgen, y

la capacidad de adoptar el uso de herramientas matemáticas, son tan importantes como el tipo de conocimiento matemático tradicionalmente impartido en las escuelas. (Geiger, 2013, p. 272)

Resulta interesante observar que, a diferencia de los primeros modelos de formación Teoría-Aplicación (modelo de Laplace) y de uno más reciente Teoría-Modelación matemática (evidenciado en Pollak, 1988), en estas contribuciones se identifica una nueva perspectiva en la que la modelación es el puente entre teoría y aplicaciones.



A su vez, es importante notar la creciente preocupación por integrar la modelación matemática en la formación de futuros profesionales. Existe un énfasis particular sobre la ingeniería, debido a que se trata de una disciplina que ha ejercido el papel de puente entre la matemática y otras disciplinas usuarias. Sin embargo, actualmente la matemática no solo ha encontrado ya un lugar en múltiples campos, sino que continúa abriéndose camino de manera constante en nuevas disciplinas. Es por ello, que la modelación matemática surge como una necesidad en las formaciones de todo tipo de profesionales.

El análisis muestra que implementar efectivamente la modelación matemática en las aulas conlleva ciertas dificultades: el tiempo limitado por el currículo, la familiaridad del estudiante con la forma más tradicional de enseñanza teórica, el aislamiento que ha tenido la formación matemática con respecto a los contextos, así como el desconocimiento por parte del profesor, de la matemática que se utiliza en la práctica profesional y las características específicas de su uso laboral en áreas particulares. Se considera así, que la forma de llevar la modelación al aula requiere, en primer lugar, considerar la manera en que la modelación se lleva a cabo en los contextos laborales. Para el interés de nuestro trabajo, eso significa concentrarnos en los procesos de modelación de los ingenieros, y en específico, del ingeniero en sistemas computacionales.

También se requiere que los profesores adquieran competencias para el desarrollo de un currículo basado en la modelación matemática, incluso que cuando los profesores demandan determinado campo laboral en que se insertarán los egresados de las carreras (según perfiles de egreso), y que se les pida que los resuelvan.

#### 1.4 Modelación matemática en la formación

El análisis para determinar el lugar de la modelación matemática en el proceso de formación de ingenieros (en computación) nos lleva a considerar algunas contribuciones que permiten reconocer las características generales de la actividad de modelación, así como las condiciones necesarias para su uso sistemático en la enseñanza, entre las, tres contribuciones del estudio ICMI 14, dedicado a la modelación y las aplicaciones en matemática educativa y publicado en el 2007, hechas por Henning y Keune, Artaud y Wake. De igual forma, considerar contribuciones que tocan aspectos de nuestro interés particular, como la de Bissell y Dillon (2000), la cual permite reconocer los elementos principales de la relación entre formación práctica profesional de ingenieros.

### 1.4.1 Proceso de modelación, competencia de modelación y condiciones para su implementación en el aula

En este estudio se analiza la modelación y las aplicaciones de matemáticas, como la relación existente entre la matemática y el mundo extra-matemático. Para esto, en primer lugar se define el proceso de modelación matemática desde el punto de vista de la matemática, Niss, Blum y Galbraith (2007) presentan la siguiente esquematización del proceso:

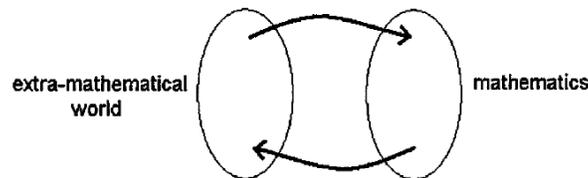


Figura 1. Proceso de modelación matemática (Niss, Blum y Galbraith, 2007)

Los autores explican que el modelo matemático consiste tanto en el dominio de interés extra-matemático  $D$ , como de algún dominio matemático  $M$  y el correspondiente mapeo entre ambos. Es decir:

El término modelación se refiere al proceso entero y todo lo que involucra: la estructuración de  $D$ , la elección del dominio matemático  $M$  adecuado, el mapeo de  $D$  a  $M$  (matematización), el trabajo matemático en  $M$ , la interpretación y la evaluación de los resultados del trabajo matemático con respecto a  $D$ , y la repetición del ciclo varias veces de ser necesario. (Niss, Blum y Galbraith, 2007)

Al **observar** esta definición de modelación y **considerarla como válida**, cabe preguntarse ¿qué necesita el estudiante para poder llevar a cabo este **proceso o actividad** de modelación?

Henning y Keune (2007), proponen tres niveles de competencias de modelación, necesarias para que el estudiante pueda desarrollar actividades de modelación, para definir competencia de modelación se recurre a Blum:

La habilidad de estructurar, matematizar, interpretar y resolver problemas, y, adicionalmente, la habilidad de trabajar con modelos matemáticos, validar los modelos, analizarlos críticamente y valorarlos así como sus resultados, comunicar los modelos, observar y ajustar por su cuenta el proceso de modelación. (Blum, citado en Henning y Keune, 2007)

A partir de esta definición, se establecen los siguientes niveles:

Nivel 1: Reconocimiento y comprensión de la modelación

Nivel 2: Modelación independiente

Nivel 3: Meta-reflexión sobre la modelación

Las definiciones MÁS QUE OBSERVARSE ...se analizan, se comprenden. porque VER u OBSERVAR es insuficiente ya que "no es a ojo"

¿Qué criterios se usaron para su validez? y por otro lado: qué fin tiene su validez en este trabajo? en todo caso uno puede temático en M, la interpretación

2007) y proceso son cosas distintas y aquí la consideran como lo mismo.. REVISAR

Siempre que el profesor sea capaz de comprender este proceso de modelación y favorezca que el estudiante transite por el.

Se explica que el nivel 1, es caracterizado por las habilidades para reconocer y describir el **proceso de modelación**. El nivel 2, comprende las habilidades para analizar y estructurar problemas, abstraer cantidades, adoptar diferentes perspectivas, proponer modelos matemáticos, trabajar con ellos, interpretar resultados y validarlos, así como validar el proceso entero. En este nivel, se dice que el estudiante será capaz de resolver un problema por cuenta propia, así como adaptar el modelo a cambios en el mismo. Para el nivel 3, los autores caracterizan las habilidades para juzgar y reconocer relaciones. El concepto de modelación está bien comprendido, y el estudiante identifica la importancia de los modelos en diferentes áreas, así como su uso en la ciencia. Lo que implica, que incluso los modelos “terminados” o “completos”, puedan ser analizados.

Estos niveles de competencias pretenden describir las habilidades que el estudiante obtiene al llevar a cabo procesos de modelación, ofreciendo una medida que permita evaluar su grado de desarrollo o aprendizaje. Aunque, como notan los autores, existen otras perspectivas de competencias de modelación, que buscan, de la misma manera, establecer una posible medida para valorar el grado de avance del estudiante. Existen perspectivas como la de PISA (*Programme for International Student Assessment*), que utiliza la noción de “alfabetismo matemático”, el cual define como:

La capacidad del individuo de identificar y comprender el rol que juegan las matemáticas en el mundo, para hacer juicios matemáticos bien fundamentados, e interactuar con las matemáticas, de formas que cumplan con las necesidades del individuo, como un ciudadano reflexivo, interesado y constructivo. (Organization for Economic Co-operation and Development (OECD), citado en Henning y Keune, 2007).

Los autores consideran que esta definición conecta la matemática con el tratamiento de problemáticas “realistas”, por lo que las competencias de modelación que ellos proponen, pueden ser vistas como parte del alfabetismo matemático. Asimismo, reconocen la existencia de otras definiciones de competencias de modelación, basadas en otras perspectivas. Esto causará que la elección y uso de una definición particular, dependa de la perspectiva de modelación que se utilice.

Una problemática retomada por el estudio, es que a pesar de la existencia de trabajo en matemática educativa enfocado a las aplicaciones y modelación, éstos no tienen grandes repercusiones en el aula. Artaud (2007), cuya contribución se enfoca en las condiciones que considera necesarias para que la modelación pueda existir en el aula, también lo señala. Su investigación, se desarrolla en el marco de la Teoría Antropológica de lo Didáctico, considerando dos aspectos principales:

- Lo que se enseña y aprende (visto en términos de praxeologías matemáticas)
- Las actividades de aprendizaje y enseñanza (vistas en términos de praxeologías didácticas)

Una praxeología, como lo detallamos en el capítulo 2 de esta tesis, es una unidad de análisis de la actividad humana, que tiene cuatro componentes, tipo de tarea  $T$ , técnica  $\tau$ , tecnología  $\theta$  y teoría  $\Theta$ . El tipo de tarea define lo que se hace, la técnica cómo se hace, la tecnología produce, explica y justifica la técnica, la teoría tiene el mismo rol, pero a un nivel más general. Las praxeologías matemáticas son actividades matemáticas y las didácticas, describen el proceso de estudio que lleva a la construcción o reconstrucción de una praxeología matemática.

En esta contribución, el proceso de modelación es visto como un medio para que las praxeologías matemáticas emerjan, a través de los seis “*momentos de estudio*”, de los cuales el estudiante debe pasar al menos por un “*primer encuentro*”, un “*momento exploratorio*” del problema o tarea, “*elabora una técnica*” que resuelva la tarea, “*construir un entorno tecnológico-teórico*” que fundamente y justifique la técnica, y que “*trabaje la técnica*”, es decir, que analice los resultados y refine la técnica en la medida de lo posible, de manera que la haga más efectiva. Estos cinco momentos aunados al momento de institucionalización, dentro de la teoría, son los que permiten la construcción o reconstrucción de una praxeología, en este caso sería de modelación matemática.

Según la autora, al utilizar situaciones reales para lograr esta emergencia de las praxeologías matemáticas, es vital que no se utilicen estos contextos como fachada o “*vestidura*”, de la que el estudiante tenga que desprenderse lo más pronto posible. Este contexto debe ser parte del proceso de estudio y al momento que el profesor institucionalice la praxeología matemática estudiada, no debe dejar de lado el contexto en el cual fue trabajada y desarrollada, sin importar que éste sea completamente no-matemático. La autora menciona que en ocasiones es difícil reconocer que la matemática es relevante al analizar situaciones del mundo real, por ello, estos contextos no-matemáticos resultarán de gran importancia para que el estudiante mantenga presente el papel que la matemática tiene en situaciones reales, como se ilustra en la siguiente cita:

En este aspecto, hablar de aplicaciones de la matemática contribuye a oscurecer la modelación que tiene que hacerse: esto lleva, por ejemplo, a pensar que el enseñar al estudiante una respuesta matemática, y resolver problemas en los cuales dicha respuesta es utilizada o juega un papel importante, será suficiente porque entonces existirá en los estudiantes una habilidad para transferir tal conocimiento. (Artaud, 2007, p. 376)

Debido al fuerte énfasis sobre el contexto y su importancia para desarrollar actividades de modelación, surge la cuestión, ¿dónde y cómo encontramos contextos útiles para llevar a cabo procesos de modelación? Consideramos, asimismo que los contextos tienen la función de ilustrar el papel y la relevancia que la matemática tiene en el mundo laboral. Por lo tanto, nos cuestionamos, ¿cómo podemos identificar y estudiar contextos que tengan relación tanto con la formación como con la práctica laboral del futuro ingeniero?

Las respuestas NO SE ENSEÑAN ...  
Como lectora me preocupa que  
poco se reflexione sobre los  
argumentos de otros autores, ya que  
se asumen como verdades  
presentes en este asunto de ENSEÑAR UNA  
RESPUESTA MATEMÁTICA.

Esta es una de las funciones.... evitar  
plantearlo como el sólo fuese la única  
función... incluso es una postura.

### **1.4.2 La modelación matemática en el área laboral y su posible impacto sobre el aula**

La contribución de Wake (2007), estudia la actividad laboral, vista desde la perspectiva de la modelación matemática, argumentando que, histórica y culturalmente, la matemática escolar parece preparar de manera insuficiente al estudiante para su actividad profesional. El autor observa que a menudo se ha estudiado el mundo laboral, con la intención de estudiar la matemática utilizada por los trabajadores e identificar si ésta puede ser parte del diseño curricular. Algunos estudios, buscaron identificar contenido matemático de acuerdo a su uso, la comprensión requerida para su uso, así como su necesidad. Estos trabajos han progresado hasta el punto en que los investigadores toman el rol de etnógrafos, de manera que puedan explorar adecuadamente el campo laboral. Sin embargo, identificar la matemática existente no implica identificar la modelación matemática y sus posibles implicaciones para la matemática escolar.

Sin embargo, a pesar de que los modelos matemáticos se vislumbran, más implícita que explícitamente, dentro de las actividades de los trabajadores. Se ha puesto, hasta ahora, poco énfasis en lo que podemos aprender sobre la investigación sobre modelación matemática en el área de trabajo y las implicaciones que podemos obtener para la enseñanza y el aprendizaje de las matemáticas en la escuela, ya que éstas no han tenido gran desarrollo. (Wake, 2007, p. 396)

Al analizar dos casos de estudio, en los que se investigó el proceso de modelación llevado a cabo en dos contextos laborales, se identificó que la actividad matemática del trabajador implica el uso de modelos “estándar” o “aceptados”, que pueden haber sido desarrollados por el mismo trabajador o simplemente adoptados. Además de esto, dicha actividad implica la constante interpretación de la información producida por el uso de estos modelos, sobre un enorme sustento contextual, conocimientos provenientes de la experiencia laboral, por ejemplo. A su vez, es posible observar que estos conocimientos permiten al trabajador un cierto grado de libertad en la manera en que emplean tales modelos, pudiendo modificar dicha manera, en caso de considerarlo necesario. Por lo que su actividad principal, será la interpretación de los resultados. Un punto importante es la observación del autor sobre la necesidad de una profunda comprensión del área de trabajo, para poder explicar y entender los modelos a utilizar.

Finalmente, Wake sugiere que <sup>cómo</sup> esta área requiere de mayor investigación, de manera que pueda explorarse adecuadamente **como** la actividad laboral resulta útil para las aulas de todos niveles, y cómo el currículo podría preparar de mejor manera al estudiante. Derivando de esto, la importancia que tiene no solo identificar la matemática del área de trabajo, sino de todo el contexto en que se desenvuelve y considerar que la actividad no concluye solamente con el desarrollo de un modelo que permita resolver un problema, sino que se complementa con su uso, su interpretación y su adaptación constante.

Lo anterior, exige saber **¿cómo se llevan a cabo los procesos de modelación en la práctica de los ingenieros?** Para conseguir una respuesta, enseguida se analiza el trabajo realizado al respecto de Bisell y Dillon, con el propósito de identificar las características de la modelación en ingeniería.

Esta pregunta puede tener más de una respuesta que puede ser tangencial a otras. Dependerá de un enfoque teórico, perfil del ingeniero, etc.... Y es que aparecen preguntas con respuestas que pareciera es la única manera de explicar o comprender una **ciencia**

### 1.4.3 Modelación matemática en Ingeniería

En el trabajo de Bissell y Dillon (2000) se aborda la modelación matemática tanto en la formación como en la práctica de ingenieros. Los autores señalan que tradicionalmente se reconocen dos formas de representar la actividad de modelación: el ciclo rígido (Fig. 2) y el ciclo flexible (Fig. 3), los cuales se ilustran enseguida:

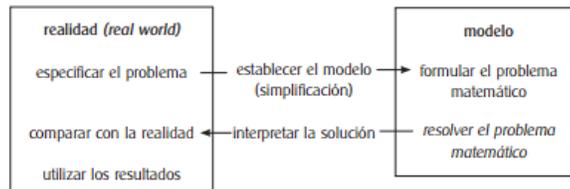


Figura 2. Ciclo Rígido

El ciclo rígido (Fig. 2) se caracteriza por 4 etapas:

- Describir el sistema físico (*physical modelling*)
- Describir el sistema matemático (*model construction*)
- Analizar la descripción matemática (*model solution*)
- Interpretar y sacar provecho de esta descripción (*system design*)

Bisell y Dillon destacan que aun cuando el ciclo luce “engañosamente simple”, presenta varias dificultades al momento de ponerlo en práctica:

- ¿Cómo se especifica el problema?
- ¿Cómo se decide la simplificación requerida?
- ¿Qué se considera como “resolver el problema matemático”?
- ¿Cómo se validan los resultados?

Este ciclo, además hace suponer que siempre es posible modelar la realidad, como los autores lo señalan. Esta perspectiva ha permanecido arraigada con fuerza, como puede observarse en el trabajo de Tatsien Li (2013), analizado en la sección anterior. Por su parte, el ciclo flexible (Figura 3), consiste también en una iteración de ciclos de modelación, pero se presenta como una versión más ‘suave’. También se observa que los procesos de generación, manipulación y evaluación que contiene el modelo no se definen, pero puede suponerse que, de ser empleado correctamente, el ciclo terminará por asegurar una solución.

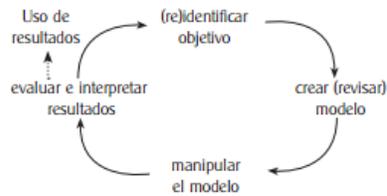


Figura 3. Ciclo Flexible

Bisell y Dillon plantean entonces una serie de consideraciones de la actividad del ingeniero:

- La creación de modelos no forma parte del cotidiano de la práctica del ingeniero, se trata de una situación inusual el crear un modelo desde cero.
- Los ingenieros a menudo seleccionan de un conjunto de “modelos estándar”, cuyas soluciones son bien conocidas y que sólo deben ser ajustados, adaptados o ligeramente modificados.
- El procedimiento de modelación es a menudo incremental, basado en el refinamiento de modelos existentes, y en la práctica y la experiencia – incluyendo experiencias con modelos que han fallado.
- La modelación no es un proceso algorítmico, sino uno subjetivo, que requiere usualmente de conocimiento y de habilidades únicas para cada disciplina.
- La intuición es importante, los modeladores exitosos presentan una habilidad muy bien desarrollada para intuir el tipo de modelo, con mayores probabilidades de éxito en una situación determinada.
- Un modelo matemático es útil, sólo si puede ser empleado exitosamente, es decir, que un modelo menos preciso pero más sencillo de resolver puede ser preferido por sobre uno más sofisticado pero que requiera más recursos para su resolución.

Todas estas características muestran una actividad de modelación asociada a la práctica, a la manera de enfrentar un problema práctico y solucionarlo a través del uso de modelación matemática. La eficacia del modelo está asociada tanto a elementos teóricos como prácticos.

## 1.5 Conclusión

Como producto de nuestro análisis sobre los distintos roles de la modelación matemática, es posible destacar los siguientes <sup>¿por qué sólo a la industria? y es</sup> <sup>su rasgos</sup> <sup>que no se remiten a</sup> <sup>"ingeniero" lo asumen en</sup> <sup>entado...</sup> aspectos:

1) La modelación matemática no ha suplantado completamente el enfoque *teoría-aplicación*, sino a tomar el papel de “puente” entre ellas. 2) La modelación también es vista como puente entre la **matemática e industria**, por lo que ésta resulta de gran importancia en la formación de futuros profesionistas de diversos ramos. 3) El avance industrial y científico ha modificado las tareas laborales, provocando en efecto la necesidad de profesionistas matemáticamente competentes. Esta necesidad se convierte en una demanda, cada vez más apremiante, para las instituciones educativas, cuya responsabilidad es preparar al estudiante para enfrentar eficazmente las problemáticas de su futura vida laboral. Pero, históricamente la enseñanza de las matemáticas ha

tenido un peso considerable en las formaciones ingenieriles, lo que puede ocasionar una mayor dificultad para adaptar dicha enseñanza a las nuevas necesidades de la industria. Por ejemplo, el rol de la tecnología informática es abismalmente distinto en la industria que en la enseñanza de las matemáticas. Estos aspectos, muestran que el proceso de integración de la modelación a la formación puede resultar demasiado complejo. Consideramos que una vía que parece ser ineludible para esta integración, es el análisis de contextos laborales y del rol que la modelación matemática tiene en éstos. Dicho análisis debe ser, sobre un proceso de búsqueda de formas para acortar las brechas (identificadas) entre formación y práctica profesional de los ingenieros. Tal proceso de búsqueda se realizaría por parte de los profesores de matemáticas (a través de distintos mecanismos puestos a su disposición), con el propósito de generar la innovación educativa requerida.

las nuevas disposiciones de la RAE dicen que ahora los pronombres se escriben sin tilde, así como SOLO.

Como ha sido posible observar, la problemática de la integración de actividades de modelación a la formación del ingeniero es muy amplia. Por lo tanto, nos resulta necesario enfocarnos a una formación particular, para lo cual elegimos al Programa Educativo de Ingeniería en Sistemas Computacionales (ISC) del Instituto Tecnológico de Chilpancingo (ITCH). Conforme a lo expuesto, se estudiará el contexto profesional particular de esta ingeniería, buscando identificar y analizar modelos matemáticos que resulten susceptibles de ser llevados al aula en un diseño de actividades de modelación matemática. Planteado de otra manera, tenemos como objetivo el diseñar actividades de modelación matemática que permitan vincular una actividad de modelación, proveniente de la práctica profesional, con la enseñanza de las matemáticas en el Programa Educativo de ISC-ITCH.

¿PREGUNTA que sólo atiende a una comunidad de ingeniero con perfil de egreso específico,

Se plantea entonces la siguiente pregunta de investigación: ¿De qué forma puede integrarse la modelación matemática en una formación de futuros ingenieros? Una posible respuesta a esta interrogante proviene de la consideración de que tal proceso de integración, puede lograrse estudiando el papel que desempeña la modelación en la actividad profesional y como puede relacionarse con la formación de ingenieros. En consecuencia, se puede asumir como objetivo de esta investigación: el diseño de actividades de modelación matemática que permitan vincular una actividad de modelación proveniente de la práctica profesional con la formación matemática de los ingenieros del Programa Educativo ISC-ITCH.

OBJETIVO

Para desarrollar esta investigación, de corte antropológico, hemos elegido como marco la Teoría Antropológica de lo Didáctico y en particular, algunas de sus herramientas, las cuales detallamos en el siguiente capítulo.

¿por qué de corte antropológico? ... por la teoría que sustenta el estudio? inicio la revisión y poca claridad de la elección de los aspectos teóricos

## 2 ELEMENTOS DE LA TEORÍA ANTROPOLÓGICA DE LO DIDÁCTICO

SUGIERO breve justificación del uso del marco.... porque se introdujo sin ningún preámbulo de por qué adopta

### 2.1 Introducción

En este capítulo se presentan elementos de la Teoría Antropológica de lo Didáctico (TAD), que **consideramos necesarios** para alcanzar el objetivo de esta investigación: el diseño de actividades de modelación matemática que permitan vincular una actividad de modelación proveniente de la práctica profesional con la formación matemática de los ingenieros del Programa Educativo ISC-ITCH. Para esto, como se expuso en el capítulo previo, resulta necesario el estudio de la actividad del ingeniero en sistemas computacionales en algún contexto de su práctica profesional. Dicho análisis se enmarcará dentro de la TAD, ya que ésta propone un modelo epistemológico para el estudio de la actividad humana.

Dentro de esta teoría, se estudia la actividad en el marco de la institución en la que se realiza, es decir, la organización social que la posibilita y regula. Distinguiremos instituciones, que nos permiten situar la problemática presentada en el capítulo 1, como son la formación de ingenieros y la industria. Esto nos permitirá analizar cómo estas instituciones determinan, posibilitan y norman las actividades de modelación.

La noción de praxeología es la unidad mínima de análisis de la actividad humana y será considerada para estudiar la actividad de modelación del ingeniero, primeramente en un contexto profesional. Asimismo, consideramos utilizar el modelo praxeológico extendido, ya que el análisis de la actividad de modelación se hará en contextos profesionales, cuya regulación a diferencia de la matemática involucra fuertemente elementos prácticos (técnica eficaz vs técnica teóricamente más óptima). El uso del modelo praxeológico extendido nos parece, permitirá reconocer las tecnologías prácticas (justificaciones, validaciones, elementos que favorecen su uso, etc.) asociadas al uso de los modelos matemáticos, así como las condiciones y restricciones prácticas que pesan sobre el modelo. Este análisis servirá de sustento para el diseño de actividades didácticas de modelación para la clase de matemáticas en una formación específica de ingenieros. Estas actividades didácticas de modelación serán vistas desde la TAD, como procesos de construcción o reconstrucción de praxeologías matemáticas, que también podemos denominar como praxeologías u organizaciones didácticas. Estos procesos de estudio se encuentran estructurados por “momentos didácticos”, que nos servirán para diseñar las organizaciones didácticas.

Este enfoque institucional nos permitirá identificar algunas posibles formas para la integración de la modelación matemática en la formación de ingenieros. Presentamos a continuación los elementos de la TAD que han sido considerados para desarrollar esta investigación.

## 2.2 Nociones básicas

La TAD, es un modelo epistemológico que permite el estudio de la actividad humana dentro de su dimensión institucional. A continuación presentamos sus nociones básicas: institución y praxeología.

### 2.2.1 Institución

La institución es definida de la siguiente manera:

Se trata de organizaciones sociales estables, que enmarcan las actividades humanas y simultáneamente las hacen posibles a través de los recursos que estas instituciones ponen a disposición de sus sujetos. Estos recursos materiales e intelectuales han sido producidos por comunidades, a lo largo de procesos de enfrentamiento a situaciones problemáticas, para resolverlas con regularidad y eficacia. (Castela y Romo, citado en Romo-Vázquez, 2014, pp.323-324)

Desde el punto de vista de la TAD, toda actividad humana se desarrolla en el marco de instituciones, las cuales se caracterizan por poseer un cierto grado de estabilidad. Cada institución tiene una historia propia, nace, se desarrolla, vive sus momentos de auge y también de decaimiento o “conversión” a fin de no perecer. Cada una de estas etapas atiende a razones de ser, que están influenciadas en mayor o menor medida por un contexto social particular y específico. Particularmente, nos interesamos en dos tipos de instituciones, la práctica profesional y la formación del ingeniero. Estas dos instituciones, como hemos visto en el capítulo precedente, han ido evolucionando con el tiempo, modificando sus formas de hacer a partir de los cambios sociales, como han sido la evolución científica y tecnológica. La formación del ingeniero, se asume, aporta al estudiante elementos matemáticos e ingenieriles para su futura práctica profesional. Dicha formación puede verse traducida en términos de instituciones y relaciones entre éstas, como aparece en Romo-Vázquez (2014), de la siguiente manera:

- Instituciones productoras de saberes, matemáticas  $P(M)$  y de disciplinas intermediarias  $P(DI)$ . Estas instituciones son las encargadas de producir modelos, vistos como praxeologías, de validarlos internamente y de asegurar su coherencia con otras praxeologías que conforman la disciplina en cuestión. Se trata de asegurar una pertinencia teórica.
- Instituciones de enseñanza matemáticas  $E(M)$  y de disciplinas intermediarias  $E(DI)$ , son las responsables de difundir las praxeologías. Se encargan por tanto de operar sobre las praxeologías para adaptarlas a las condiciones y restricciones particulares de la enseñanza (lógica escolar), es decir, realizar las transposiciones que resulten necesarias.
- Instituciones usuarias, son las instituciones que representan a la práctica profesional  $Ip$ , o a las actividades prácticas  $Ap$ . Mientras que  $Ip$  representa la práctica profesional del ingeniero,  $Ap$  comprende actividades similares que se desarrollan en el marco de su formación, usualmente en forma de proyectos. Son las que ponen en funcionamiento las praxeologías para atender las necesidades propias de la práctica. Puede tratarse tanto de comunidades de ingenieros como de los proyectos mismos que éstos realizan.

Se asume una relación entre las instituciones de la enseñanza y de la práctica profesional. Esta relación supone que la enseñanza forma ingenieros capaces de atender problemáticas de la práctica profesional, sin embargo, no se está cumpliendo. Para analizar estas instituciones y sus relaciones, se analizan particularmente algunas actividades de modelación que tienen lugar en ellas.

### 2.2.2 Praxeología

La praxeología es la unidad mínima de análisis de la actividad humana, y está conformada por cuatro componentes:

- El tipo de tarea  $T$ , define un grupo de tareas, por ejemplo generar un algoritmo de optimización de una búsqueda en internet es un tipo de tarea  $T$  y generar el algoritmo PageRank es una tarea  $t$ , donde  $t \in T$ .
- La técnica  $\tau$ , es la manera en que se lleva a cabo la tarea.
- La tecnología  $\theta$ , son los discursos que explican, producen y justifican las técnicas, permitiendo reconocer su razón de ser.
- La teoría  $\Theta$ , son discursos más generales que explican, producen y justifican a su vez las tecnologías.

Estos componentes pueden ser organizados en dos bloques:

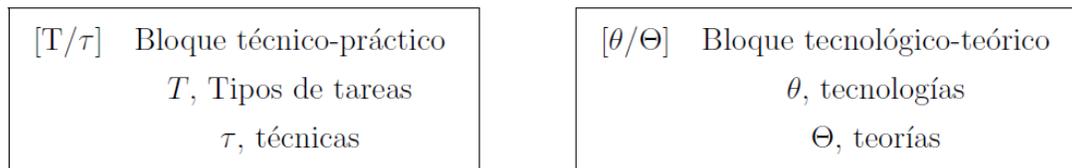


Figura 1. Bloques praxeológicos

El primer bloque, “técnico-práctico” representa el “saber hacer”, e incluye la tarea a resolver y la forma en que ésta se realiza (técnica). El segundo bloque, “tecnológico-teórico” representa el “saber”. La praxeología puede ser utilizada para analizar todo tipo de actividad humana: actividad matemática, de modelación o de ingeniería, por mencionar algunas.

### 2.2.3 Praxeologías matemáticas y mixtas

La actividad del ingeniero, como hemos visto en el capítulo 1, se caracteriza por hacer intervenir diferentes instituciones, por ejemplo, la matemática, la industria, etc. Esto nos hace suponer que en el análisis de la actividad de modelación de una práctica profesional de ingenieros, podremos encontrar diferentes tipos de praxeologías. Por ejemplo, praxeologías matemáticas y mixtas.

Las praxeologías matemáticas, están conformadas por un tipo de tarea matemática ( $T$ ), que requiere de una técnica matemática para ser resuelta ( $\tau$ ), la cual está justificada/validada por una tecnología matemática ( $\theta$ ), la cual a su vez encuentra sustento en una teoría matemática ( $\Theta$ ) que, por extensión, valida también a la técnica.

Las praxeologías mixtas, tienen entre sus componentes elementos de, al menos, dos instituciones, por ejemplo matemático se ingenieriles. Así, una praxeología mixta puede estar conformada por un tipo de tarea no matemático ( $T$ ), enfrentado a través de técnicas matemáticas ( $\tau$ ), que son sustentadas por tecnologías matemáticas, así como de la ingeniería ( $\theta$ ) y teorías ( $\Theta$ ) matemáticas y de ingeniería. Estas teorías provenientes de la ingeniería pueden sustentar técnicas y tecnologías matemáticas utilizadas dentro de la ingeniería, como es, por ejemplo, estructura de datos.

### 2.3 Modelo praxeológico extendido

El modelo praxeológico extendido define dos componentes de la tecnología: componente teórica ( $\theta^{th}$ ) y componente práctica ( $\theta^p$ ), que pueden esquematizarse de la manera siguiente:

$$\left[ \begin{array}{l} T, \tau, \theta^{th}, \Theta \\ \theta^p \end{array} \right] \begin{array}{l} \leftarrow P(S) \\ \leftarrow Iu \end{array}$$

Figura 2. Modelo praxeológico extendido

Las flechas observadas en el esquema simbolizan las instituciones que producen dichas componentes, la tecnología teórica y la teoría son producidas y validadas por la institución productora de saberes  $P(S)$ . La tecnología práctica es validada por la institución usuaria. La práctica es vista como institución usuaria  $Iu$ , que emplea las praxeologías producidas por  $P(S)$  para realizar ciertos tipos de tareas.

La componente teórica ( $\theta^{th}$ ), mantiene el papel original de la tecnología y concentra los discursos (justificaciones y explicaciones) de origen teórico, que en este caso puede ser matemático, mientras que la noción de tecnología práctica ( $\theta^p$ ), tiene seis funciones: (Castela y Romo-Vazquez, citado en Romo-Vazquez, 2014, pp. 327-328):

1. **Describir el tipo de tareas y la técnica.** La producción de un discurso que caracteriza el tipo de tarea y los pasos que componen una técnica son considerados como una pieza de saber no identificable a la maestría de la técnica en sí misma. Las acciones en juego y el contexto donde se sitúa la praxeología, en un sistema compartido, se pueden identificar en la elaboración de un sistema de representaciones verbales y, más ampliamente, simbólicas. La producción de estos lenguajes, y la descripción que ellos permiten, constituye una componente decisiva del proceso de transmisión de una invención técnica.
2. **Validar la técnica.** La función considerada corresponde a lo que en general se entiende bajo el término justificar, en los textos que definen la noción de praxeología. Los saberes considerados establecen que la técnica produce bien lo que ella dice que produce, que los pasos que la componen permiten conseguir los objetivos que le son asignados. En el caso de las matemáticas, esta función es generalmente asegurada por los saberes justificados por las teorías matemáticas. [...] Sin embargo, en otros contextos, los saberes validados experimentalmente

en laboratorio o empíricamente en el uso pueden validar una técnica. Éste es particularmente el caso cuando se trata de validar las adaptaciones de la técnica.

3. **Explicar la técnica.** Se trata de saberes que permiten analizar cómo la técnica y sus diferentes pasos permiten conseguir los objetivos pretendidos. Es cuestión de una inteligencia de las causas. Después de la diatriba de los geómetras en torno a los métodos analíticos de Descartes, se sabe que existen –incluso en matemáticas– validaciones que no explican. Existen también explicaciones que no validan, porque éstas no respetan completamente las normas de la validación en la institución que examina esta cuestión de la validez, apoyándose por ejemplo en analogías. Contribuyen a la comprensión de las causas de los sujetos y, por tanto, están sumamente relacionadas a su cultura compartida.
4. **Facilitar la aplicación de la técnica.** Los saberes considerados en esta función permiten a los usuarios utilizar con eficacia, pero también con un cierto confort, la técnica. Éstos son portadores de mejoras pero también de advertencias que permiten evitar errores y torpezas conocidas como frecuentes. Este dominio de saberes es el terreno privilegiado de las elaboraciones tecnológicas de los usuarios. Dicho dominio produce efectos retomados de descripciones que lo especifican al adaptarlo a las condiciones particulares del contexto institucional de utilización y al enriquecimiento de la memoria de las experiencias acumuladas.
5. **Motivar la técnica y los pasos que la componen.** Estos saberes están orientados hacia la práctica. Participan de una inteligencia de los fines: son los objetivos esperados que justifican racionalmente los pasos, mostrando su razón de ser. Se trata de escribir una historia de la técnica que sitúe sus componentes, los unos en relación con los otros: ¿por qué (¿para hacer qué?) se realiza tal paso en tal momento? Los saberes de motivación son frecuentemente saberes relacionados con el tipo de tareas, puesto que ellos analizan los objetivos. Permiten anticipar las etapas esperadas y juegan, por tanto, un papel heurístico importante cuando la aplicación de la técnica necesita adaptaciones.
6. **Evaluar la técnica.** Los saberes considerados aquí tienen que ver con el dominio, las condiciones y los límites de una técnica en relación con las tareas del tipo T. Ellos pueden igualmente concernir la ergonomía de la técnica desde el punto de vista de sus usuarios. Las funciones evaluar, facilitar y motivar están a veces muy relacionadas: la puesta en evidencia de ciertas dificultades (evaluar) puede provocar, al cabo de cierto tiempo, la producción de mejoramientos (facilitar); la motivación está dada por la evaluación.

Estas seis funciones tecnológicas nos sirven para reconocer el rol de los discursos prácticos, permitiendo identificar aquellas condiciones que determinan el funcionamiento y el uso del modelo, los elementos o restricciones provenientes del contexto, que provocan adaptaciones y los conocimientos que facilitan la toma de decisiones.

Las herramientas teóricas hasta este momento expuestas, nos permitirán analizar la actividad de modelación del ingeniero en sistemas computacionales para identificar modelos susceptibles de ser llevados al aula. Es decir, posibles relaciones entre formación matemática y actividad práctica de ingenieros. Con el propósito de mostrar cómo estas praxeologías pueden ser estudiadas dentro de la enseñanza de las matemáticas, presentamos a continuación, los niveles praxeológicos, la co-determinación entre lo matemático y lo didáctico así como los momentos del estudio.

## 2.4 Niveles de organización praxeológica, proceso de estudio y niveles de co-determinación

Dentro de la TAD se reconocen organizaciones praxeológicas de diferentes niveles según la complejidad de la praxeología en cuestión, como se muestra a continuación:

- **Praxeologías puntuales**, si están generadas por lo que se considera en la institución como un único tipo de tareas  $T$ . Esta noción es relativa a la institución considerada y está definida, en principio, a partir del bloque práctico-técnico  $[T/\tau]$ .
- **Praxeologías locales**, resultado de la integración de diversas praxeologías puntuales. Cada praxeología local está caracterizada por una tecnología  $\theta$ , que sirve para justificar, explicar, relacionar entre sí y producir las técnicas de todas las praxeologías puntuales que la integran.
- **Praxeologías regionales**, se obtienen mediante la coordinación, articulación y posterior integración alrededor de una teoría matemática común  $\Theta$ , de diversas praxeologías locales. La reconstrucción institucional de una teoría matemática requiere elaborar un lenguaje común que permita describir, interpretar, relacionar, justificar y producir las diferentes tecnologías ( $\theta_j$ ) de las praxeologías locales ( $PL_j$ ) que integran la praxeología regional.
- **Praxeologías globales**, que surgen agregando varias praxeologías regionales a partir de la integración de diferentes teorías. (Bosch, Gascón, García, Higuera, 2006, p. 39)

Estos niveles de organización nos permiten observar que la actividad puede tener diferentes niveles de complejidad y de generalidad. El análisis de la actividad puede hacerse puntualmente o globalmente, eso dependerá de los objetivos de la investigación, siendo útiles estos niveles para descomponer la actividad. Por ejemplo, estudiar una actividad de modelación en una práctica profesional, va requerir de dividir la actividad para poder comprender “partes” que la conforman. Las praxeologías puntuales permitirán dar cuenta de actividades más pequeñas que en su conjunto conformarán la actividad estudiada, que consideramos debe ser al menos de un nivel local. Analizar únicamente praxeologías puntuales nos ofrecería un análisis muy “parcelado” de la actividad de modelación, pero analizar una praxeología regional o global nos parece ir más allá de las posibilidades de este estudio.

### 2.4.1 Proceso de estudio y sus momentos

Construir o reconstruir una praxeología es realizar un proceso de estudio. Bosch, et al (2006), observan que las praxeologías escolares, por ejemplo, no son creadas de manera

súbita, ni se consideran completas en ningún momento, sino que son resultado de un largo proceso de construcción o proceso de estudio, que resulta en la praxeología matemática. En estos procesos de estudio pueden identificarse *momentos*, que permiten dar una estructura a cualquier proceso de construcción y/o reconstrucción de las praxeologías matemáticas. Estos “momentos didácticos”, son los siguientes:

**1) Momento del primer encuentro**

Se trata del primer contacto de los estudiantes con algún componente de la praxeología, o con alguna situación problemática que puede ser respondida o recibir alguna aportación de la praxeología. El propósito del primer encuentro, es concientizar a la comunidad de estudio sobre la existencia de los componentes de la praxeología, y de las cuestiones que motivan su construcción y uso, es decir, que son su razón de ser.

**2) Momento exploratorio**

Se explora la situación problemática, se identifica la tarea a resolver, y se construye o plantea una técnica para resolverla. En este momento se identifican dos etapas: la investigación de técnicas o mecanismos que puedan solucionar las cuestiones problemáticas, y en la segunda, se consideran problemas concretos dentro de la tarea y se resuelven con técnicas matemáticas. En estos dos primeros momentos, el estudiar problemáticas se convierte en el medio que permite construir y poner en práctica una técnica, la cual se establecerá como el medio de resolución de otras problemáticas del mismo tipo.

**3) Momento de construcción de un entorno tecnológico-teórico**

Se constituye la tecnología y la teoría que sustenta a la técnica o técnicas que se construyen o surgen durante el proceso de estudio. Se trata de dar respuesta a cuestiones sobre el funcionamiento, validez y resultados arrojados por la técnica. Estas respuestas pueden requerir realizar nuevas tareas matemáticas que se integrarán a la praxeología en estudio

**4) Momento de trabajo de la técnica**

Este momento se inicia buscando que el estudiante las utilice de manera rutinaria para así fortalecer su dominio sobre ellas. Es en este proceso que la técnica puede desarrollarse, arrojando así técnicas relativamente nuevas para la comunidad de estudio.

**5) Momento de la institucionalización**

En este momento se precisa y formaliza la praxeología, se definen los elementos que formaron parte de su construcción, y que ahora forman parte de ella, separándolos de aquellos otros que pudieron haber participado en el proceso de construcción pero que finalmente no necesitan ser integrados a la praxeología.

**6) Momento de la evaluación**

Es el momento en que se cuestiona la calidad de los componentes que conforman la praxeología. La tarea, si está bien definida o si se asocia a las cuestiones trabajadas. La técnica, si es fiable, si es la más eficiente para atender la problemática. La tecnología, si justifica de manera adecuada el funcionamiento y resultados de las técnicas, y más importante aún, si está

definida de manera clara y la información que ofrece permite comprender la técnica de manera que puedan derivarse o construirse nuevas técnicas.

Estos momentos no están estructurados de manera lineal, por lo que pueden repetirse las veces que sea necesario, el enfoque puede estar centrado sólo en algunos de ellos y pueden darse incluso de manera simultánea, si así se requiere. Con ellos se conforma un proceso de estudio que, como toda actividad humana, puede ser visto como una praxeología. Ésta se denomina praxeología u organización didáctica, la cual contemplará tipos de tarea, técnicas, tecnologías y teorías didácticas, con los mismos roles descritos anteriormente.

### 2.4.2 Niveles de determinación

Bosch, et al (2006), observan que existen relaciones de mutua dependencia entre las organizaciones o praxeologías matemáticas (OM) y sus organizaciones o praxeologías didácticas (OD) correspondientes. Estas relaciones están estructuradas en una sucesión de niveles denominada, *jerarquía de niveles de determinación*:

Civilización → Sociedad → Escuela → Pedagogía →  
Disciplina → Área → Sector → Tema → Cuestión

La estructura de las OM condiciona las formas posibles de organizar su proceso de estudio. La naturaleza y funciones de los dispositivos didácticos, en el nivel al que pertenezcan, condicionan a su vez el tipo de OM que puede construirse/reconstruirse en él. En otras palabras, toda cuestión a estudiarse en una institución didáctica está integrada en un tema, que pertenece a un sector, que se encuentra incluido en una disciplina. Si la disciplina en cuestión es Matemáticas, podemos renombrar estos niveles como, “niveles matemáticos”, y aquellos que se encuentran más allá de la disciplina son llamados, “niveles pedagógicos”. Estos niveles siguen ejerciendo influencia sobre los inferiores, determinando que praxeologías se han de considerar como parte de ellos.

Consideramos que estos niveles de determinación pueden corresponderse con los niveles de organización praxeológica, de la siguiente manera:

Cuestión → Praxeología Puntual  
Tema → Praxeología Local  
Sector → Praxeología Regional  
Área → Praxeología Global

Podemos observar que queda definido, de manera clara, el tipo de OM que puede reconstruirse en cada nivel, actividad que se realizará a través de un proceso de estudio (OD). Este proceso de estudio puede ser visto como una Actividad de Estudio e Investigación, la cual parte de una cuestión generatriz  $Q$  que permite identificar un tipo de problemas y una técnica de resolución de dichos problemas, así como una tecnología apropiada para justificar y comprender mejor la actividad matemática que se ha llevado al cabo (Chevallard, citado en Bosch, et al, 2006, pp. 59). Podemos observar dentro de este proceso, la presencia de los tres primeros momentos didácticos: *momento del*

*primer encuentro, momento exploratorio y momento de construcción de un entorno tecnológico-teórico.*

Si me identifico o construyo se corresponden (no en FUTURO)

Consideramos que las praxeologías **identificadas** en el análisis de la actividad de modelación del ingeniero, se **corresponderán** a algún nivel de organización praxeológica. Esto nos permitirá identificar aquellas praxeologías matemáticas, pertenecientes a E(M), con las que pueda establecerse una posible relación o incluso generarla dentro de una nueva praxeología de modelación. Se buscará que esta praxeología sea de nivel local, que será la base para proponer una OD y por tanto un proceso de estudio asociado.

## 2.5 Conclusión

¿Por qué como conclusión? debiera ser del capítulo y me parece que en

Los **elementos teóricos** presentados a lo largo de este capítulo, nos permiten **primeramente considerar la formación** de ingenieros en términos de instituciones. Particularmente, **la práctica profesional**, vista también como institución, tiene una relación con la formación, en el sentido que esta última prepara a futuros profesionales. De esta manera, se busca reconocer posibles relaciones entre estas instituciones.

¿Por qué teóricos? se refieren al texto que la conclusión en algunos casos. Aquí lo que se derivan son acciones por ejemplo la que más allá es una función y debiera plantear en pasado.

Para analizar la actividad de modelación del ingeniero en sistemas computacionales, **se utilizará** la noción de praxeología, con el objetivo de identificar los tipos de tareas que enfrentan, las técnicas que utilizan y las validaciones y justificaciones asociadas al uso de los modelos. Consideramos que, una vez identificada una praxeología en la práctica analizada, cuyo modelo matemático permita plantear una relación con E(M), será posible diseñar un proceso de estudio, u organización didáctica que permita llevar la actividad de modelación al aula.

Consideramos una propuesta metodológica, que nos permita realizar lo antes señalado, para el diseño de actividades de modelación, basada en la TAD y **que presentamos** en el siguiente capítulo.

que se son presente

## 3 CONSIDERACIONES METODOLÓGICAS

### 3.1 Introducción

En el presente capítulo se expone la metodología propuesta para el desarrollo de esta investigación. En primer lugar, reconocemos las instituciones que participan en la formación de futuros ingenieros de sistemas computacionales y las del Instituto Tecnológico de Chilpancingo (ITCH), como son enseñanza de las matemáticas y enseñanza de las disciplinas intermedias. En segundo lugar, se presenta una metodología para el diseño de actividades basadas en modelación matemática, sustentada en la TAD e inicialmente propuesta en Macias (2012). Finalmente, se muestra como las cuatro fases que conforman dicha metodología: *Elección de un contexto extra-matemático*, *Análisis praxeológico e identificación de un modelo matemático*, *Análisis del modelo en E(M)* y *Diseño de la actividad didáctica*, fueron utilizadas y adaptadas para el desarrollo de esta investigación, en la que se elige analizar la programación web, contexto de la práctica profesional del ingeniero en sistemas computacionales, para reconocer en ésta la actividad de modelación matemática. Este primer análisis debe conducir a analizar a su vez actividades de modelación en la institución enseñanza de las matemáticas, para así generar una base para el diseño didáctico, como se muestra en la fase 3 de la metodología aquí presentada. Asimismo, se presentan las condiciones de la institución educativa ITCH y las razones de su elección para implementar el diseño didáctico producido.

### 3.2 Instituciones presentes en la formación del ingeniero en sistemas computacionales

En esta investigación consideramos la formación de ingenieros en sistemas computacionales, que se imparte en el Tecnológico Nacional de México, ubicado en Chilpancingo de los Bravo, Guerrero. Dicha formación se analiza en términos de las instituciones que la componen y de las relaciones que pueden establecerse entre éstas. Para analizarla, se utiliza el modelo de instituciones que participan en una formación de ingenieros (Romo-Vázquez, 2014), presentado en el capítulo previo. A partir de este modelo consideramos específicamente tres instituciones:

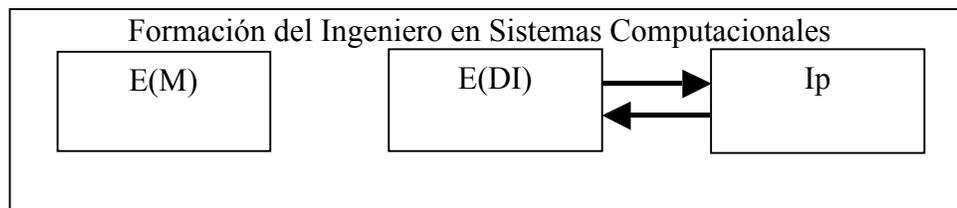
- Enseñanza de la matemática,  $E(M)$ , conformada por el conjunto de asignaturas matemáticas presentes en la formación del ingeniero en sistemas computacionales del ITCH:  
Cálculo diferencial, Matemáticas discretas, Cálculo integral, Álgebra lineal, Probabilidad y estadística, Cálculo vectorial, Ecuaciones diferenciales y Métodos numéricos.
- Enseñanza de la especialidad,  $E(DI)$ , conformada por el conjunto de asignaturas que la especialidad presentes en la formación del ingeniero en sistemas computacionales del ITCH:

Fundamentos de programación, Programación orientada a objetos, Estructura de datos, Sistemas operativos, Tópicos avanzados de programación, Fundamentos de bases de datos, Taller de sistemas operativos, Fundamentos de telecomunicaciones, Taller de base de datos, Simulación, Fundamentos de ingeniería de software, Arquitectura de computadoras, Lenguajes y autómatas, Redes de computadoras, Administración de base de datos, Graficación, Ingeniería de software, Lenguajes de interfaz, Lenguajes y autómatas II, Sistemas programables, Programación lógica y funcional, Programación Web, Inteligencia artificial

POR QUÉ "SE SUPONE" sugiero que si el programa de enseñanza lo plantea más bien decir que se supone que ese automatizado proponen que los ingenieros en sistemas, porque SUPONE, es hipotético desde la postura de Julian, y no de lo oficial. Si es así, entonces decir, que DEBIERA....

- Práctica Profesional, *Ip*, el contexto laboral del ingeniero en sistemas computacionales, donde resuelve problemáticas propias de su especialidad.

Las instituciones de enseñanza de especialidad  $E(DI)$  se **supone** deben brindar al ingeniero en sistemas computacionales las herramientas necesarias para desarrollarse en su práctica profesional *Ip*, estableciendo una relación entre estas instituciones,  $E(DI)$  e *Ip*. Esta relación puede esquematizarse de la siguiente manera:



POR QUÉ "ENSEÑAR" me remite a la crítica de lo "tradicional" o acaso eso plantea la TAD ?

Figura 1. Relaciones entre  $E(DI)$  e *Ip*

La flecha que se dirige de  $E(DI)$  a *Ip*, representa a las praxeologías que son **enseñadas** con el objetivo de que el ingeniero pueda ponerlas en funcionamiento en *Ip*, por ejemplo, las pertenecientes a Estructura de Datos, que preparan al estudiante para desenvolverse en los contextos laborales de la programación, programación Web y Gestión de base de datos. Asimismo, para la flecha que va de *Ip* a  $E(DI)$ , consideramos que representa la existencia de miradas, estudios e investigaciones, que analizan las praxeologías que existen en *Ip* y que pueden proporcionar contextos de enseñanza para  $E(DI)$ . Esto implicaría por ejemplo, que se utilizaran actividades de los contextos mencionados previamente para la enseñanza de las praxeologías de Estructura de datos. Lo anterior, hace emerger algunas cuestiones, ¿cómo estas relaciones pueden reconocerse/analizarse de manera sistemática? ¿En qué medida es posible crear otras relaciones y con qué fines?

De manera similar a  $E(DI)$ , se asume que la enseñanza de la matemática  $E(M)$ , tiene la función de proporcionar al estudiante las praxeologías que pondrá en funcionamiento dentro de  $E(DI)$ , y a su vez, en *Ip*. Esto puede observarse en el plan de Estudios del ITCH, donde se muestra que ciertas asignaturas de  $E(M)$  son requisito para cursar algunas asignaturas de  $E(DI)$ . Por ejemplo, la asignatura de Álgebra lineal es un requisito para cursar algunas asignaturas de  $E(DI)$ , como Estructura de datos e Investigación de operaciones. Esto nos muestra que existe una relación entre  $E(M)$  y  $E(DI)$ , que puede esquematizarse de la manera siguiente (Fig. 2):

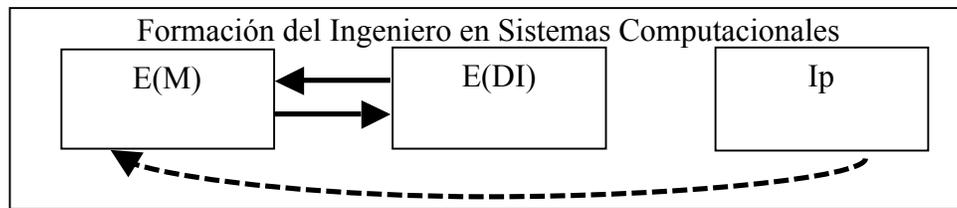


Figura 2. Relaciones entre  $E(M)$  y  $E(DI)$

La flecha que va de  $E(M)$  a  $E(DI)$ , representa las praxeologías proporcionadas por  $E(M)$  que serán utilizadas por el estudiante en  $E(DI)$ , la flecha que va de  $E(DI)$  a  $E(M)$  representa posibles contextos de uso de las matemáticas. Es de nuestro interés generar una relación de  $Ip$  a  $E(M)$  que represente praxeologías provenientes de la práctica, que bajo un análisis y diseño didáctico, puedan ser consideradas en  $E(M)$ . Ésta vincularía en efecto la enseñanza de la matemática con un contexto de la práctica profesional del ingeniero en sistemas computacionales, como por ejemplo programación, programación Web o Gestión de base de datos.

Para provocar esta relación  $Ip \rightarrow E(M)$ , resulta necesario buscar y, en la medida de lo posible, reconocer una praxeología que pueda relacionar las tres instituciones  $E(M)$ ,  $E(DI)$  e  $Ip$ . Para esto, se analizará un contexto de  $Ip$  donde tenga lugar praxeologías de modelación del ingeniero en sistemas computacionales. Las praxeologías observadas en el análisis pueden también, dada su naturaleza, ser parte de  $E(M)$  y de  $E(DI)$ . Para poder determinar estas relaciones con claridad, deberá hacerse un análisis fino del contexto de  $Ip$  que permita identificarlas, ya sea a través del estudio de libros de texto, documentos especializados o entrevistas con sujetos expertos de  $Ip$ .

### 3.3 Metodología para el diseño de actividades de modelación para formación de futuros ingenieros

Consideramos la metodología propuesta inicialmente por Macias (2012), para el diseño de actividades basadas en modelación matemática. Dicha metodología se desarrolla en el marco de la TAD, por lo que comparte los elementos teóricos considerados para esta investigación. Asimismo, el objetivo de la metodología es diseñar actividades de modelación que permitan llevar praxeologías provenientes de contextos de  $E(DI)/Ip$ , a  $E(M)$ , de manera que se genere la relación que buscamos entre tales instituciones. Esta metodología continúa en proceso de desarrollo y refinación, estableciendo actualmente las siguientes cuatro fases:

1) *Elección de un contexto extra-matemático*

Tras identificar el nivel educativo en el que se da la enseñanza sobre la que se quiere intervenir, en esta fase se busca reconocer y seleccionar un contexto en el cual se encuentre en funcionamiento la matemática que es objeto de enseñanza  $E(M)$ , según planes y programas, en tal nivel. Se expresa que en el caso particular de la ingeniería,  $Ip$  y  $E(DI)$  son los contextos naturales, en que  $E(M)$  se utiliza.

2) *Análisis praxeológico e identificación de un modelo matemático*

En esta fase se realiza un análisis praxeológico del contexto seleccionado para identificar praxeologías, ya sean matemáticas o mixtas (de modelación). Esto con el objetivo de identificar un modelo que se muestre susceptible de ser llevado a la enseñanza de la matemática,  $E(M)$ , y que sea el objeto de análisis para la siguiente fase.

3) *Análisis del modelo en  $E(M)$*

Se analiza la o las praxeologías identificadas en la fase previa, empleando el modelo praxeológico extendido, utilizando las funciones de la tecnología práctica para reconocer elementos tecnológicos que puedan ser parte del diseño de actividades ( $OD$ ). A su vez, se analizan las instituciones de  $E(M)$ , para identificar praxeologías matemáticas con las que las praxeologías analizadas puedan tener una relación.

4) *Diseño de la actividad didáctica*

El diseño de la actividad, organización didáctica, debe estar basado tanto en el análisis praxeológico, empleando el modelo praxeológico extendido así como en el análisis del modelo en  $E(M)$ . Es decir, es necesario reconocer las praxeologías de uso y de enseñanza para poder ver las relaciones entre éstas. El tipo de organización didáctica a diseñar es flexible. Pero esta flexibilidad no representa transgresión alguna del marco teórico aquí utilizado, pues, en el marco de la TAD, considerar un proceso de estudio que permita la construcción/reconstrucción de las praxeologías analizadas, es una elección lógica metodológicamente hablando.

La manera en que estas cuatro fases fueron consideradas para desarrollar esta investigación, se presenta a continuación.

### 3.3.1 Primera fase: Elección de un contexto extra matemático

Se elige como contexto de análisis la programación web, ya que tal y como lo propone la metodología, **ésta** forma parte de la práctica profesional del ingeniero en sistemas computacionales. Para probar que este contexto resultaba idóneo para el análisis se hizo una primera búsqueda de documentos propios a este contexto, encontrándose dos:

1) “El secreto de Google y el Álgebra lineal”

Se trata de un artículo proveniente de  $Ip$ , con carácter de difusión. En este documento se presenta el método PageRank y tiene por objetivo presentar el modelo que lo **sustenta**. El título del documento sugiere que ciertos elementos del álgebra lineal van a figurar en la realización de algunas tareas llevadas a cabo por Google <sup>consideró</sup> para presentar resultados de una búsqueda, por lo que se **decidió** considerarlo para el análisis. Sin embargo, no permitió observar con claridad las tecnologías que sustentan al modelo, debido a que éste se encuentra ya validado por  $Ip$ , por lo que exponer los elementos tecnológicos no resulta importante. Debido a esto, se buscó otro documento que permitiera clarificar los aspectos tecnológicos.

2) “El Álgebra lineal detrás de los buscadores”

Se trata de un documento de carácter didáctico, elaborado por un matemático y profesor de álgebra lineal, debido a estas características, el documento hace

intervenir diferentes instituciones:  $P(M)$ ,  $E(DI)$  y  $E(M)$ . Esto nos ofrece una perspectiva muy amplia, ya que hace evidente elementos de estas instituciones, según la situación que se esté abordando. El objetivo del documento es analizar y transparentar la matemática que sustenta al algoritmo PageRank así como los elementos que lo conforman, al tiempo que busca mostrar a los estudiantes de informática el uso que tienen las matemáticas dentro de esta disciplina. Debido a estas características, el documento permite observar los elementos tecnológicos, teóricos y prácticos, que el documento anterior no mostraba.

Un breve análisis de estos documentos permitió identificar la problemática de la *optimización de una búsqueda Web*, problemática enmarcada dentro del área de la programación Web, y que es resuelta a través del método PageRank. Este análisis también permitió reconocer los siguientes modelos matemáticos: grafos, matriz de incidencia, matrices, sistemas de ecuaciones lineales, valores y vectores propios y cadenas de Markov. De estos modelos, los grafos, así como sus representaciones matriciales, son enseñados en Matemáticas discretas, mientras que las matrices y sistemas de ecuaciones lineales, son enseñados en Álgebra lineal. Aun cuando los demás modelos no encontraron representación en alguna  $E(M)$ , el análisis de los modelos que sí, en el contexto de la programación Web a través de estos dos documentos permitiría reconocer relaciones con instituciones de  $E(M)$ .

### **3.3.2 Segunda fase: Análisis praxeológico e identificación de un modelo matemático**

El análisis del modelo se realizó a través de la comparación de los elementos presentados por los dos documentos seleccionados, los cuales tienen, aún cuando se ubican en instituciones diferentes, el mismo objetivo, mostrar el método con una estructura similar: representación matemática de la definición del criterio y cálculo de las importancias en algunos procesos de búsqueda en la Web. Esto permitió observar la manera en que el modelo es construido, así como sus elementos matemáticos. Asimismo, se consideró el modelo praxeológico extendido, utilizando sus funciones para analizar los elementos prácticos de la praxeología seleccionada y todas las que pudieran surgir en el análisis. Es decir, que dentro de esta fase concentramos el análisis propuesto por la metodología para las fases dos y tres. El análisis permitió dar cuenta de la existencia de una macro-praxeología, conformada por el tipo de tarea: *ordenar los sitios que integran a la Web*, cuya técnica contiene al método PageRank.

Para determinar si las praxeologías identificadas pueden ser llevadas al aula, a través del diseño de una organización didáctica, resulta necesario identificar posibles relaciones con las praxeologías de  $E(M)$ . Para esto, se procede a la tercera fase.

### **3.3.3 Tercera fase: Análisis del modelo en $E(M)$**

En esta etapa se consideran las praxeologías analizadas en la fase 2, para identificar si pueden encontrar relación con praxeologías en  $E(M)$ , de manera que la organización didáctica resultante pueda llevarse a esta institución. Este análisis se realizó sobre los modelos matemáticos identificados como existentes en  $E(M)$  en la primera fase: Grafos,

Matriz de incidencia, Matrices y Sistemas de ecuaciones lineales. Los dos primeros en Matemáticas Discretas y los dos segundos en Álgebra lineal

Se analizó también, de manera breve, el artículo “*Standing in the shoulders of giants*”, documento que busca evidenciar la presencia del modelo subyacente a PageRank en contextos pertenecientes a otras áreas: sociometría, bibliometría, econometría, deportes. Estos contextos se asumieron como importantes, debido a la posibilidad de utilizarlos para el diseño de la organización didáctica.

### **3.3.4 Cuarta fase: Diseño de la actividad didáctica**

El diseño de la organización didáctica considera el análisis praxeológico de la actividad, los elementos prácticos aportados por el modelo praxeológico extendido y las relaciones identificadas con  $E(M)$  y  $E(DI)$ . Para este diseño, se consideran las praxeologías seleccionadas en la etapa anterior como viables para ser llevadas al aula. Se **considerará** un proceso de estudio que lleve al estudiante a la reconstrucción de las praxeologías elegidas. El diseño de la organización didáctica se presenta en el capítulo 5.

### **3.3.5 Consideraciones para la implementación en el ITCH**

Para la implementación del diseño de la organización didáctica, se eligió el ITCH. Esta institución tiene desde su creación, el 2 de Octubre de 1984, el propósito de impartir educación superior tecnológica, con el objetivo de formar profesionales capaces de solucionar las problemáticas planteadas por una sociedad en constante transformación, no solo social, sino tecnológica. Para esto el ITCH se propone como misión: “Preservar, innovar, trascender y aplicar el conocimiento científico-tecnológico en la formación de profesionistas, con responsabilidad social y capacidad de investigar, que desarrollen y apliquen propuestas de solución a las problemáticas de la sociedad”. Podemos observar en esta declaración, la necesidad de que los estudiantes utilicen los conocimientos que la formación les brinde, en la resolución de problemáticas en sus áreas respectivas.

Para llevar a cabo la implementación en el ITCH, se identificó en qué período son impartidas las materias de Álgebra lineal y Matemáticas discretas. Debido a que se imparten en segundo y primer semestre, se dio preferencia a Álgebra lineal, considerando que cualquier estudiante que la curse, habrá pasado ya por Matemáticas discretas. Para conseguir un grupo en el que se pudiera implementar el diseño, se realizaron dos entrevistas con un docente de la institución, que imparte asignaturas matemáticas  $E(M)$ , de la formación de Ingeniería en sistemas computacionales. Sin embargo, el docente no se encontraba impartiendo Álgebra lineal durante el semestre, por lo que se recurrió a un grupo de Investigación de operaciones, asignatura de  $E(DI)$  que se imparte en tercer semestre, y que presenta como prerrequisito el haber cursado Álgebra lineal, por lo que se asumió como correcto. Por cuestiones de disponibilidad de tiempo del grupo elegido, que se encontraba en evaluaciones, la actividad tuvo que planearse para dos o tres sesiones de una hora, como máximo. Debido a esto, la actividad no pudo ampliarse demasiado y se diseñó con un formato de cuestionario.

### 3.4 Conclusión ¿?

La propuesta metodológica presentada en este capítulo, ha sido desarrollada con el objetivo de generar una relación entre  $E(M)$  y  $E(DI)/Ip$ . Para conseguirlo, se propone el diseño de organizaciones didácticas sustentadas en el análisis de contextos extra-matemáticos siguiendo la metodología propuesta inicialmente por Macias (2012) y aquí adaptada. Así, uno de nuestros objetivos es basar el diseño de la organización didáctica en el análisis praxeológico de la programación web, utilizando para ello el modelo praxeológico extendido, para posteriormente completarlo con un análisis de  $E(M)$ , particularmente del Álgebra Lineal y de Matemáticas Discretas. Estos análisis deben permitirnos evidenciar posibles relaciones entre  $Ip$  y  $E(M)$  así como generar nuevas. Asimismo, se busca poner a prueba los elementos tecnológicos obtenidos en el análisis praxeológico y su utilidad para un futuro diseño didáctico. Diseño, que se pretende funja como herramienta para el docente, ejemplificando cómo generar y desarrollar un proceso de modelación en el aula. Asimismo, se pretende poner a prueba una Actividad de Estudio e Investigación (AEI). Recordamos que una AEI es un proceso de estudio que considera solo los primeros tres momentos didácticos (momento de primer encuentro, momento exploratorio y momento de construcción del entorno tecnológico-teórico) como base para su diseño. En la implementación de la AEI (organización didáctica) se espera observar el tránsito del estudiante por estos momentos. Sin embargo, a pesar de que el interés del diseño radica en los primeros tres momentos, no se descarta que los demás puedan aparecer durante la implementación.

En el siguiente capítulo se presenta el análisis de la programación web ( $Ip$ ) y particularmente de la construcción de la técnica PageRank así como del uso de modelos matemáticos involucrados. Esto permitirá mostrar cómo el análisis praxeológico puede sustentar el diseño de una organización didáctica.

## 4 ANÁLISIS PRAXEOLÓGICO DE LA MODELACIÓN MATEMÁTICA EN EL MÉTODO PAGERANK

### 4.1 Introducción

En la propuesta metodológica para el diseño de actividades basadas en modelación, presentada en el capítulo previo, se observa que el análisis praxeológico de un contexto extra-matemático posee un papel fundamental. Esto se debe a que resulta indispensable para sustentar el diseño de una organización didáctica. En este capítulo se presenta dicho análisis praxeológico, que constituye la segunda fase de la metodología que empleamos, utilizando para ello el modelo praxeológico extendido. Este análisis fue realizado dentro del contexto de la programación Web, eligiendo el denominado “método PageRank” como praxeología de nuestro interés. Para esto se utilizaron dos documentos de literatura especializada. El primero, un representante de  $I_p$ , que permite observar rasgos generales del funcionamiento del método, sin profundizar en los aspectos tecnológicos. A pesar de esto, su análisis permite comenzar el análisis de las praxeologías. En el segundo documento, que hace convivir elementos de  $P(M)$ ,  $E(M)$ , e  $I_p$ , permite observar, aclarar y reforzar los aspectos tecnológicos que no fueron abordados con profundidad en el primero. Como resultado del análisis se presenta una arborescencia praxeológica, que nos permite organizar las praxeologías encontradas (mixtas y no matemáticas) según su nivel de complejidad. Se aborda entonces la tercera fase de nuestra metodología, seleccionando las praxeologías susceptibles de ser llevadas al aula en una organización didáctica y sustentando esta decisión con un análisis a  $E(M)$ , así como un breve análisis de otros contextos en los que el modelo es utilizado, y que puedan ser de utilidad para el diseño de la organización didáctica.

### 4.2 Análisis del método PageRank

Se presenta a continuación el análisis praxeológico realizado al método PageRank, comparando como es presentado en los dos documentos elegidos: “*El secreto de Google y el Álgebra Lineal*” (D1, en adelante) y “*El Álgebra lineal detrás de los buscadores de Internet*” (D2, en adelante). El análisis busca transparentar los elementos tecnológicos teóricos y prácticos presentes en la actividad de modelación, para lo cual se emplea el modelo praxeológico extendido, resaltando las funciones de las tecnologías identificadas a lo largo del análisis. Se inicia describiendo a continuación la problemática que origina la construcción del método PageRank, y las tecnologías que van a condicionarlo desde el comienzo. Estas tecnologías, que condicionan al método, pueden ser motivadas por las exigencias de la problemática, así como de la práctica profesional desde la cual se está abordando. Esto podrá observarse con mayor claridad dentro del análisis.

#### 4.2.1 La problemática asociada a los buscadores Web

Dentro de D1, la problemática de la organización y ordenación de información es vista en términos de las necesidades del buscador Web, necesidades que deben ser cumplidas

para alcanzar el objetivo principal que Google persigue: que el usuario encuentre su respuesta en los primeros 10 resultados. Para ello, se hace la siguiente consideración:

Tenemos los resultados de una búsqueda: ¿cómo los mostramos, en qué orden? Necesitamos un **criterio** de ordenación, una asignación de **importancias** a cada sitio de la red.

Esto lleva a identificar una tarea orientada a la resolución de la problemática: *ordenar los resultados de una búsqueda Web*. Es decir, cuando un usuario hace una búsqueda sobre la palabra institución, el buscador utiliza ciertos “criterios de búsqueda” para identificar esa palabra en todos los sitios que la contengan. Sin embargo, esta obtención de sitios que contienen la palabra institución es aleatoria y se hace necesario tener un criterio de ordenación que nos permita ordenar los resultados obtenidos para identificar los más relevantes. Entonces surge la cuestión: ¿Qué criterio o criterios son necesarios para ordenar los resultados? El autor presenta lo que denomina “un par de ingredientes” que Google utiliza para ordenar sus búsquedas (Fig. 1).

NOTA. Hay un par de ingredientes que Google combina con el que veremos aquí:

- no es lo mismo que un cierto término aparezca en una página en el título, en negrita, en un tipo de letra pequeña, etc.
- Para búsquedas combinadas, tampoco es lo mismo que los términos buscados aparezcan “cerca” o “lejos” unos de otros.

Figura 1. Consideraciones de búsqueda

Nótese que de estas consideraciones se desprenden, no sólo dos criterios de ordenación, sino también una categorización entre búsquedas, “simples” (de una sola palabra) y combinadas (dos o más palabras). Aparece aquí la tarea, *ordenar los sitios de la red*, que da origen al método PageRank, el cual es objeto de nuestro análisis en la siguiente sección.

Por su parte en el documento D2 se comienza describiendo la historia del buscador Google, presentando la **motivación** principal del mismo: ofrecer un servicio de búsqueda que fuera *rápido* y *eficiente*. Contrariamente a D1, en este documento se dedica una sección completa a la problemática bajo el título: “Los buscadores de internet”, presentando una metáfora para explicar la tarea a la que se enfrenta el buscador.

Uno podría comparar el trabajo de un buscador con el de un bibliotecario. Para hacerlo más explícito, digamos que se trata de un bibliotecario de las épocas en las que no había ordenadores. Si uno acudía a la biblioteca en aquellos cada vez más lejanos tiempos intentando encontrar información sobre algún tema en particular, se iba a encontrar con un gran fichero o catálogo enorme, impreso, conteniendo toda la información existente en esa biblioteca hasta la última actualización. Con un poco de suerte además había también alguna especie de catálogo-diccionario, relacionando libros con algunas palabras clave. Supongamos ahora que mi biblioteca contiene más de mil millones de libros, y que bajo la

palabra clave “jirafa” hay cuatro millones de textos que tienen algo que decir al respecto, y que para enumerarme uno por uno todos estos textos -a razón de un texto cada 10 segundos- el bibliotecario demoraría casi 463 días. Yo claramente no necesito leer los cuatro millones de libros para hacer el trabajo que me toca, quizás con 10 de ellos ya me alcance. Pero entonces... ¿cuáles 10? (D’Andrea, 2012, p.3)

Es como el autor explica el objetivo que el buscador, y por extensión el método, persigue: que dentro de los primeros 10 resultados se encuentre la respuesta esperada. Prosigue entonces **describiendo** al buscador en términos de la situación del bibliotecario, utilizando esta analogía con la intención de facilitar su comprensión:

Un buscador de internet esencialmente es una especie de catálogo de biblioteca junto con un bibliotecario que te recomienda qué libros leer. El éxito de este buscador depende justamente de tener una buena base de datos, ordenada de acuerdo a palabras clave de una manera razonable, y también un buen recomendador, ya que uno quiere acceder a la información de manera rápida y eficiente. (D’Andrea, 2012, p.4)

Con esto podemos establecer una relación con la motivación original del buscador: *eficiencia y rapidez* en su funcionamiento, pero al mismo tiempo, también las condiciones que deben cumplirse para conseguirlo, ambas con la misma importancia: una base de datos ordenada de acuerdo a palabras clave (indexada), y un “buen recomendador”. Se introduce aquí el término “recomendador” que nos permite traducir la tarea T, identificada en D1, que el método deberá cumplir en otros términos: recomendar sitios al usuario en los que pueda encontrar la respuesta que requiere. Y, ¿cómo podemos asegurar que sea un “buen recomendador”? En términos de los objetivos: será uno que pueda entregar esa respuesta al usuario en sólo 10 sitios. Entonces surge la pregunta que define la problemática a resolver: “¿Cómo se hace el trabajo de bibliotecario? Es decir, ¿cómo decido qué páginas mostrar primero cuando alguien pone en el buscador la palabra “jirafa”?”, es así como plantea entonces analizar el funcionamiento del algoritmo PageRank.

Notamos que existen diferencias entre ambos documentos, que nos permiten identificar su naturaleza y objetivos, D1 no profundiza en la problemática, haciendo explícita la tarea que se realiza a partir del método PageRank: *ordenar el conjunto de sitios*, planteando la *asignación de importancias* como parte de este método. En cuanto a D2, identificamos, a partir del tratamiento extenso que da a la problemática, el uso de analogías con un objetivo didáctico que permiten contextualizarla para una mejor comprensión. Esta forma de abordar la situación refleja la intención del autor de explicitar con el mayor detalle posible la necesidad de ordenar la información (relativo al análisis de D1). A continuación se analizará el proceso de construcción de la técnica PageRank que permite resolver la tarea ordenación de los sitios web en ambos documentos, deteniéndonos en las diferentes etapas del proceso y la manera en que éstas son descritas, analizadas y validadas.

### 4.2.2 Construcción de la técnica (método) PageRank en D1

Para iniciar el análisis del método, que nosotros identificaremos como técnica, PageRank y del modelo que lo subyace, los cuales permiten realizar la tarea: *ordenar los sitios que conforman la Web*, consideramos primeramente el documento D1. Como primer paso, se reconoció el criterio de ordenación, la *importancia de un sitio*, lo que conlleva *asignar importancias a cada sitio de la red*. Para esto, en primer lugar se proponen dos variables, sitios, denotada con  $P$  e importancias, denotada con  $x$  (Fig. 2).

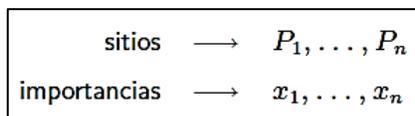


Figura 2. Asignación de variables a sitios e importancias

Tras la asignación de estas dos variables, aparece como siguiente paso de la técnica: *representar matemáticamente la estructura de la Web*, tener lo que el autor llama: “El modelo” (Fig. 3).

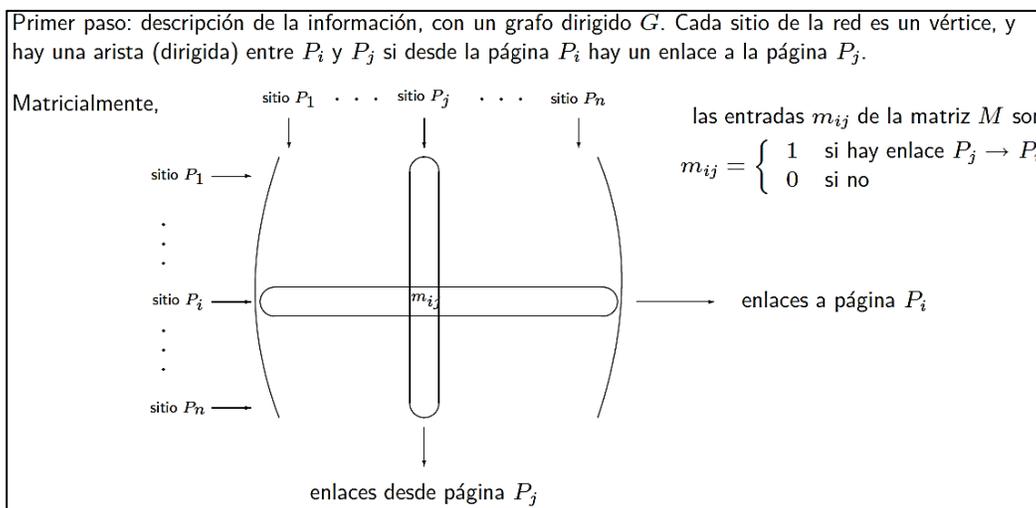


Figura 3. “El modelo”

Este modelo es mostrado “naturalmente”, sin explicar si se trata del único, el más usual o el más conveniente. Sin embargo, parece ser que su elección es **motivada** por la facilidad con que permite representar los sitios y las relaciones entre ellos, a través de un modelo matricial, asociado a un grafo dirigido  $G$ . El uso de matrices ofrece facilidades, del tipo computacional, en su operación y manejo, por lo que parece ser común en el área. Es posible que debido a que se trata de un documento perteneciente a  $I_p$ , se considere innecesario explicitar esta clase de **justificaciones** asociadas a la elección del modelo. Consideramos que por su naturaleza, este modelo y su uso para representar la red, podrían ser objeto de enseñanza en un curso de Álgebra Lineal o en un curso de Estructuras de Datos en una formación de ingenieros en sistemas computacionales. Lo anterior, nos hace pensar que esta praxeología puede ser común a las instituciones  $I_p$ ,  $E(M)$  y  $E(DI)$  y por tanto se convierte en un objeto primario de

análisis. Con el objetivo de comprender cómo este modelo es utilizado en la técnica de PageRank se continúa con el análisis del documento.

Una vez presentado el modelo en D1, el autor propone una relación proporcional entre  $x$  y  $P$ , llamada: “**Primer intento**:  $x_j$  es proporcional al número de páginas que enlazan con  $P_j$ ”. Es decir, que la importancia de una página debería resultar proporcional a la cantidad de enlaces que apuntan hacia ella. Implícitamente parece plantearse la siguiente pregunta: *¿Es suficiente esta relación entre  $x$  y  $P$  para resolver el problema de ordenación?* cuestión que trata de responderse al **evaluar** este paso de la técnica, planteando la siguiente consideración: no es suficiente sólo la cantidad de enlaces que recibe un sitio, sino también la importancia de las páginas de las que parte dicho enlace. Por ejemplo, un enlace desde un sitio “importante”, como Amazon o Microsoft, tendrá un valor inherentemente mayor que un enlace proveniente de un sitio “pequeño” como un blog, o un sitio de reciente creación. Esto **motiva** una adaptación al modelo, para lo cual el autor reconoce la necesidad de combinar ambos aspectos:

- Páginas muy citadas.
- Páginas poco citadas, pero desde sitios importantes.

La consideración de estos dos nuevos rasgos, involucra nuevos conceptos matemáticos, esto hace que la relación entre  $x$  y  $P$  se vuelva más compleja, pero interpreta mejor las exigencias del problema tratado; dicha relación se describe como sigue:

“**Segundo intento**:  $x_j$  es proporcional a la suma de las importancias de las páginas que enlazan con  $P_j$ ”.

Es importante notar que se transforma el planteamiento de la asignación de importancias, de uno de sumatorias a uno de un sistema de ecuaciones lineales, que se plantea como más **fácil** de operar (Fig. 4) aunque no se aborda algún fundamento para esta conclusión.

Por ejemplo, la página  $P_1$  es citada desde las páginas  $P_2$ ,  $P_{25}$  y  $P_{256}$ , mientras que  $P_2$  sólo se cita desde  $P_1$  y  $P_{256}$ , etc. Nuestra asignación  $x_1, \dots, x_n$  debe cumplir que

$$\begin{aligned}x_1 &= K(x_2 + x_{25} + x_{256}), \\x_2 &= K(x_1 + x_{256}), \\&\vdots\end{aligned}$$

Figura 4. Reformulación de la asignación de importancias

Esta reinterpretación de la asignación de importancias permite utilizar un modelo matricial para modelarla, herramienta de la que el autor hace uso de manera inmediata (Fig. 5). Podemos asumir que en este paso existe una tecnología teórica que sustenta la transformación del modelo en uno matricial, proveniente de P(M)-E(M), así como una tecnología práctica que **evalúa** el modelo previo y **motiva** su reformulación para **facilitar** su uso, proveniente de  $I_p$ .

Matricialmente,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = K \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$P_1$   $P_2$   $P_{25}$   $P_{256}$   
 $\downarrow$   $\downarrow$   $\downarrow$   $\downarrow$

Figura 5. Modelo matricial de importancias

Este modelo matricial no sólo **facilita** el reinterpretar nuevamente la asignación de importancias, y por lo tanto definir nuevamente la relación entre  $x$  y  $P$ , sino que permite transformar una vez más el modelo, que el autor asume como venido de la inspiración, “como mágicamente, hemos transformado el problema en uno de autovalores y autovectores”:

$$Mx = \lambda x$$

Con esta reformulación también plantea que  $x$  debe ser un autovector de  $M$ , sin describir quién es  $x$ , porqué se trata de un autovector y porqué debemos centrar nuestra atención en él. Es decir, que el autor considera que estos elementos tecnológicos resultan “transparentes” desde un inicio. Continúa estableciendo dos consideraciones para este modelo:

- Que las entradas de  $x$  sean no negativas ( $x \geq 0$ ).
- Que  $x$  fuera un autovector único.

Este segundo punto es señalado por el autor sólo como “conveniente”, pero para ninguno de estos dos elementos tecnológicos se presenta alguna justificación que sustente la necesidad de que se cumplan. Asumimos que tales justificaciones, por la naturaleza de las condiciones y del modelo, que parece situarse en lo puramente matemático, tendrían que provenir de  $P(M)$ . A continuación, bajo el título “Mejor en proporciones” el autor propone reescribir la matriz  $M$ . Como en el caso de las consideraciones del paso anterior esta transición aparece de manera natural, la cual consiste en tomar en cuenta no sólo la importancia de la página de la que proviene el enlace, sino también la cantidad de enlaces que parten de la misma (Fig. 6).

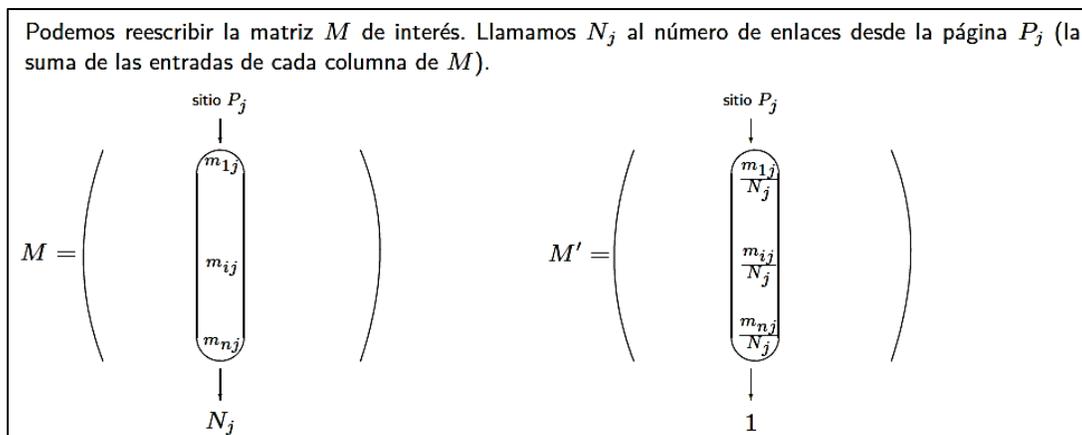


Figura 6. Reescritura de la matriz  $M$

Tras haber hecho la reescritura, el autor explica que la matriz  $M$  ha sido transformada en una matriz estocástica, para establecer entonces que se hará una formulación alternativa, utilizando para ello la cadena de Markov y señala que esto se debe a la riqueza que plantea la reformulación. Ésta parece ser la conclusión del proceso de construcción del modelo. Consideramos que dado que el objetivo de este documento es presentar un algoritmo que es *funcional*, validado desde la práctica ( $I_p$ ) no parece necesario justificar cada etapa de su construcción. Para complementar este análisis, buscando principalmente identificar aquellas tecnologías que se encuentran ausentes en D1, se analiza a continuación el modelo visto desde D2.

#### 4.2.3 Construcción de la técnica (método) PageRank en D2

Recordemos que D2 tiene por título: “El álgebra lineal detrás de los buscadores de Internet” y ha sido elegido debido a que su autor, matemático y profesor de álgebra lineal en el área de informática, busca mostrar las herramientas del álgebra lineal que son utilizadas en el método o algoritmo Page Rank. Para esta investigación resulta muy interesante debido a que su análisis sobre el método no está <sup>¿por qué los supuestos?</sup> centrado en el algoritmo mismo sino en los elementos del álgebra lineal  $E(M)$  que intervienen, pero además orientándolo a estudiantes de informática. Por lo que **suponemos** que los elementos tecnológicos provenientes de  $I_p$ , implícitos en D1, aquí se hacen explícitos.

El análisis del método comienza en la sección 4 del documento, titulada: “El modelo PageRank. Vectores y Valores Propios”, este título sugiere que encontraremos un modelo matemático subyacente al método PageRank, y por lo tanto, que la modelación matemática formará parte del proceso de construcción del mismo. El autor señala que una de las preguntas primordiales para poder comprender el funcionamiento del algoritmo PageRank es: “¿cómo se modela matemáticamente un buscador de internet?” Se plantea lo siguiente: “la *teoría de grafos* nos ayudará a modelar nuestra situación”. La situación es **descrita** como: “*modelar la Web*”, y la identificamos como el primer paso de la técnica PageRank: *representar matemáticamente (modelar) la estructura de la Web*. Es posible observar en ambos documentos D1 y D2 el uso del *grafo dirigido* como herramienta de modelación, y de manera similar es presentado como la única opción para llevar a cabo este paso. En D2, se **explica** lo que representa cada uno de sus elementos y cómo es que éstos representan la Web: viendo a los sitios Web como *nodos* y a los enlaces entre sitios como *aristas orientadas*, esto último se debe a que los enlaces son forzosamente unidireccionales, es decir, no se puede realizar un recorrido “de regreso” sin un enlace que vaya en esa dirección. Se nos ofrece a partir de esto una representación gráfica de lo que en adelante el autor llamará el *grafo de internet* (Fig. 7). En ambos documentos es notorio aun cuando no se hace explícito, la consideración de dejar fuera del análisis el caso imposible o inexistente: un sitio no puede enlazarse a sí mismo.

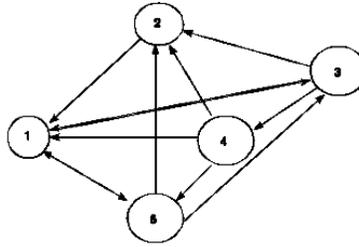


Figura 7. El grafo de internet

¿por qué plural "las importancias"

Es tras la propuesta de este grafo, que el autor retoma nuevamente **las importancias** para entonces tratar de definir cómo asignarlas. Para esto, de manera implícita, utiliza el grafo dirigido para hacer un análisis de la forma en que están relacionados los sitios. Esto resulta en suponer que una manera “razonable” de asignar las importancias será considerar el número de enlaces que un sitio recibe como medida de su importancia. Este razonamiento es similar al planteado en D1 y nos lleva a la misma hipótesis: “*cuantos más enlaces reciba una página, más importante será*”, antes de continuar el autor reconoce esta aproximación como análoga al *índice de citas*, modelo conocido y utilizado en ambientes académicos y de investigación, que sirve en este punto para ejemplificar su uso en otro contexto.

Se analiza nuevamente la situación con la intención de **evaluar** el modelo obtenido, señala: “*El problema con este modelo aparentemente sensato es que uno podría “inflar” rápidamente la importancia de una página Web determinada simplemente creando varias páginas que tengan enlaces con la misma*”. Notemos que su evaluación es hecha calificando al modelo de “aparentemente sensato”, pero reconociendo su falta de eficiencia ya que resulta fácilmente manipulable y por ende poco representativo de la relación entre los sitios. Esto **motiva**, como sucedió también en D1, la necesidad de considerar no sólo el *número* de citas o enlaces recibidos, sino también la *importancia* de los sitios que las emiten. Se **explica** entonces que esto reduciría la influencia que páginas de reciente creación (que dentro del modelo serían poseedoras de una menor importancia) pudieran tener sobre el grafo. Se **valida** este nuevo modelo al observar cómo se distribuyen las importancias en el “grafo de importancias” (Fig. 8), dado que a pesar de que “F” y “C” reciben un mismo número de enlaces (lo que les daría la misma importancia en el modelo previo), la importancia de las páginas que los emiten es diferente, y ya que “C” está siendo enlazado por el sitio con la mayor importancia de la red se le considerará más importante que “F”.

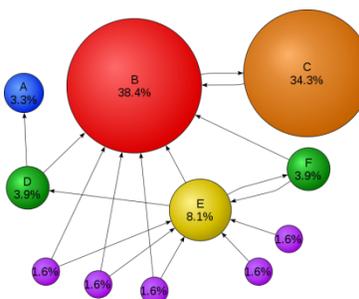


Figura 8. Grafo de importancias

Una vez hecha esta adaptación, el modelo resultante es **descrito** a través del llamado *postulado de PageRank*, que es presentado de la siguiente manera: “La importancia  $x_i$  de la página  $P_i$  es directamente proporcional a la suma de las importancias de las páginas que enlazan con ella”. Podemos reconocer que se trata de la segunda aproximación presentada en D1, sin embargo a diferencia de D1, no se hace la reinterpretación inmediata al modelo matricial, sino que el proceso de aproximación al problema es explicado en términos del *grafo de internet*. El autor plantea entonces que lo siguiente es “traducir matemáticamente” lo anterior para observar como el álgebra aparece “naturalmente” dentro del contexto. La modelación es vista como una *traducción* de una formulación/hipótesis a un lenguaje matemático.

Antes de considerar el modelo matricial y su uso, se trata de justificar la necesidad del mismo a través de algunas condiciones que surgen de la situación. La primera de ellas es que la Web está conformada por una cantidad masiva de sitios y enlaces entre ellos, y tan sólo el tratar de representarlas resultaría en un enorme gasto de recursos desde el punto de vista computacional (es decir, que se trata de una tecnología práctica), tales como: espacio de almacenamiento, tiempo de ejecución, uso de memoria, entre otros. Además, aunque la representación del grafo pueda resultar ilustrativa para el estudio y la comprensión de la situación (ejemplificado por las dos primeras proposiciones), el volumen de información hace que no resulte útil para el problema que se está abordando. Cabe hacer una pausa en este punto para comentar que, al tratar con el modelo matricial estamos entrando al área de interés del documento, esto puede notarse al observar que a diferencia del grafo de internet empleado para modelar la Web, en este caso sí se presentan las justificaciones de su uso. Las condiciones planteadas responden la pregunta implícita: ¿por qué usar un modelo matricial en lugar del grafo dirigido? Podemos observar de manera clara su papel como tecnología si planteamos la pregunta en otros términos: ¿Qué **motiva** el cambio de modelo?

Una vez expuestas las justificaciones mencionadas, se establece que resultará necesario utilizar la denominada “*matriz de incidencia*” del grafo (identificada como perteneciente a  $E(DI)$ ) como nuevo modelo de la situación, la cual es **descrita** entonces como: “la matriz cuadrada de tamaño igual a la cantidad de nodos del grafo”. Se **explica** también que los valores  $(i, j)$  de la matriz serán iguales a 1 en caso de que exista un enlace desde el sitio  $P_i$  a  $P_j$  y de 0 si es que no lo hay, asegurando de esta manera que la matriz representa la estructura de la Web que había sido modelada previamente con el grafo dirigido (Fig. 9).

$$M_0 = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Figura 9. Matriz de incidencia del grafo de Internet

Con la introducción del modelo matricial el autor señala la aparición del Álgebra Lineal en el problema, así como la de los valores y vectores propios, y se dispone a profundizar en ellos, confirmando así el foco de su interés: transparentar los elementos del Álgebra

Lineal que subyacen a la técnica PageRank y que forman parte de la tecnología del mismo. Con este objetivo se presenta el siguiente teorema:

**Teorema 4.1.** Si  $M_I$  es la matriz de incidencia del grafo de internet, y

$\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}_{\geq 0}^n$  el vector de importancias, entonces se cumple

$$M_I^t \mathbf{x}^t = \lambda * \mathbf{x}^t$$

donde  $\lambda \in \mathbb{R}_{> 0}$  es la constante de proporcionalidad.

Podemos notar que nos encontramos frente al mismo modelo de valores y vectores propios que fue planteado en D1, sin embargo es importante observar que, a diferencia de D1, en D2 se identifican y detallan los elementos que conforman el modelo. El autor utiliza un lenguaje matemático matizado por términos del problema tales como: *matriz de incidencia* y *vector de importancias*. Debido a esto consideramos que tales elementos pertenecen a  $P(DI)$  o a  $I_p$ , y por lo tanto, el modelo planteado en el Teorema (palabra que evoca  $P(M)$ ) pertenecería también a estas instituciones. Inmediatamente después se presenta la definición de vector propio, elemento que el autor señala como proveniente del Álgebra Lineal clásica y que considera “conveniente” recordar. El énfasis en este punto parece radicar en cómo este elemento será utilizado tanto en la modelación de la situación como en la construcción de la técnica PageRank, sirviendo como justificación teórica del funcionamiento del mismo. Mostramos a continuación la definición.

**Definición 4.2.** Dados una matriz cuadrada  $M$  de tamaño  $n \times n$ , un vector no nulo  $\mathbf{x} \in \mathbb{R}^n$  (o  $\mathbb{C}^n$ ) y un número  $\lambda \in \mathbb{R}$  (o  $\mathbb{C}$ ), el vector  $\mathbf{x}$  se dice *vector propio* de  $M$  con *valor propio asociado*  $\lambda$  si y solo si se verifica

$$M\mathbf{x}^t = \lambda * \mathbf{x}^t$$

Podemos confirmar el papel de esta definición como sustento teórico del Teorema 4.1 tras observar el lenguaje con el que se aborda y describe este modelo (el cual se corresponde con el expresado en el Teorema previo), es expresado matemáticamente  $P(M)$ - $E(M)$ : *matriz cuadrada, valor propio y vector propio*. Al no abordar el modelo en términos de la situación, esta definición se corresponde con la que podríamos encontrar en un libro de Álgebra lineal (documento perteneciente a  $E(M)$ ). Para ilustrarlo, consideramos la siguiente definición, presentada en Lay (2001):

Un **vector propio** de una matriz  $A$   $n \times n$  es un vector  $\mathbf{x}$  diferente de cero tal que  $A\mathbf{x} = \lambda\mathbf{x}$  para algún escalar  $\lambda$ . Un escalar  $\lambda$  se llama **valor propio** de  $A$  si existe una solución no trivial  $\mathbf{x}$  de  $A\mathbf{x} = \lambda\mathbf{x}$ ; una  $\mathbf{x}$  tal se llama *vector propio correspondiente a  $\lambda$* . (p. 297)

Con esto podemos corroborar, en primer lugar, que la Definición 4.2 proviene de la institución  $P(M)$ - $E(M)$ . Por lo tanto, la misma puede ser considerada como una tecnología teórica asociada al Teorema 4.1, el cual a su vez es el modelo matemático que subyace al algoritmo PageRank. Aquí resulta interesante examinar lo siguiente: si

consideramos que los modelos presentados en el Teorema 4.1 y la Definición 4.2 son esencialmente el mismo, pero cuyas características nos hacen reconocerlos en instituciones diferentes ( $P(M)-E(M)$ , y  $P(DI)-Ip$ , respectivamente) podemos observar entonces que el modelo matemático se presenta como una praxeología matemática transpuesta por  $P(DI)$  o  $Ip$ . O bien, que esta praxeología puede ser susceptible de operarse para transponerse de  $E(M)$  a  $Ip$  y de  $Ip$  a  $E(M)$ .

A continuación el autor nos presenta el siguiente Corolario derivado del Teorema y Definición:

**Corolario 4.3.** *El vector de importancias de las páginas web es un vector propio (positivo) de la matriz  $M^t$ , y la constante de proporcionalidad  $\lambda$  es el valor propio asociado a este vector.*

Este Corolario permite llegar a una conclusión sobre lo analizado previamente, cumpliendo con el papel de justificación, o mejor dicho, tecnología del modelo. Explicita la relación existente entre la Definición 4.2 y el Teorema 4.1, y en consecuencia, la relación establecida entre las instituciones a las que ambos pertenecen ( $P(M)-E(M)$  y  $P(DI)-Ip$ , respectivamente) a través del modelo.

A continuación el autor plantea el Ejemplo 4.4, con el que procede a mostrar cómo se lleva a cabo la asignación de importancias de acuerdo al postulado de PageRank. Estableciendo de antemano a  $\lambda$  como la constante de proporcionalidad (considerada en el teorema), y utilizando la información del *grafo de internet* que hemos venido considerando (Fig. 7), se obtiene lo siguiente (Fig. 10):

$$\begin{aligned} 0 \cdot x_1 + 1 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 + 1 \cdot x_5 &= \lambda x_1 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 + 1 \cdot x_5 &= \lambda x_2 \\ 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 1 \cdot x_4 + 1 \cdot x_5 &= \lambda x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5 &= \lambda x_4 \\ 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 1 \cdot x_4 + 0 \cdot x_5 &= \lambda x_5, \end{aligned}$$

Figura 10. Sistema de ecuaciones lineales

Observamos entonces que la asignación de importancias es vista como un *sistema de ecuaciones lineales*, dentro del cual podemos observar que se están cumpliendo las condiciones establecidas por el postulado: que la importancia  $x_i$  correspondiente a un sitio  $P_i$  resulta directamente proporcional a la sumatoria de las importancias de los sitios que enlazan a  $P_i$ . Tras una primera observación podemos identificar los coeficientes 0 y 1 que aparecen junto a las importancias como representantes de la existencia o ausencia de un enlace entre los sitios. Esta función de los coeficientes (representar los enlaces entre sitios), nos lleva a determinar que se está operando con la *matriz de incidencia* (Fig. 9). Sin embargo, después de una comparación entre los valores que aparecen como coeficientes del sistema de ecuaciones y aquellos que encontramos en la matriz de incidencia, podemos observar que los coeficientes de las importancias en cada ecuación se corresponden con los valores de las columnas de la matriz de incidencia y no con los valores de las filas como podría esperarse, lo que es expresado por el autor de la siguiente manera: “*si en las filas de la matriz de incidencia del grafo de internet uno puede leer cuántos enlaces salen de una página dada, justamente en las columnas aparecerán tantos unos como enlaces haya hacia la página indexada por esa columna.*”

Es decir, que cada ecuación utiliza los coeficientes de cada columna para asegurar que está operando con los enlaces recibidos por  $P_i$ . Encontramos entonces que el sistema de ecuaciones lineales se encuentra asociado con la *traspuesta* de la matriz de incidencia, en lugar de la matriz por sí misma.

Se procede entonces a representar el sistema de ecuaciones lineales tratado previamente en notación matricial:

$$M_0^t \mathbf{x}^t = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}^t \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_7 \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_7 \end{pmatrix} = \lambda \cdot \mathbf{x}^t.$$

Figura 11. Sistema en notación matricial

Al representar en notación matricial el sistema de ecuaciones, podemos identificar cada uno de los elementos que son descritos en el Teorema 4.1, tanto la matriz  $M_1^t$  (traspuesta de la matriz de incidencia), como el *vector de importancias*  $\mathbf{x}$ , así como la *constante de proporcionalidad*  $\lambda$ . Notemos que aquí se está utilizando la *transformación lineal*  $P(M)-E(M)$ . Regresando al modelo, la identificación de los elementos del Teorema en la representación matricial del sistema de ecuaciones, cumple con el objetivo de comprobar que el modelo enunciado en el Teorema coincide con lo establecido en el postulado de PageRank, una tecnología que **valida** el nuevo modelo. Es así como se ilustra y justifica la llegada al modelo de vectores y valores propios que sustenta al algoritmo PageRank a partir de una reinterpretación del modelo previo:

$$M\mathbf{x} = \lambda\mathbf{x}$$

Posteriormente el autor utiliza Mathematica para encontrar soluciones para el ejemplo del *grafo de internet* propuesto (Fig. 7). Dentro de las soluciones encontradas para  $\lambda$  sólo una cumple con ambas condiciones: tratarse de un número real y positivo, a continuación se observa el vector de importancias asociado al valor de  $\lambda$  que se identifica como probable solución, para analizar los valores que arroja, es decir, las importancias que asigna a cada sitio. Estos valores determinan un orden de importancias que coincide con lo que esperaríamos tras observar el grafo, colocando en primer lugar al sitio  $P_1$ , y enseguida a  $P_2$  y  $P_3$  con una importancia de idéntico valor, pues aún cuando  $P_2$  recibe un voto más que  $P_3$ , este último recibe un enlace de un sitio más importante y que el enlace de  $P_2$ . Esto permite afirmar que el modelo representa de manera adecuada la situación planteada, **validándolo**.

La solución encontrada es cuestionada por el autor: “¿habrá un único valor propio positivo, y asociado al mismo un solo vector propio positivo (solución) para cualquier matriz cuadrada con ceros y unos?”. Esta pregunta permitirá **evaluar** si es posible confiar en las soluciones que arroja el modelo. Para responderla presenta un ejemplo (Fig. 12) de una matriz 2x2 cuyo único valor propio es  $\lambda=0$ , agregando que: “También es fácil construir matrices de tamaño más grande con más de un valor propio positivo”.

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Figura 12. Ejemplo

Se plantea hacer una modificación al modelo “*pequeña pero de vital importancia*”, **motivada** por lo siguiente:

Con el transcurrir del tiempo y el advenimiento de las redes sociales (muy propensas a producir hechos “en cadena”, en varios lugares y al mismo tiempo) se encontró una falla en el modelo previsible desde un primer momento: si una página tiene un solo enlace, este enlace vale lo mismo que cualquier otro enlace de otra página que produzca un millón de enlaces. Es como -si bien producir enlaces desde mi propia página no aumenta mi importancia- cuantos más enlaces produce mi página, más afecta a toda la red. (D’Andrea, 2012, p.8)

Nos encontramos entonces ante una situación que permite evaluar al modelo, el cual resulta aún influenciado, situación que es atendida con una nueva adaptación del mismo. Ésta requiere de una modificación, planteada como “*sencilla*”, a la matriz. La cual consiste en sustituir todos los unos de la matriz de incidencia por valores de  $1/(\# \text{de enlaces desde } P_i)$ . **Explicando** que de esta manera podemos reducir la “autoridad total” o “poder de voto” de un sitio a un valor total de 1, y este valor es entonces distribuido uniformemente de acuerdo a la cantidad de enlaces salientes, de manera que quede reducida la capacidad de un sitio para influenciar la estructura de la Web. Esto lleva a modificar la matriz de incidencia de manera que ahora se encuentra compuesta por números no negativos que suman 1 en cada fila, asemejándose a una distribución de probabilidad (Fig. 13). Es al observar estas características del nuevo modelo matricial que el autor presenta la *matriz estocástica por filas*, denominada en Álgebra Lineal ( $P(M)-E(M)$ ) como *matriz estocástica derecha*, y que se define como “una matriz cuadrada, cada una de cuyas filas está formada por números reales no negativos, sumando cada fila 1”. El autor menciona que este tipo de matriz aparece con frecuencia en el modelado de ciertos procesos, y que éstos serán mencionados posteriormente.

$$M_{0,E} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \end{pmatrix}$$

Figura 13. Reinterpretación de la Matriz de Incidencia, “Matriz de Transición”

Si recordamos el proceso llevado a cabo en D1 se llega a un modelo similar a éste (Fig. 6), sin embargo esta modificación de la matriz (denominada en D1 como *matriz de transición*,  $P(M)-E(M)$ ) es realizada sin una buena explicación de las razones o justificaciones para tal acción, a diferencia de lo que ocurre en D2, donde se perciben claramente las condiciones que **motivan** dicha modificación. Esto se debe principalmente a que el interés de los autores de ambos documentos radica en áreas diferentes, mientras en D1 se busca mostrar la existencia de algunos elementos provenientes del Álgebra Lineal dentro del algoritmo PageRank, en D2 se busca transparentar tales elementos y justificar tanto su uso como algunas de las razones para emplearlos (tecnología). Es importante notar que ambos modelos matriciales, a pesar de parecer similares a primera vista, presentan una diferencia importante: mientras en D2

se llega a una *matriz estocástica derecha*, en D1 se llega a una *matriz estocástica izquierda* (“una matriz cuadrada, cada una de cuyas columnas está formada por números reales no negativos, sumando cada columna 1”). Sin embargo, a pesar de esto ambos modelos llegan a los mismos resultados, cabe preguntarse entonces: ¿cómo es posible esto?

Veamos primero el proceso llevado a cabo en D2: al tiempo que se reescribe el sistema de ecuaciones lineales en notación matricial, resultando esto en el modelo de vectores y valores propios, se hace *explícito* que este último resulta de operar con la traspuesta de la matriz de incidencia  $M'_I$  (Teorema 4.1), indicando el uso de la traspuesta en cada ocasión. Ahora observemos el proceso llevado a cabo en D1: al hacer la reescritura del sistema de ecuaciones lineales en notación matricial, se muestra inmediatamente la matriz traspuesta en uso (Fig. 5), sin señalar la trasposición de la matriz de incidencia ni las razones que la motivan. Es debido a estas diferencias en el tratamiento que se da al modelo en ambos documentos que, a pesar de que ambos arriban a la reformulación de la matriz como una estocástica, el modelo matricial resultante es relativamente diferente.

Se procede nuevamente a proponer el uso de herramientas computacionales para determinar la solución del modelo al operar con esta nueva matriz, solución que es obtenida con el propósito de contrastarla con la encontrada para el modelo previo, es decir, que se busca comprobar la **validez** del nuevo modelo. Una vez llegado al punto de la comparación se concluye que los resultados del nuevo modelo son “mejores”, esto se debe a que muestran una mayor congruencia con la situación. Recordemos que con el primer modelo los sitios 2 y 3 recibían una misma importancia a pesar de tener un diferente número de enlaces. En este nuevo modelo al sitio 3 se le asigna una mayor importancia que al 2, debido a que los enlaces que apuntan al sitio 2 (provenientes de los sitios 3, 4 y 5) ven disminuido su valor, ya que los tres sitios emiten 3 votos cada uno, de tal forma que no sólo es menos importante que el sitio 3, sino que se ve superado también por el sitio 5; mientras tanto, uno de los enlaces que recibe el sitio 3, no sólo proviene del sitio más importante de la red, sino que el mismo sólo emite dos enlaces, otorgando por ello un mayor valor a cada uno. Este nuevo modelo no sólo resuelve la situación planteada, sino que refleja de una manera más completa la situación que el modelo formulado previamente.

Siendo ésta la segunda ocasión en el documento en que se propone el empleo de herramientas computacionales, consideramos importante analizar el uso que se hace de las mismas. En ambas ocasiones estas herramientas son sugeridas con el propósito de obtener soluciones al problema planteado, para el posterior análisis de tales soluciones y por lo tanto, las emplea para **validar** el modelo. Sin embargo, podemos observar que a pesar de que se da libertad al lector de emplear el software de su elección (el autor sólo hace una sugerencia), inmediatamente se presentan los resultados esperados. Consideramos que esto tiene como propósito que sea posible comparar y validar los resultados que el lector pudiera obtener, tratando de evitar que errores en los cálculos pudieran resultar un obstáculo. Esto viene a reafirmar indirectamente el interés principal del autor, siendo que no ahonda en los procesos de cálculo necesarios para obtener

soluciones del modelo, sino que centra su atención en cómo estos procesos sirven como herramienta para **validar** el modelo a través de su análisis con respecto a la situación modelada.

Finalizando esta sección el autor hace una observación respecto a que el valor propio de la matriz  $\lambda$  resulta ser 1, expresando lo siguiente: “*el hecho de que 1 sea un valor propio de la matriz estocástica no es casualidad, ya que es fácil comprobar que el vector  $(1,1,\dots,1)$  es siempre un vector propio de toda matriz estocástica por columnas, asociado al valor propio  $\lambda = 1$* ”. Nuevamente, podemos notar que esta cuestión parece alejarse del interés del autor, por lo que no se adentra en ella, dejando al aire la pregunta ¿qué posibles dificultades o facilidades puede generar esto al momento de buscar soluciones al modelo? (**facilitar**) Es con esto que de manera similar a lo sucedido en D1 parece concluir la fase de construcción del modelo PageRank.

A continuación se procede a analizar la tercera sección de los documentos, donde se aborda la validación e implementación del modelo, que a su vez lo validará de manera práctica. Se observa también que durante la implementación surgen algunas adaptaciones al modelo, que identificamos también como pasos de la construcción de la técnica PageRank. Debido a que esta sección se mueve en gran medida dentro del contexto de la informática, es en esta parte que esperamos ver una relación más fuerte con  $I_p$ , es decir, una mayor presencia de tecnologías prácticas.

#### 4.2.4 Validación e Implementación del Modelo de Importancias en D1

En esta sección analizamos, en cada uno de los documentos, cómo se **valida** el modelo construido antes de su implementación, la cual a su vez es también una **validación** en sí misma, hecha por  $I_p$ . Recordemos que en D1 y D2 el modelo matricial estocástico es propuesto como el modelo final, aunque en D1 se hace una reinterpretación adicional del mismo, planteándolo como una cadena de Markov.

Con el propósito de **evaluar** si el modelo al que se ha llegado hasta este momento es una adecuada representación de la estructura y funcionamiento de la web, se presenta en D1 un escenario con el título: “El surfista aleatorio” que el autor utiliza para describir a grandes rasgos el comportamiento de un usuario al navegar por la Web. Pero, ¿quién es el surfista aleatorio? Se trata de un usuario cualquiera de la red, que se está desplazando entre sitios, para ello utiliza los enlaces del sitio en el que se encuentra, por lo que el siguiente sitio que visite tendrá que estar forzosamente relacionado con el actual, es decir, enlazado. Dado que no podemos conocer el objetivo de la búsqueda del usuario podemos considerar que la probabilidad de que visite cualquiera de los sitios enlazados es igual para cada uno de ellos, es decir, se trata de una distribución de probabilidad uniforme (tecnología que **facilita** la técnica). Estableciendo una relación con el modelo, el autor plantea representar la situación de la siguiente manera: “Los estados son los vértices del grafo  $G$ . La matriz  $M'$  es la **matriz de transición** del sistema: cada entrada  $m'_{ij}$  es la probabilidad de pasar del estado (vértice)  $P_j$  al estado (vértice)  $P_i$ .” Podemos reconocer en esta situación una tecnología que viene a **describir** la reformulación de la matriz de incidencia como matriz de transición, y al funcionamiento de la Web como tecnología que la **motiva**.

Se **explica** entonces que las entradas  $m'_{ij}$  serán  $1Nk$ , o bien 0, donde el denominador  $Nk$  representa el total de enlaces salientes de un sitio, sin embargo existe dentro de este razonamiento una cuestión originada en la situación: ¿Qué sucede si el usuario llega a un sitio sin enlaces salientes? Esta situación, representada por los casos en que  $m'_{ij}$  sea 0, y que en la práctica es posible e inevitable, invalidaría la interpretación que hasta ahora se ha venido construyendo, dado que un sitio sin enlaces salientes tendría una columna compuesta sólo de ceros, por lo que la matriz  $M$  dejaría de ser estocástica. Aquí podemos identificar otra tecnología práctica que describe dificultades en la implementación del modelo (**facilita**), y que a su vez **motiva** una alteración al mismo. La solución que emplea Google es **descrita** de la siguiente manera: “*la sustituimos por una columna con valores  $1n$ , así que si el surfista llega a una página de la que (antes) “no podía salir”, ahora sorteamos (uniformemente) entre todas las de la red*”. Podemos observar que de manera indirecta esto resulta en que un sitio sin enlaces “vota” por todos los demás sitios de la Web, esto constituiría una alteración no contemplada en el modelo hasta ahora, ya que no forma parte de la situación que representa originalmente ¿Sería necesario reformular el modelo tomando en cuenta esta consideración? Las condiciones del mismo modelo se encargan de dar respuesta a esta cuestión, no sólo el volumen de datos causa que  $1n$  resulte en un valor minúsculo, sino que al sumarse a la totalidad de valores de la matriz no afectará el “orden” de las importancias. De manera implícita encontramos aquí una tecnología que **facilita** la técnica, al considerar que el objeto de atención del problema no es el valor de las importancias, sino el “orden” de las mismas. Finalmente se **explica** también que esta solución permite mantener a  $M$  como estocástica.

A continuación, bajo el título: “¿Y la cuestión computacional?” el autor aborda la implementación del modelo y con ello presenta una limitación importante para el funcionamiento del mismo: el volumen de datos. La matriz empleada por Google contiene un volumen de información gigantesco, de forma que para obtener el autovalor que nos interesa (mayor en módulo a todos los demás) resultaría necesario encontrar el autovector asociado, para esto tendríamos que calcular **todos** los autovectores de la matriz para después seleccionar el que nos interesa. Esto es poco práctico, sin embargo, el autor menciona que la propia estructura del problema nos **facilita** la tarea, pasando a explicar a continuación el “*método de las potencias*”, que nos permitirá resolver el último paso de la técnica: *calcular las importancias*. En esta situación encontramos expresada una tecnología de tipo práctico que al abordar las dificultades de implementación (volumen de datos) cumple con la función de **facilitar** la implementación del modelo, a la vez que **motiva** la elección de una técnica para resolver el último paso. Se nos presenta brevemente una nota con la definición de *matriz primitiva*, sin embargo no se explica la razón de ser de la misma, por lo que su papel tecnológico (se identifica como una tecnología teórica) no es del todo claro, sin embargo se asume que el autor presenta este concepto dado que  $M'$  debe cumplir con la condición de tener un único autovalor dominante, por lo que se trata de una primitiva, aunque en esta parte del documento sólo se asume que esta condición se cumple.

NOTA. Una matriz  $A \geq 0$  es **primitiva** si tiene un único autovalor dominante. Esto ocurre cuando, por ejemplo,  $A^k > 0$  para cierto entero positivo  $k$ .

Figura 14. Definición de Matriz primitiva

Podemos observar lo anterior cuando el autor describe los elementos aportados por la estructura que considera **facilitan** la obtención de resultados a través del método, los cuales maneja como suposiciones: “*Supongamos, por simplificar, que  $A$  es diagonalizable. Los autovectores  $\{v_1, \dots, v_n\}$  son una base de  $\mathbb{R}^n$ . Están ordenados de manera que los autovalores correspondientes vayan en orden decreciente de tamaños:*”

$$\underbrace{|\lambda_1|}_{\text{autovalor dominante}} > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

Esto resulta en dos suposiciones importantes:

- Que existen las condiciones que garantizan la existencia de un autovalor mayor en valor (módulo) que todos los demás autovalores.
- Que la matriz es diagonalizable.

Al tiempo que podemos identificar ambos puntos como tecnologías de tipo teórico, se observa que al menos dentro del documento, la segunda resulta incompleta, puesto que permanece en carácter de suposición y deja al aire cuestiones como: ¿Qué condiciones deben cumplirse para que la matriz sea diagonalizable? ¿Es posible asegurar que la matriz de Google cumple con tales condiciones? En el caso de la primera suposición surgen también preguntas similares: ¿por qué es importante que exista un autovalor mayor en valor que todos los demás de la matriz? ¿Qué condiciones aseguran la existencia de dicho autovalor? Y, ¿la matriz de Google cumple con esas condiciones? Sin embargo estas preguntas son respondidas hasta cierto punto a través de un apartado que explicita parte del sustento matemático del modelo (**tecnologías teóricas**): el *Teorema de Perron* y el *Teorema de Perron-Frobenius*. El primero de éstos especifica las condiciones que aseguran la existencia de un autovalor positivo, mayor en valor absoluto a todos los demás: que se trate de una *matriz cuadrada* y que sus entradas sean positivas. Ambas condiciones están, hasta este punto, aseguradas de manera implícita (no existen valores negativos en la matriz, se aseguró previamente que no existan ceros, y las dimensiones de la matriz son  $n \times n$ ). Sin embargo, las condiciones expresadas por el segundo Teorema establecen que la matriz deberá ser irreducible, condición que no ha sido considerada hasta el momento, para lo que, en la siguiente nota, se presenta la definición a considerar (**tecnología teórica**).

NOTA. ¿Qué quiere decir irreducible?:

- 1.- no existe ninguna permutación (de filas y columnas) que transforma  $A$  en una matriz del tipo
 
$$\left( \begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array} \right),$$
 donde  $A_{11}$  y  $A_{22}$  son matrices cuadradas;
- 2.-  $(I + A)^{n-1} > 0$ .
- 3.- El grafo subyacente está fuertemente conectado.

(Obsérvese que si  $A > 0$ , entonces es  $A \geq 0$  e irreducible).

Figura 15. Definición de irreducibilidad

Haciendo un paréntesis, es importante notar que la totalidad de las implicaciones de los Teoremas para con el problema no son tratadas en profundidad. La necesidad de mantener a la matriz como estocástica se ha venido justificando por la intención de observar al modelo como cadena o matriz markoviana, sin embargo, otra de sus **motivaciones** es asegurar que la matriz sea positiva, permitiendo el uso del teorema de Perron para garantizar que el autovalor dominante (único) exista, y esto es dejado implícito. Esto puede ser observado en la sustitución de valores hecha previamente, que garantiza que no existan ceros dentro de la matriz.

Con la pregunta ¿Estamos realmente en una situación ideal? que da título a la siguiente sección, se busca evaluar al modelo dando respuesta a la pregunta que surge naturalmente tras la presentación de las definiciones de irreducibilidad: la matriz de Google ¿es irreducible? Para responder se considera el punto 3 y se retoma la estructura de la Web. Si en términos de grafos, irreducible significa que debe ser fuertemente conexo, entonces: ¿Es el grafo de la Web, fuertemente conexo? Según un estudio citado por el autor (Broder, 1999) la respuesta parece ser negativa. De las 203 millones de páginas censadas el 90% de las mismas se encuentran dentro de una componente, gigantesca, pero débilmente conexa (Fig. 16), esto nos sirve como tecnología que **motiva** una nueva adaptación al modelo.

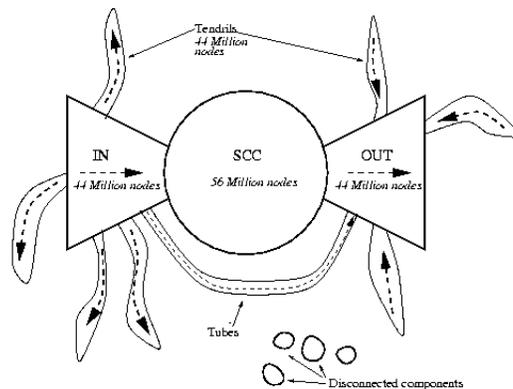


Figura 16. Grafo de la Web (1999)

Esto hace también que surjan otras cuestiones a las que el autor trata de dirigirnos con este planteamiento: ¿queda entonces invalidado el modelo? y si así fuera, ¿existe alguna solución posible? Es interesante notar que el modelo construido no necesita estar “equivocado” para ser invalidado o desechado. Sabemos hasta este momento que el modelo representa la situación, y que ofrece soluciones que resuelven la tarea, sin embargo para *Ip* resulta de mayor relevancia su implementación que su validez matemática, podemos considerar esto último como una tecnología con el papel de **facilitar** la técnica.

Es en este punto donde se nos **describe** una posible solución, que consiste en agregar a la formulación del modelo una serie de probabilidades de salida a todos los vértices de la matriz (Fig. 17). Esta solución es empleada por Google, aunque esto no es mencionado por el autor, por lo que podemos considerarla como **validada** por la práctica.

$$M'' = cM' + (1 - c) \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} (1, \dots, 1)$$

Figura 17. Nueva formulación del modelo

Se recurre nuevamente al escenario del “surfista aleatorio” para **explicar** esta formulación: “En términos del surfista aleatorio, estamos añadiendo la posibilidad de que (con probabilidad  $1 - c$ ) se “aburra” de seguir los enlaces y opte por saltar a otras páginas (con arreglo a cierta distribución de probabilidad)”. Este “salto” es la razón por la que se le conoce en algunas fuentes como “*teleportation matrix*” (matriz de teleportación). De manera similar a lo ocurrido previamente al reformular la matriz de incidencia como matriz de transición, la alteración de los valores de la matriz, que en última instancia alterarán los valores de las importancias, no llama la atención del autor, sin embargo se asume que se trata de algo natural. Esto debido a que, como se mencionó previamente, si la modificación altera a la matriz entera de manera uniforme, el orden de las importancias no se verá alterado.

Es con esta última adaptación que el documento finaliza realmente el proceso de construcción del modelo, podemos asumir que este último paso arroja el modelo terminado o final, pasando entonces el autor a ejemplificar brevemente usos de matrices no negativas en otros contextos:

- Modelos de evolución probabilística.
- Modelos dinámicos discretos (matriz de Leslie).
- Modelos económicos (modelo de producción y consumo de Leontiev).

Nuevamente recurrimos a analizar D2 para complementar el análisis de esta sección, principalmente con la intención de identificar tecnologías que no hayan sido observadas en D1.

#### 4.2.5 Validación e Implementación del Modelo de importancias en D2

El autor procede a retomar la pregunta planteada previamente: “¿habrá un único valor propio positivo, y asociado al mismo un solo vector propio positivo (solución) para cualquier matriz cuadrada con ceros y unos? Y en caso de que la respuesta sea negativa ¿cómo asegurar que esto se cumpla?” en una nueva sección titulada: “Solución del problema... ¿Unicidad?”.

Dentro de esta sección se plantea que para poder encontrar solución al problema de unicidad resultará necesario acudir al *Análisis Funcional*, y procede entonces a presentar el sustento teórico (tecnologías) necesario: el Teorema de Perron y el Teorema de Frobenius (mostrados ambos en D1).

**Teorema 5.1** (Perron, 1907).

Sea  $M$  una matriz cuadrada con todos sus coeficientes **positivos**. Entonces

- existe un valor propio simple  $\lambda > 0$  tal que  $M \cdot \mathbf{v}^t = \lambda \cdot \mathbf{v}^t$ , donde  $\mathbf{v}$  es el vector correspondiente, y tiene todas sus coordenadas **positivas**

En qué sentido se habla de PROBLEMA A RESOLVER? PORQUE LOS TEOREMAS SE DEMUESTRAN. Si se refiere a que eso se constituye en un problema, desde un punto de vista cognitivo pues, al o a qué se refiere desde el matemático, se demuestra o se valida una tesis

- *este valor propio es mayor, en módulo, que todos los demás valores propios de M*
- *cualquier otro vector propio positivo de M es un múltiplo de v*

Podemos observar que, de inicio, **el Teorema nos presenta un problema a resolver**: si la matriz representa la estructura de la Web, resultará imposible que todas sus entradas sean positivas, dado que no existirán enlaces entre todos los sitios, y aún cuando existieran, también resultaría necesario que todos los sitios se enlazaran a sí mismos, esto entra en conflicto con la situación que se está modelando. Se trata entonces de una tecnología práctica con el papel de **facilitar** la técnica. A pesar de que este Teorema nos ofrece una solución única: el vector propio positivo asociado al valor propio de mayor módulo, no podemos emplearlo, debido a que hasta este momento el modelo no cumple con las condiciones que requiere.

Se presenta entonces el Teorema obtenido por Frobenius, que es aplicable a matrices no negativas, como la que resulta del modelo, por lo que se espera que arroje alguna solución aplicable al problema:

**Teorema 5.2** (Frobenius, 1908 1912).

*Sea M una matriz cuadrada con entradas **no negativas**. Si M es **irreducible**,*

- *existe un valor propio simple  $\lambda > 0$  tal que  $M \cdot \mathbf{v}^t = \lambda \cdot \mathbf{v}^t$ , donde  $\mathbf{v}$  es el vector correspondiente, y tiene todas sus coordenadas **positivas**;*
- *este valor propio es **mayor o igual**, en módulo, que todos los demás valores propios de M;*
- *cualquier otro vector propio con entradas **no negativas** de M es un múltiplo de  $\mathbf{v}$ .*

Nos enfrentamos ahora a una nueva condición: que la matriz sea *irreducible*. Debido a su importancia para continuar el análisis, resulta necesario para abordar la definición de irreducible, tecnología teórica proveniente del Álgebra Lineal, para posteriormente presentar una proposición derivada del Teorema:

**Definición 5.3.** Una matriz M se dice **irreducible** si no existe ninguna permutación de sus filas y columnas que la transforme en otra matriz del tipo

$$\begin{pmatrix} M_{11} & M_{12} \\ \mathbf{0} & M_{22} \end{pmatrix},$$

Donde  $M_{11}$  y  $M_{22}$  son matrices cuadradas.

**Proposición 5.4.** *Si M es la matriz de incidencia de un grafo dirigido, entonces M irreducible es equivalente a que el grafo sea “fuertemente conexo”, es decir que dados dos nodos cualesquiera del mismo, es posible encontrar una sucesión de aristas que lleven de uno a otro.*

Entonces, de manera similar a D1 llegamos a la conclusión de que el modelo requiere que el grafo de la Web sea *fuertemente conexo* para que el modelo pueda aplicarse, considerando de inicio que esto es altamente improbable (tecnología que **facilita**).El

autor retoma también el estudio tratado en D1 que demuestra que esta condición no se cumple. Esto significa que ninguno de los dos Teoremas puede aplicarse a la situación, por lo que no puede garantizarse la unicidad, esto **motiva** una nueva transformación al modelo. Apareciendo las siguientes preguntas: ¿Qué hacemos entonces? ¿Qué hace Google? Para responder a estas cuestiones, se recurrirá a un concepto proveniente de la *matemática computacional*  $P(DI)$  y el *Álgebra Lineal Numérica*  $P(M)$ : una “perturbación” (tecnología teórica) .

Se señala que en ambas ramas se acostumbra trabajar con datos aproximados, siendo inevitable referirnos en este punto a la similitud que tiene esta “costumbre” con la práctica profesional de la ingeniería, en donde no es tan necesaria la exactitud de un dato o un resultado, como la **eficiencia** al resolver problemas, esta “costumbre” tiene el papel de **facilitar** la aplicación de la técnica.

Aquí lo que haremos será algo muy ingenuo pero **eficiente**, reemplazaremos nuestra matriz estocástica (que denotaremos con  $M_{I,E}$ ) por una matriz a la que haremos positivos todos sus coeficientes sumándole una matriz conveniente. El principio subyacente a esta idea es que la función “importancia” es continua, y si puedo calcularla “cerca” de la situación donde me encuentro, ya me alcanza para lo que quiero que es ordenar las importancias y no realmente calcularlas. (D’Andrea, 2012, p.11)

Se observa cómo se recalca el carácter “eficiente” de la solución que **justifica** su implementación. La importancia de la eficiencia reafirma que no resulta necesario centrar nuestra atención en el proceso y exactitud del cálculo de las importancias (recalcado esto por lo observado previamente en el uso de herramientas para llevar a cabo dichos cálculos). Dado que el foco real de nuestro interés es el orden de las mismas. Es así como se resuelve de manera efectiva el problema planteado. Simbólicamente se **describe** de la siguiente manera (Fig. 18), con un  $\varepsilon$  (mayor que) 0, muy pequeño:

$$M_{I,E}^{\varepsilon} := (1 - \varepsilon) \cdot M_{I,E} + \frac{\varepsilon}{n} \cdot \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

Figura 18. Modelo “perturbado”

El resultado de esto es, nuevamente, una matriz estocástica por filas, **explicando** que esta nueva matriz sólo tiene entradas positivas, con un valor mayor o igual a  $\varepsilon/n$ . Con esto es posible garantizar el cumplimiento de la condición establecida por el Teorema de Perron y, por lo tanto, se resuelve la cuestión de unicidad del problema, podemos decir entonces que la solución del problema es la que se obtiene según este enunciado para un  $\varepsilon$  prefijado (se especifica que en la práctica Google emplea un valor de 0.15).

A continuación se pide calcular la matriz “*perturbada*” del grafo que se ha venido trabajando (Fig. 7), con el siguiente resultado:

$$M_{0,E}^{\epsilon} = \begin{pmatrix} 0,03 & 0,03 & 0,455 & 0,03 & 0,455 \\ 0,88 & 0,03 & 0,03 & 0,03 & 0,03 \\ 0,313 & 0,313 & 0,03 & 0,313 & 0,03 \\ 0,313 & 0,313 & 0,03 & 0,03 & 0,313 \\ 0,313 & 0,313 & 0,313 & 0,03 & 0,03 \end{pmatrix}$$

Figura 19. Cálculo de la matriz “perturbada”

Posteriormente, se pide calcular el vector propio de esta matriz perturbada (sugiriendo nuevamente auxiliarnos de una herramienta) asociado al valor propio positivo de  $\lambda = 1$ , que fue calculado previamente para la matriz sin perturbar, esto finalmente entrega un orden de importancias igual al encontrado para el modelo previo. Podemos reconocer en estas actividades el mismo propósito de **validación** del modelo previamente identificado en las anteriores, al presentar solamente los resultados esperados de manera que el lector pueda comprobar los que haya obtenido por su cuenta, manteniendo el enfoque sobre el significado que tienen estos resultados para con el modelo y la tarea que resuelve, particularmente al final, al comparar solamente el *orden* de las importancias que arrojan ambos modelos sin prestar atención a los valores en sí mismos.

Hasta este momento tenemos un modelo que representa adecuadamente la situación que enfrentamos, es decir, que funciona bajo todas las condiciones que han surgido del análisis de la misma. Se trata de un modelo que ha contrastado sus resultados con la realidad que está tratando de modelar, validándolos en sucesivas ocasiones, y que en los momentos en que no ha cumplido con las condiciones planteadas ha sufrido modificaciones o reinterpretaciones que le han permitido adaptarse al problema que busca representar. Sin embargo, surge una cuestión de suma importancia para la ingeniería (*Ip*): ¿es **éste** un modelo **eficiente**?

Desde ya desaconsejamos a cualquier optimista intentar utilizar las técnicas aprendidas (o por aprender) en el curso de álgebra que involucrarían calcular un polinomio característico de grado mayor que un billón, encontrar todas sus raíces, y detectar entre todas ellas la que tiene módulo máximo, y luego resolver un sistema lineal enorme para calcular el vector de importancias. Incluso si supiéramos el valor de  $\lambda$  (que en los casos estocásticos es 1), resolver un sistema lineal del orden de un billón es una tarea dantesca que no puede ser realizada en poco tiempo ni siquiera por los ordenadores más rápidos que hay disponibles en este momento. (D’Andrea, 2012, p.12)

Al resolver tales dificultades, se cumple con el papel tecnológico de **facilitar** la técnica, explorando la problemática que surge al implementarla, reforzando al mismo tiempo a la eficiencia como un aspecto de vital importancia en la práctica profesional de la ingeniería. Por ello, aun cuando el modelo sea congruente con la realidad y los resultados que ofrece sean correctos, si el proceso de cálculo de las soluciones consume tantos recursos, como puede preverse tras lo expresado por el autor, no resultaría aplicable en la práctica y tendría que ser desechado, ¿qué sucede entonces?

Se nos presenta entonces el *método de las potencias*, descrito por el autor como “*en apariencia bastante ingenuo de enunciar pero computacionalmente muy efectivo*”, esto

último resulta de vital importancia para nosotros, puesto que se trata entonces de un método (o praxeología) que ha sido **validado** ya por Ip, además de que la “efectividad” demostrada **motiva** también su uso. Con esto podemos identificar el último paso de la técnica: *calcular las importancias*. Éste podría ser observado a su vez como una praxeología puntual, en cuyo caso el método de las potencias constituiría su técnica.

El método es descrito por el autor de la siguiente manera:

Si una matriz cuadrada  $M$  es diagonalizable y tiene todos sus vectores propios  $\{v_1, \dots, v_n\}$  numerados de tal manera que los autovalores correspondientes cumplan lo siguiente

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|,$$

partiendo de  $v_0 \geq 0$  tal que

$$v_0 = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

con  $\alpha \neq 0$ , entonces se tendrá

$$M^k v_0 = \alpha_1 \lambda_1^k v_1 + \dots + \alpha_n \lambda_n^k v_n.$$

Luego

$$\lim_{k \rightarrow \infty} M^k v_0 = \alpha_1 \lambda_1^k v_1$$

Consideramos que este método podría ser de gran importancia, debido a que resulta de utilidad para obtener información específica sobre conjuntos de datos contenidos en estructuras matriciales. Se pueden encontrar en documentos pertenecientes a  $E(DI)$ , por ejemplo menciones a un “método de las potencias” en una versión mucho más sencilla que permite determinar la existencia de caminos entre nodos:

*El número de caminos  $C_{i,j}(k)$ , atravesando  $k$  aristas desde el nodo  $i$  al nodo  $j$ , viene dado por un elemento de la potencia  $k$ -ésima de la matriz de adyacencia:*

$$C_{i,j}(k) = [A^k]_{ij}$$

Es con la presentación de este método que se concluye el documento, comentando el autor que **éste** es el método que utiliza Google en la práctica (reafirmando la validación por parte de Ip) con resultados *razonables*, ofreciéndonos un último recordatorio de que el objetivo es obtener resultados eficientes con respecto a los recursos que se emplean para llegar a ellos, y que **éstos** resuelvan la tarea. Es interesante notar que el método **sólo** es descrito, no se trata de validar el mismo, dado que como se comentó previamente, **éste** ya se encuentra validado por la disciplina  $P(DI)$ -Ip, además de que **éste** es un método para obtener resultados del modelo PageRank, su importancia radica en que se utiliza para solucionar las dificultades de *implementación* del modelo, dificultades que por sí mismas no pueden solucionarse y que entonces cumplen también con el papel de **motivar** el método. Como parte de las reflexiones finales el autor nos recuerda una de las motivaciones más importantes de su trabajo: “*cualquier persona que quiera trabajar resolviendo problemas necesita de recursos, de herramientas. Y*

cada una de estas áreas de la matemática y de la informática tiene que ser para el ingeniero precisamente eso, una herramienta”.

### 4.3 Del análisis praxeológico de los buscadores web hacia el diseño de una actividad didáctica

Una vez realizado el análisis praxeológico, resultó posible identificar la existencia de varias praxeologías relacionadas, conformando una arborescencia praxeológica, la cual nos permite organizarlas en niveles de complejidad, y que será descrita. Sin embargo, no todas las praxeologías identificadas pueden ser consideradas para un diseño de actividades, por lo que se seleccionan las praxeologías “viables” y para ello se realizó a su vez un análisis sobre  $E(M)$  para justificar la viabilidad de tal elección.

#### 4.3.1 Arborescencia praxeológica

El análisis praxeológico presentado sobre los buscadores web permite evidenciar la existencia de una arborescencia praxeológica que puede ser vista esquemáticamente de la siguiente manera:

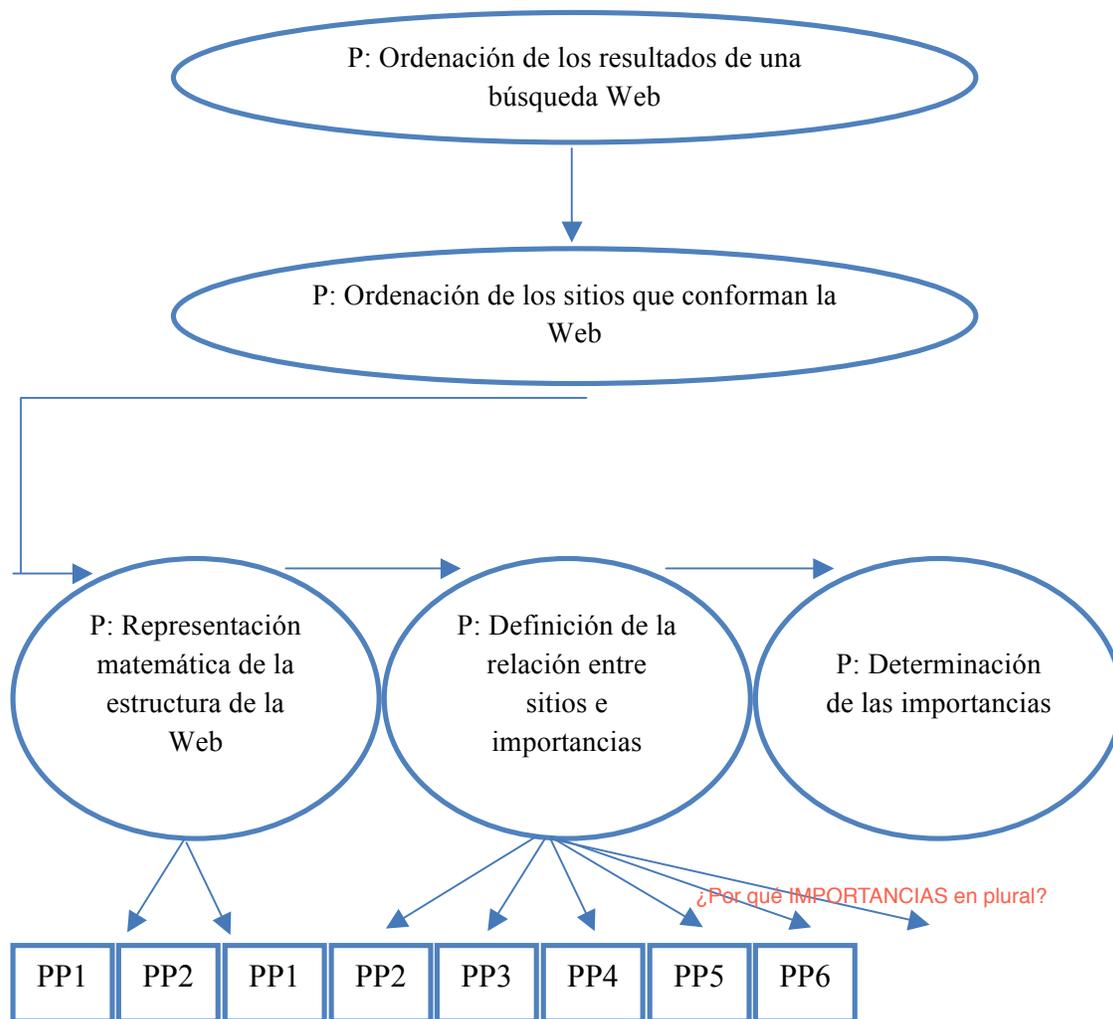


Figura 20. Arborescencia praxeológica

Las praxeologías en la base, denotadas como PP, son modelos reconocidos como praxeologías puntuales que por cuestiones de espacio no son nombradas, pero son

tratadas en la descripción individual de cada praxeología. Esta representación esquemática permite observar ciertos niveles de complejidad, que se corresponden con los niveles de organización matemática:

- Se observan praxeologías en la base, que se resuelven (cada una) con una sola técnica. Se trata de *praxeologías puntuales*.
- En el siguiente nivel se observan dos praxeologías que agrupan a las puntuales, se trata de tareas que pueden resolverse con más de una técnica, pero que comparten la misma tecnología, por lo tanto se trata de *praxeologías locales*.

Esto haría suponer que en el siguiente nivel estamos tratando con una *praxeología regional*, que sería *Estructura de Datos*, considerando que agrupa al menos dos praxeologías locales. Sin embargo no disponemos de suficiente información como para asegurarlo, es decir, que desconocemos si existe una teoría única que sustente la tecnología de las praxeologías locales. Esto se debe principalmente a que en dicho nivel se trata de una praxeología no matemática, así como en el nivel superior.

Observamos entonces, que en los dos niveles superiores encontramos praxeologías no matemáticas, que nos permiten observar el proceso de análisis que parte desde la problemática inicial, el cuál motiva, o es la razón de ser, de todas las praxeologías que integran la arborescencia.

Identificamos que cada praxeología requiere que se realicen ciertas tareas, las cuales conforman el nivel inmediatamente inferior, es decir que para *ordenar los resultados de una búsqueda Web*, se identifica que debe *ordenarse el conjunto de sitios que conforman la Web*, y para realizar esta última nos encontramos con una técnica separada en tres pasos, que son observados como praxeologías para posibilitar un análisis más fino. Esta separación sólo es de tipo metodológico, ya que la sólida relación entre ellas causa que no puedan ser realizadas de manera independiente. La *Representación matemática de la estructura de la Web* trata de resolver una necesidad de análisis y comprensión de la estructura, que sirve de sustento tecnológico vital para la *Definición de la relación entre sitios e importancias*, por lo que, de resolverse de manera completamente independiente, perdería por completo su razón de ser. A su vez, en esta última se construye el modelo cuyas tecnologías sustentan a su vez la técnica utilizada en la *Determinación de las importancias*. Esto nos resulta de vital importancia al pensar en un diseño didáctico, dado que nos permiten identificar aquellas tecnologías que no pueden dejarse fuera del mismo, además de que nos muestra que será necesario al menos trabajar con dos de las tres praxeologías en el proceso de estudio.

A continuación se presentan las praxeologías previamente esquematizadas.

### **Praxeología 1. Ordenación de los resultados de una búsqueda Web**

#### **T: Ordenar los resultados de una búsqueda Web**

##### **Técnica:**

- Se definen categorías de búsqueda, dependiendo de la cantidad de términos a buscar: **simple**, de una sola palabra, y **combinada**, de dos o más.

- Se definen criterios de ordenación, en el caso de la búsqueda simple: la ubicación y características del término dentro de la página (título, negrita, tamaño de fuente, etc.). En el caso de la búsqueda combinada: además del criterio de búsqueda simple para cada palabra, la “cercanía”/“lejanía” entre las palabras dentro de la página.
- Ordenar los sitios de la Web.

### **Tecnología:**

- Las búsquedas consisten en una o más palabras, las cuales, dentro de la página tienen un tipo y tamaño de fuente, pueden estar en negrita, cursiva y/o subrayado, así como ser parte del título o del texto principal.
- El conjunto de sitios está enlistado en una base de datos, que los cataloga de acuerdo a ciertas palabras clave de manera constante.
- El ordenar los sitios Web permitirá tener un orden de búsqueda predeterminado, que disminuiría el tiempo de ejecución.
- Si se establece un orden adecuado para los resultados, tendremos un buen nivel de confianza de que los primeros 10 resultados de la búsqueda sean suficientes para el usuario.

Se observa que **ésta** es una tarea no matemática, en la cual no aparecen elementos tecnológicos teóricos. Se asume que la técnica identificada está incompleta, debido a que el resto de la misma no influye o no se relaciona con el objetivo de los documentos. Sin embargo, identificamos como parte de la técnica: *ordenar los sitios de la Web*. La importancia de este paso radica en que encontramos al método PageRank como su vía de resolución, por lo que este paso se vuelve objeto central del análisis. Para facilitar y permitirnos analizar los elementos tecnológicos del método, decidimos considerar este paso como una sub-tarea, cuya técnica será el método PageRank.

### **Praxeología 1.1. Ordenación de los sitios que conforman la Web**

#### **T: Ordenar los sitios de la Web**

##### **Técnica:**

1. Se reconoce el criterio de ordenación a utilizar como, *importancia de un sitio*. Y se asignan variables, sitios:  $P_1, \dots, P_n$  e importancias:  $x_1, \dots, x_n$
2. Se representa matemáticamente la estructura de la Web.
3. Se define la relación entre sitios e importancias, es decir, como se asignará la importancia de cada sitio.
4. Se determinan, o calculan, las importancias de cada sitio.

##### **Tecnología:**

- La importancia de un sitio como criterio de ordenación proviene de la idea de “autoridad”, si un sitio es importante, significa que su contenido es valioso, y que debería ser consultado. Una analogía posible es un profesor, el cual tendrá un cierto grado de “importancia” en función, en este caso, de las credenciales

que lo avalan. Es por esto que resulta importante definir como se asignará la importancia de un sitio.

Estos son los pasos que identificamos como la técnica, en los cuales podemos observar también el proceso de construcción y resolución del método PageRank. Así como previamente se identificó *ordenar los sitios de la Web* como sub-tarea de manera que pudiera ser analizada, dentro de la técnica PageRank observamos que tres de sus pasos (pasos 2, 3 y 4) se mostraban susceptibles de recibir el mismo tratamiento. De manera que fueron considerados como sub-tareas individuales, esto permitió realizar un análisis más fino de los elementos tecnológicos que son de nuestro interés.

### Praxeología 1.1.1 **Representar matemáticamente la estructura de la Web (RMEW)**

#### **T: Representar matemáticamente la estructura de la Web.**

##### **Técnica:**

- Se identifican los elementos de la Web a representar: sitios y enlaces.
- Se representan los elementos mostrados a través de un *grafo dirigido*  $G$ : sitios como *vértices* o *nodos*, y enlaces como *aristas orientadas*. (PP1)
- Se representan los elementos mostrados a través de una *matriz de incidencia*: sitios emisores como filas, sitios receptores como columnas, valores de 0 si no hay un enlace entre ambos y 1 si existe. (PP2)

##### **Tecnología Práctica:**

- La Web está constituida por un conjunto de sitios, que se relacionan con el resto a través de direcciones o enlaces, estos enlaces sólo pueden recorrerse en una dirección (del emisor al receptor), en caso de querer hacer un recorrido en dirección contraria será necesario otro enlace (del receptor al emisor). Además, un sitio no se enlaza a sí mismo, ya que esto no le representa ninguna utilidad práctica.
- El *grafo* debe ser *dirigido* debido a que de esta manera puede representar la unidireccionalidad de los enlaces a través de *aristas orientadas*.
- El volumen de datos (número de sitios y enlaces entre ellos) hace que el grafo tenga dimensiones masivas, por lo que resultaría demasiado costoso de manipular. Aunque es útil para tener una mejor comprensión de las relaciones entre sitios, una representación matricial es mucho más eficiente.
- El uso de matrices es recomendado por sobre otras estructuras para representar un conjunto de datos por las facilidades de operación que permite, así como un menor gasto de recursos: tiempo de ejecución, espacio de almacenamiento, etc.
- Es necesario representar matemáticamente la estructura de la Web para poder utilizar esta información para ordenarla.

##### **Tecnología Teórica:**

- Grafo dirigido.
- Matriz de Incidencia.
- Matriz Cuadrada.

## Teoría:

- Matemáticas Discretas
- Estructura de Datos

La función principal de esta tarea es permitir un mejor análisis y comprensión de la forma en que se relacionan los sitios, apareciendo elementos provenientes de Matemáticas Discretas  $P(M)-E(M)$  que han sido traspuestos por  $P(DI)-E(DI)$ , y que serán utilizados en el resto de la técnica PageRank. Es decir, que si retomamos esta tarea como paso de la técnica, estaríamos hablando de una técnica tecnológica, que se va justificando a sí misma. Las dos representaciones, grafo dirigido y matriz de incidencia, pueden ser vistas como dos modelos individuales, o como dos técnicas que permiten resolver la misma tarea. Por lo tanto, podemos considerar que estamos observando dos praxeologías puntuales, lo que convertiría a ésta en una praxeología local.

## Praxeología 1.1.2. Definir criterio de importancias (DCI)

T: Definir la relación entre sitios e importancias.

### Técnica:

- Se propone una relación en base a la estructura de la Web: *La importancia de un sitio es proporcional a la cantidad de enlaces que recibe*. Es decir, que la importancia está en función de la sumatoria de enlaces. (PP1)
- Se propone una segunda aproximación a la relación tomando en cuenta la importancia del sitio que emite el enlace:  $x_j$  es proporcional a la suma de las importancias de las páginas que enlazan con  $P_j$  y se representa con un sistema de ecuaciones lineales, tantas como sitios existan. (PP2)
- Se reescribe el sistema de ecuaciones lineales en notación matricial. (PP3)
- Se reformula como un modelo de valores y vectores propios. (PP4)
- Se modifican los valores de la *matriz de incidencia* (reemplazando los 1 por  $1/(\#$  de enlaces emitidos por  $P_i$ ), transformándola en una *matriz de transición*. (PP5)
- Se reemplaza la *matriz de transición*  $M_{I,E}$  por una matriz a la que haremos positivos todos sus coeficientes sumándole una matriz conveniente, considerando un  $\varepsilon > 0$  muy pequeño, resultando en la *matriz perturbada*. (PP6)

$$M_{I,E}^{\varepsilon} := (1 - \varepsilon) \cdot M_{I,E} + \frac{\varepsilon}{n} \cdot \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

### Tecnología Práctica:

- Cuando un sitio enlaza a otro es equivalente a “referenciarlo”. Podemos entonces suponer que si un sitio es referenciado por muchos es porque su contenido es “bueno” o “confiable”, por lo tanto que mientras más referencias, “importancia”, tenga un sitio en particular, mayor es la probabilidad de que tenga información valiosa.

- Si sólo se considera la cantidad de enlaces para determinar la importancia de un sitio, ésta podría manipularse tan sólo con crear numerosos sitios que lo enlacen. Además, el considerar el número de enlaces implica que todos los enlaces tienen el mismo valor, sin embargo, es muy diferente recibir un enlace de una página nueva a recibir uno de un sitio como Amazon o Microsoft, que sabemos son sitios importantes, el valor de un enlace depende de quien lo emita.
- Computacionalmente, cuando se tiene la opción de emplear una representación matricial se da preferencia a ésta. Por lo que en caso de que sea posible se buscará llevar a cabo dicha reinterpretación.
- Si en las filas de la matriz de incidencia uno puede leer cuántos enlaces salen de una página dada, justamente en las columnas aparecerán tantos unos como enlaces haya hacia la página indexada por esa columna.
- Si una página tiene un solo enlace, este enlace vale lo mismo que cualquier otro enlace de otra página que produzca un millón de enlaces. Es como -si bien producir enlaces desde mi propia página no aumenta mi importancia- cuantos más enlaces produce mi página, más afecta a toda la red, esto se soluciona disminuyendo el “poder de voto” total de cada sitio a uno y dividiendo ese valor entre la cantidad de enlaces emitidos.
- Para la ingeniería la eficiencia en la resolución de un problema tiene un peso mucho mayor que la exactitud, es decir, se prefiere una técnica que arroje resultados aceptables por sobre de una que produzca resultados sumamente exactos pero que consuma una mayor cantidad de recursos.
- Para que una técnica se considere confiable, es necesario asegurar que consistentemente entregará el mismo resultado, es decir, que garantice que siempre encontrará soluciones únicas.
- Así como no se necesitan valores exactos, el cálculo de las importancias no es el centro de interés, sino en el orden que éstas presentan. Por esto, sólo es necesario calcular una aproximación a las mismas que nos permita determinar tal orden.
- Se utilizan herramientas computacionales para realizar el cálculo de los valores y vectores propios, analizando los resultados obtenidos para evaluar si el modelo que los arroja representa de manera adecuada la situación.
- “El surfista aleatorio” Un surfista en la red está, en cierto instante de tiempo, en la página  $P_k$ . En el siguiente instante de tiempo, estará en una página de entre las posibles (aquéllas a las que envía  $P_k$ ). La elección de una u otra sigue una distribución de probabilidad uniforme (esto es, probabilidad  $1/N_k$  para cada una

$$\left( \begin{array}{c} m'_{1,k} \\ \vdots \\ m'_{i,k} \\ \vdots \\ m'_{n,k} \end{array} \right) \left( \begin{array}{c} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array} \right) = \left( \begin{array}{c} m'_{1,k} \\ \vdots \\ m'_{i,k} \\ \vdots \\ m'_{n,k} \end{array} \right)$$

de las admisibles).

- “Adaptación-no enlaces salientes”

### Tecnología Teórica:

- Matriz de Incidencia

- Sistema de ecuaciones lineales
- Matriz Traspuesta
- Si  $M_I$  es la matriz de incidencia del grafo de internet, y  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}_{\geq 0}^n$  el vector de importancias, entonces se cumple

$$M_I^t \mathbf{x}^t = \lambda * \mathbf{x}^t$$

donde  $\lambda \in \mathbb{R}_{>0}$  es la constante de proporcionalidad.

- Dados una matriz cuadrada  $M$  de tamaño  $n \times n$ , un vector no nulo  $\mathbf{x} \in \mathbb{R}^n$  (o  $\mathbb{C}^n$ ) y un número  $\lambda \in \mathbb{R}$  (o  $\mathbb{C}$ ), el vector  $\mathbf{x}$  se dice *vector propio* de  $M$  con *valor propio asociado*  $\lambda$  si y solo si se verifica

$$M\mathbf{x}^t = \lambda * \mathbf{x}^t$$

- El vector de importancias de las páginas web es un vector propio (positivo) de la matriz  $M_I^t$ , y la constante de proporcionalidad  $\lambda$  es el valor propio asociado a este vector.
- Valores y Vectores propios
- Matriz estocástica
- Matriz estocástica izquierda y derecha
- Los estados son los vértices del grafo  $G$ . La matriz  $M^t$  es la **matriz de transición** del sistema: cada entrada  $m^t_{ij}$  es la probabilidad de pasar del estado (vértice)  $P_j$  al estado (vértice)  $P_i$ .
- Perturbación
- **Teorema 5.1** (Perron, 1907).

*Sea  $M$  una matriz cuadrada con todos sus coeficientes **positivos**.*

*Entonces existe un valor propio simple  $\lambda > 0$  tal que  $M \cdot \mathbf{v}^t = \lambda \cdot \mathbf{v}^t$ , donde  $\mathbf{v}$  es el vector correspondiente, y tiene todas sus coordenadas **positivas**; este valor propio es mayor, en módulo, que todos los demás valores propios de  $M$ ;*

*cualquier otro vector propio positivo de  $M$  es un múltiplo de  $\mathbf{v}$*

- **Teorema 5.2** (Frobenius, 1908 1912).

*Sea  $M$  una matriz cuadrada con entradas **no negativas**. Si  $M$  es **irreducible**, existe un valor propio simple  $\lambda > 0$  tal que  $M \cdot \mathbf{v}^t = \lambda \cdot \mathbf{v}^t$ , donde  $\mathbf{v}$  es el vector correspondiente, y tiene todas sus coordenadas **positivas**; este valor propio es **mayor o igual**, en módulo, que todos los demás valores propios de  $M$ ; cualquier otro vector propio con entradas **no negativas** de  $M$  es un múltiplo de  $\mathbf{v}$ .*

- Una matriz  $M$  se dice **irreducible** si no existe ninguna permutación de sus filas y columnas que la transforme en otra matriz del tipo

$$\begin{pmatrix} M_{11} & M_{12} \\ \mathbf{0} & M_{22} \end{pmatrix},$$

Donde  $M_{11}$  y  $M_{22}$  son matrices cuadradas.

- Si  $M$  es la matriz de incidencia de un grafo dirigido, entonces  $M$  irreducible es equivalente a que el grafo sea “fuertemente conexo”, es decir que dados dos nodos cualesquiera del mismo, es posible encontrar una sucesión de aristas que lleven de uno a otro.

- Álgebra Lineal
- Estructura de Datos
- Análisis Funcional

Es dentro de esta tarea que se desenvuelve un proceso de modelación complejo que nos permite observar las características planteadas como propias de la actividad ingenieril. Dentro de cada paso existe una forma de **evaluación** que **valida** al modelo, y en cada paso se determinan ciertas condiciones, ya sean prácticas o teóricas, que **motivan** modificaciones al mismo, cada paso es entonces un modelo funcional y válido, para ciertas condiciones, por lo que podría funcionar en otros contextos en que no existan aquellas restricciones que provocan la siguiente adaptación. Observamos entonces, como se trata de un proceso de refinación, que no sólo trata de adaptar el modelo a la situación, sino que trata de mejorar su eficiencia, para esto el ingeniero no duda en despreciar aspectos de exactitud, con tal de que el modelo pueda ser resuelto de manera mucho más sencilla y con menor gasto de recursos.

La manera en que están estructurados estos pasos nos permite hacer la misma observación que en la praxeología anterior, podemos observar cada uno de ellos, cada modelo, como una praxeología puntual, es decir, que cada modelo es una diferente técnica que puede resolver finalmente la misma tarea, lo que causa que esta praxeología se convierta en una local.

### Praxeología 1.1.3. **Determinar las importancias (DI)**

**T: Determinar/calcular las importancias**

**Técnica:**

- Partiendo de  $\mathbf{v}_0 \geq 0$  tal que
 
$$\mathbf{v}_0 = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n,$$
- con  $\alpha \neq 0$ , entonces se tendrá
 
$$M^k \mathbf{v}_0 = \alpha_1 \lambda_1^k \mathbf{v}_1 + \dots + \alpha_n \lambda_n^k \mathbf{v}_n.$$
- Luego
 
$$\lim_{k \rightarrow \infty} M^k \mathbf{v}_0 \approx \lambda_1 \mathbf{v}_1$$

**Tecnología Práctica:**

- Desde ya desaconsejamos a cualquier optimista intentar utilizar las técnicas aprendidas (o por aprender) en el curso de álgebra que involucrarían calcular un polinomio característico de grado mayor que un billón, encontrar todas sus raíces, y detectar entre todas ellas la que tiene módulo máximo, y luego resolver un sistema lineal enorme para calcular el vector de importancias. Incluso si supiéramos el valor de  $\lambda$  (que en los casos estocásticos es 1), resolver un sistema lineal del orden de un billón es una tarea dantesca que no puede ser realizada en poco tiempo ni siquiera por los ordenadores más rápidos que hay disponibles en este momento.

**Tecnología Teórica:**

- Para que el método pueda ser aplicado, debe tratarse de una matriz cuadrada  $M$  diagonalizable y que tenga todos sus vectores propios  $\{v_1, \dots, v_n\}$  numerados de tal manera que los autovalores correspondientes cumplan lo siguiente:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

Esta tarea se muestra como la más corta de las tres, una de las razones para esto es que las condiciones necesarias (justificaciones) para que la técnica pueda ser utilizada efectivamente han sido garantizadas en las dos tareas anteriores. Que junto a ésta, son pasos de la técnica PageRank, compartiendo las tecnologías que los sustentan.

Esta praxeología es un caso especial, dentro del análisis identificamos el modelo de importancias final, y en función del mismo y sus tecnologías, se elige el método que conforma la técnica de esta praxeología, sin embargo, en caso de que el modelo de importancias a trabajar fuera alguno de los anteriores, sería una técnica de resolución distinta. Es por esto que, a pesar que en el análisis pueda mostrarse como una praxeología puntual, observamos que en realidad se trata de una local.

#### **4.3.2 Elección de una praxeología para el diseño de la actividad didáctica y su vinculación con otras praxeologías de modelación**

Dado que las praxeologías 1 y 1.1 fueron reconocidas como no matemáticas, se buscó determinar entre las praxeologías: RMEW (Representar matemáticamente la estructura de la Web), DCI (Definir criterio de importancias) y DI (Determinar las importancias), aquellas que pudieran ser llevadas al aula. Esta decisión requiere un análisis previo de las asignaturas que conforman a  $E(M)$ , con el propósito de encontrar las posibles correspondencias entre praxeologías identificadas previamente como puntuales y, praxeologías matemáticas de  $E(M)$ , lo que finalmente permitiría asumir que las praxeologías locales en las que están articuladas son viables de ser reconstruidas en un proceso de estudio en el aula.

Como producto de ese análisis, se identificaron praxeologías matemáticas correspondientes, en las asignaturas de Matemáticas Discretas y Álgebra Lineal, lo que nos permitió seleccionar las praxeologías RMEW y DCI. Cabe señalar que, aunque en el caso de la praxeología RMEW ambas praxeologías puntuales que la conforman pueden ser utilizadas para el diseño de una actividad, no es así con la praxeología DCI, que articula seis praxeologías puntuales, de las cuales sólo las primeras dos encuentran correspondencia en  $E(M)$ . ¿Por qué es esta correspondencia tan importante?

El objetivo principal de este trabajo es atacar la desconexión existente entre  $E(M)$  y  $E(DI)/Ip$ , mostrando al estudiante el funcionamiento y utilidad de las matemáticas dentro de su especialidad, y el encontrar esta relación entre las praxeologías de  $Ip$  con las que habitan en  $E(M)$  nos permitirá hacerlo. Así mismo esta relación nos permite asumir que es posible llevar un proceso de reconstrucción de las praxeologías de  $Ip$  a  $E(M)$  a través de una organización didáctica. Para esto conviene recordar que una praxeología matemática está asociada a una organización didáctica en un cierto nivel de determinación, que describe el proceso de estudio (construcción, reconstrucción) de dicha praxeología, por lo tanto el determinar que una praxeología extra-matemática tiene correspondencia con una matemática, nos permite asegurar que también puede

asociarse a una organización didáctica, la cual podría ser aplicada en el mismo nivel. Esto puede representarse de la manera siguiente:

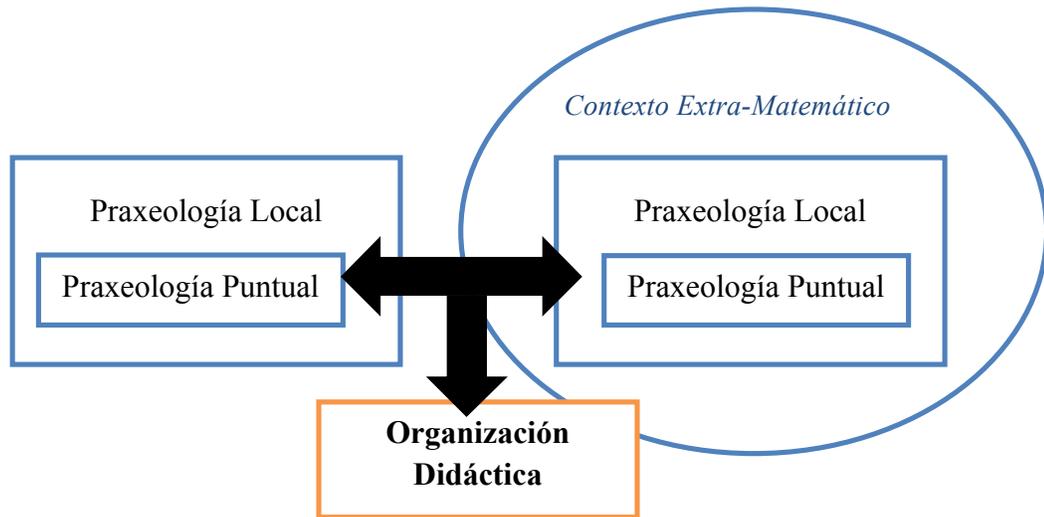
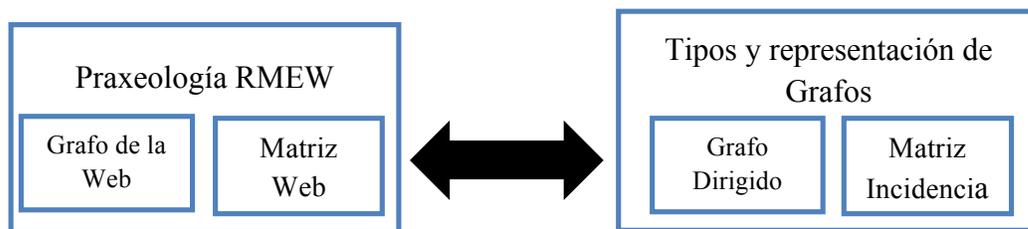
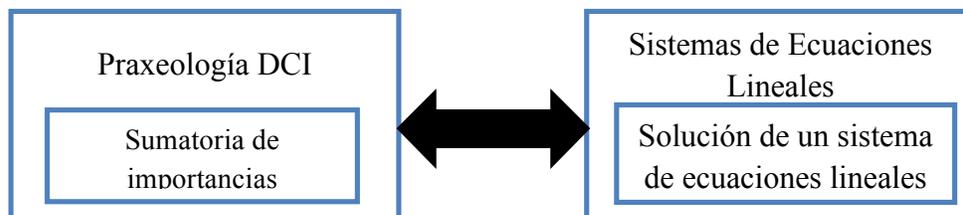


Figura 21. Niveles de co-determinación y organización didáctica

Pudo determinarse que esta relación existe entre las siguientes praxeologías:

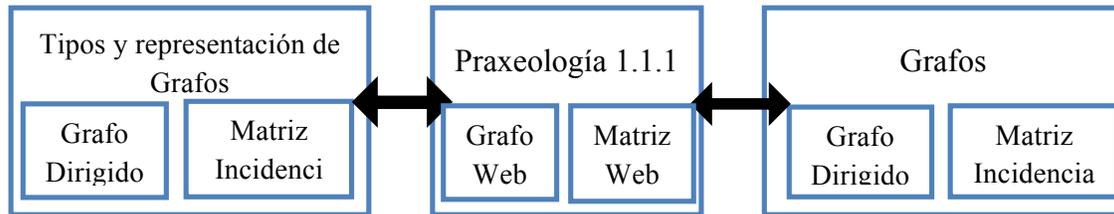


Donde la praxeología local “Tipos y representación de grafos”, pertenece a la praxeología regional “Teoría de Grafos” y ésta a su vez se inscribe en la praxeología global “Matemáticas Discretas”.



Donde la praxeología local “Sistemas de ecuaciones lineales”, pertenece a la praxeología regional “Álgebra lineal”, que se encuentra inscrita en la praxeología global “Álgebra”. Sin embargo, cabe notar que esta relación no es completamente sólida, se considera suficiente que el sustento tecnológico teórico de “Sumatoria de importancias” provenga de “Solución de un sistema de ecuaciones lineales”. Sin embargo, cabe hacer notar, que en caso de que la actividad se centre en el trabajo sobre el modelo de sumatoria de importancias, la solución de un sistema de ecuaciones lineales se convertiría en la técnica de solución de la praxeología de “Determinación de importancias”.

Durante el análisis, se hizo una revisión de los programas de aquellas asignaturas de especialidad  $E(DI)$  que tienen como requisito las asignaturas de Matemáticas Discretas y Álgebra Lineal, lo que permitió observar que, en el caso de la praxeología RMEW, no sólo existía una relación con  $E(M)$ , sino también con  $E(DI)$  en la asignatura de Estructura de Datos. Lo esquematizamos de la manera siguiente:



Donde la praxeología local “Grafos”, pertenece a la praxeología regional “Estructuras no lineales”, que se inscribe dentro de la praxeología global “Estructura de Datos”.

Esta relación permite colocar a la praxeología RMEW como punto de articulación para una relación existente entre  $E(M)$  y  $E(DI)$ , y dado que dicha praxeología fue extraída de un contexto de  $Ip$ , nos permitirá cumplir con el objetivo de generar una relación entre las 3 instituciones que forman parte de la formación del ingeniero en sistemas computacionales. Una vez establecido esto, y debido a que la praxeología DCI utiliza a la praxeología RMEW como sustento tecnológico y que permite establecer una relación con otra instancia de  $E(M)$ , se decidió que ambas constituirían un sustento conveniente para el diseño de una actividad didáctica que pudiera ser validada en un ambiente escolar.

#### 4.4 Otros contextos de uso

Una revisión al documento “*Standing in the shoulders of giants*” (Franceschet, 2011) permitió identificar otros contextos en los que estos modelos (praxeología RMEW y DCI) han sido utilizados. Estos contextos provienen de áreas diversas, tales como:

- **Bibliometría:** Modelo de Pinski y Narin, “Relevancia de una publicación”. Modelo que utiliza la relación entre artículos de investigación, dada por las referencias entre ellos: “Mientras más citado es el artículo, más importante es”, pero solo considera la cantidad de citas.
- **Sociometría:** Modelo de Seeley, “prestigio de una persona”; Modelo de Katz, “estatus de una persona”; Modelo de Hubbel. El modelo de Seeley es similar al anterior, “mientras más respaldada sea una persona (o mientras más personas lo consideren importante) más prestigiosa es”, nuevamente se trata de considerar solamente la cantidad de personas que muestran respaldo. En cambio, el modelo de Hubbel especifica “la popularidad de una persona depende de la cantidad de amistades que tenga, y la popularidad de sus amigos”, se trata del mismo modelo que utiliza PageRank en un contexto distinto.
- **Econometría:** Modelo de Leontief, “Input, Output”. Este modelo pertenece al área de Economía, salta el uso del Grafo, utilizando directamente una representación matricial para analizar las relaciones entre

sectores económicos, llamadas industrias, y su comportamiento, “una industria productiva es una que recibe insumos de otras empresas productivas”.

- Deportes: Modelo de Wei, “Ranking deportivo”.  
Este modelo, implementado al ranking de equipos de basquetbol, “un equipo es bueno, si tiene muchas victorias, y derrota a otros buenos equipos”, sin embargo, tiene como característica esencial, que un equipo puede enfrentarse más de una vez al mismo rival, por lo que el modelo debe considerarlo.
- Computación: Modelo de Kleinberg, “HITS”.  
Observa a la Web en términos de Autoridades y Hubs, estableciendo que, “buenas autoridades son páginas que son enlazadas por buenos Hubs, y los buenos Hubs son páginas que enlazan a buenas autoridades”, se trata de un modelo competidor de PageRank, con un alto nivel de complejidad.

Se considera que estos contextos (a excepción de “HITS” debido a su complejidad, que requeriría de un análisis praxeológico propio, y el modelo de Leontief, debido a la necesidad de una fuerte inmersión en el contexto de la economía, que resulta completamente ajeno al estudiante de ingeniería en sistemas computacionales, aunque no se descarta su posible utilidad) podrían resultar de utilidad para el diseño de actividades didácticas, mostrando al estudiante que el modelo no tiene una única utilidad, y que con las adaptaciones correctas puede ser empleado en áreas diversas. Esto último tendría suma importancia, para desarrollar en el estudiante la capacidad de adaptar e interpretar modelos, y que no sólo los aprenda como herramientas mecánicas.

Totalmente de acuerdo... Ahora la cuestión, es... si es necesario analizar un contexto como es en el caso de ingeniero... se entiende que programa y entonces... tiene que llevar a cabo todas esas acciones? o es la teoría de qué lo hizo. O acaso es parte de la TAD analizar niveles de complejidad? por qué?

#### 4.5 Conclusión

El análisis presentado en este capítulo, permite evidenciar la **complejidad asociada al analizar un contexto extra-matemático**, como es el caso de la programación web. El análisis del desarrollo de la técnica PageRank muestra una arborescencia praxeológica que describe la actividad de modelación a través de múltiples praxeologías de programación y mixtas. Esta descomposición de la actividad puede permitir la identificación de diferentes niveles de complejidad en las praxeologías, equiparándose a los niveles de organización matemática, puntual, local, regional, global y disciplinar. Esto se muestra en la descomposición de la técnica de PageRank en tres praxeologías: representar matemáticamente la estructura de la Web, determinar criterio de importancias y calcular las importancias. Durante el análisis de estas praxeologías se les reconoció como locales, tras identificar en sus procesos de construcción o resolución múltiples praxeologías de tipo puntual. A su vez, el uso del modelo praxeológico extendido, permite mostrar la importancia que tienen las tecnologías prácticas en la actividad, estableciendo necesidades y condiciones a cumplir, con una prioridad más alta que las tecnologías matemáticas (eficiencia sobre exactitud). Estas tecnologías prácticas resultan de gran importancia para el diseño de la organización didáctica que se sustentará en estas praxeologías. Para realizar la selección de las praxeologías que pudieran ser llevadas a  $E(M)$  en el diseño, y para sustentar adecuadamente dicha elección se realizó un análisis a  $E(M)$ , que permitió identificar praxeologías matemáticas con las que pudieran vincularse. A su vez, se hizo un primer análisis de otros contextos, sociometría, bibliometría, econometría, deportes y computación, en los

que el modelo analizado se encuentra en funcionamiento y que pudiera resultar útil en el diseño

Los análisis aquí presentados, constituyen las primeras tres fases de la metodología expuesta en el capítulo 3, sentando las bases para desarrollar la cuarta fase, diseño de la organización didáctica, que se presenta en el siguiente capítulo.

...la cuestión es que al investigador le resultó  
complejo el análisis... que puede decir al  
respecto? porque es parte de sus  
CONCLUSIONES DEL CAPÍTULO

## **5 LA ORGANIZACIÓN DIDÁCTICA basada en “PageRank”**

### **5.1 Introducción**

En este capítulo se presenta el diseño de la organización didáctica, que corresponde a la cuarta fase de la metodología presentada en el capítulo 3 de este reporte de investigación. Para este diseño, se consideran las primeras dos aproximaciones a la praxeología de modelación “determinar criterio de importancias”, que son: el modelo de sumatoria de enlaces y el modelo de sumatoria de importancias de los enlaces recibidos. Así como la praxeología “representación matemática de la estructura de la Web”, asociada a la praxeología “determinar criterio de importancias”. Estas praxeologías fueron identificadas y analizadas en la segunda fase de la metodología y seleccionadas en la tercera (ambas fases se presentan en el capítulo cuatro). Resulta importante resaltar que, en el caso de la representación matemática de la estructura de la Web, el enfoque del diseño se dirige al uso de modelos previamente enseñados: Grafo dirigido y Matriz de incidencia, como se muestra en el capítulo 4 (pág. 70-71). En el caso del modelo de importancias, éste va dirigido a la deducción de un modelo “nuevo”, es decir, que debe ser construido por el estudiante. Se pretende llevar estas praxeologías al aula utilizando los tres primeros momentos didácticos, expuestos en el Capítulo 2: *momento del primer encuentro*, *momento exploratorio* y *momento de construcción de un entorno tecnológico-teórico*. Los que permiten estructurar un proceso de estudio, es decir, una Actividad de Estudio e Investigación (AEI), cuyo diseño se presenta a continuación.

### **5.2 El diseño de la organización didáctica, “La problemática de los buscadores Web”**

La organización didáctica se generó como un proceso de estudio, que pretende emular la construcción de las praxeologías seleccionadas del método PageRank: representación matemática de la estructura de la Web y determinación del criterio de importancias. Para esto, se consideraron los tres primeros momentos didácticos: *momento del primer encuentro*, *momento exploratorio* y *momento de construcción de un entorno tecnológico-teórico*. Las tareas propuestas tienen como propósito generar un proceso de estudio que permita a los estudiantes construir las praxeologías seleccionadas. Se espera también que estas tareas motiven la validación de las tecnologías identificadas en el análisis praxeológico, particularmente las seleccionadas y transpuestas a  $E(M)$ .

Reconociendo que las tareas de la AEI motivan el estudio y la investigación, aparecen preguntas abiertas y cerradas. Las primeras tienen el objetivo de que el estudiante genere a partir de su propio proceso de estudio técnicas y tecnologías relacionadas con las de las praxeologías representación matemática de la estructura de la Web y determinación del criterio de importancias. Las segundas tienen como propósito acercarlo a las técnicas y tecnologías de dichas praxeologías.

### 5.3 El entorno tecnológico teórico base del diseño de este proceso de estudio

Para el diseño se consideró un entorno tecnológico teórico, que emerge del análisis praxeológico de la programación Web (presentado en el capítulo 4), conformado por los siguientes elementos:

*Ordenación:* que el estudiante identifique la necesidad de eficiencia y de rapidez de los buscadores Web, traducida como la necesidad de una ordenación efectiva de los resultados de una búsqueda, que puede ser atendida a través de la ordenación del conjunto de sitios.

*Importancia del sitio:* que el estudiante reconozca la "importancia" de un sitio como una medida de su confiabilidad o relevancia, para después utilizar esta noción como criterio de ordenación.

*Representación de la estructura de la Web:* motivar el uso de una representación gráfica, como noción del grafo dirigido, así como el uso de una representación tabular, como noción de la matriz de incidencia. Buscando que el estudiante identifique su utilidad como herramientas de análisis de la información representada y de su estructura.

*Criterio de importancia:* que el estudiante deduzca el criterio para la asignación de importancias del análisis de la información obtenida de las representaciones, considerando para esto las dos primeras aproximaciones identificadas en el análisis praxeológico: el valor de la importancia determinado por la cantidad de enlaces que recibe el sitio como medida de popularidad y el valor de la importancia determinado por la importancia de los sitios que emiten los enlaces, como aproximación a una ordenación más precisa.

Para que la ordenación aparezca como una tarea necesaria en el entorno, se planteó la siguiente situación problemática, “La problemática de los buscadores Web” que conforma la primera parte del proceso de estudio y se presenta a continuación.

La actividad da inicio con el planteamiento de la problemática, que cumplirá con la función de hacer pasar al estudiante por el *momento del primer encuentro*. Se comienza situando al estudiante en un contexto: un grupo de trabajo que diseña y desarrolla (programa) una aplicación a solicitud de un cliente. Ésta forma de trabajo tiene lugar en la práctica profesional del ingeniero en sistemas computacionales, por lo que es emulada constantemente en su formación, cuyas asignaturas de especialidad  $E(DI)$ , desarrollan proyectos en grupos de trabajo como parte de sus actividades. Por lo tanto, se trata de una situación familiar para el estudiante, no sólo por el contexto, sino por la forma de trabajo que se le plantea.

La situación nos permite presentar al estudiante una necesidad a satisfacer en lugar de un problema a resolver: la rapidez y eficiencia en el funcionamiento de un buscador, necesidad que será la razón de ser de la praxeología a construir.

### **La problemática de los buscadores Web**

Eres parte de un grupo de desarrolladores contratados por una empresa para trabajar en el diseño y desarrollo de un buscador Web para uso propio que sea rápido y eficiente. La preocupación principal es la eficiencia, para ello desean reducir al máximo la cantidad de sitios que deben consultar para tener información suficiente, y además, la seguridad de que dicha información sea realmente valiosa y relevante. Durante las primeras etapas de desarrollo, la principal preocupación fue resolver la manera de buscar el término introducido por el usuario, y analizar cómo las características del término en un sitio pueden ayudar a identificar un resultado valioso. Para esto se tomaron en cuenta las siguientes características:

- La cantidad de veces que el término aparece dentro de un sitio.
- Si se encuentra dentro de algún título, o dentro del texto normal.
- El tamaño de fuente, así como su comparativa con respecto al texto que le rodea.
- El tipo de fuente.
- Características del texto: negrita, cursiva, subrayado.

A continuación se le plantean las características de un primer diseño de buscador, como apoyo para el diseño de esta sección se recurrió al documento “”, documento que no fue analizado profundamente, pero que nos permite observar el panorama previo a la creación de PageRank:

En 1994, algunas personas creyeron que un índice completo de búsqueda haría posible encontrar cualquier cosa fácilmente. De acuerdo a “Best of the Web 1994 – Navigators”, “El mejor servicio de navegación debería facilitar encontrar casi cualquier cosa en la Web (una vez que se haya introducido toda la información).” Sin embargo, la Web de 1997 es muy diferente. Cualquiera que haya usado un buscador recientemente, puede testificar que la completitud del índice no es el único factor en la calidad de los resultados de la búsqueda. Los “resultados basura” a menudo ocultan cualquier resultado en el que el usuario estuviera interesado. De hecho, hasta noviembre de 1997, sólo uno de los cuatro buscadores comerciales principales se encuentra a sí mismo (regresa su propia página en los primeros 10 resultados, en respuesta a buscar su nombre). Una de las causas principales de éste problema es que el número de documentos en los índices se ha incrementado por muchos órdenes de magnitud, pero la habilidad del usuario para revisar cada uno de ellos no. Las personas solo están dispuestas a ver los primeros 10 resultados. Debido a esto, al tiempo que el tamaño del conjunto crece, necesitamos herramientas que tengan una precisión muy alta (número de documentos relevantes mostrados, digamos en los primeros 10 resultados). Queremos que nuestra noción de “relevante” solo incluya los mejores documentos dado que podría haber decenas de miles de documentos parcialmente relevantes. Una precisión muy alta es importante aún si tiene que sacrificar la recuperación de resultados (el número total de documentos relevantes que el sistema es capaz de devolver). Hay bastante optimismo recientemente en que el uso de más información hipertextual puede ayudar a mejorar la búsqueda y otras aplicaciones. En particular, la estructura y el texto de links proveen mucha información útil para emitir juicios de relevancia y filtros de calidad. Google hace uso de ambos, estructura de links y el texto enlazado (anchor text). (Brin y Page, 2012)

Asimismo se presentan los resultados de un supuesto periodo de prueba, teniendo como objetivo:

- Emular un buscador previo al desarrollo de Google, mostrando el método empleado por tales buscadores, basado únicamente en la búsqueda del término solicitado así como sus características (consideraciones de búsqueda expuestas en el Capítulo 4).
- Plantear las problemáticas surgidas a raíz de las deficiencias de los métodos de búsqueda, que motivaron el desarrollo de PageRank (Brin y Page, 1998).
- Establecer la información de la que se dispondrá para la construcción del modelo: sitios/enlaces y la estructura conformada por ellos (tecnología de la praxeología representación matemática de la estructura de la Web).

Se analizó el significado que cada una de ellas puede tener:

- Si el término se repite varias veces, es posible que sea parte importante del contenido del sitio.
- Su aparición dentro de un título indica que el contenido del sitio gira alrededor del término.
- El tamaño de fuente, un mayor tamaño “realza” el término, puede indicar que es importante para el sitio.
- El tipo de fuente, puede hacer más llamativo el término, también puede indicar cierta importancia.
- Así como los dos anteriores, la negrita, cursiva y subrayado, realzan y pueden indicar una cierta importancia del término.

Para cada característica se hizo un análisis para darle cierta prioridad a cada una, y así se estableció un “criterio de búsqueda”.

Sin embargo, al momento de realizar pruebas de desempeño para este primer prototipo de buscador se encontraron algunos problemas:

- Al realizar la búsqueda del mismo término más de una vez, las listas de resultados variaban de búsqueda a búsqueda.
- Los resultados encontrados no siempre eran de valor, apareciendo sitios “basura”, de contenido irrelevante.

Dentro de estos sitios basura, se detectó una alta ocurrencia de “spam”. Sitios que aparecían dentro de los resultados no tenían contenido real sobre el término a buscar, sino que contenían grandes cantidades de palabras repetidas múltiples veces, engañando con esto a nuestro buscador. Como parte del diseño, la información de la que disponemos son las direcciones de los sitios, así como los enlaces existentes entre ellos.

1. Con la información disponible, ¿puedes proponer alguna solución?
2. ¿Cómo podrías asegurar que el buscador entregue siempre la misma lista de resultados, sin importar cuantas veces hagas la búsqueda?

Una vez pasado el primer momento se busca que el estudiante pase por el momento *exploratorio*. Para ello, en la pregunta 1 se pide al estudiante una primera propuesta de solución para las problemáticas planteadas, con el objetivo de averiguar a qué herramientas recurre en un principio, en ausencia de alguna orientación. Se busca también saber cuál es la tarea que identifica a raíz de la información que se le proveyó. A continuación, la pregunta 2, busca causar que el estudiante argumente la validez de su propuesta, pasando por el momento de *construcción de un entorno tecnológico-teórico*, identificando la necesidad de ordenar, o al menos clasificar los sitios de acuerdo a un criterio. Es posible que en este momento el estudiante también evalúe la técnica

propuesta, sentando la base para refinarla o reemplazarla, con lo que consideramos posible el emerger de la función tecnológica de evaluación de la técnica, así como de los momentos de *trabajo de la técnica y evaluación de la praxeología construida*.

### **Análisis de la problemática**

3. Si tienes una tabla con la lista de sitios en un orden cualquiera, y siempre inicias la búsqueda desde el primero de la lista, ¿entregarás siempre la misma lista de resultados?
4. ¿Esta búsqueda te parece eficiente? ¿Por qué?
5. ¿Es seguro que obtendrías los resultados más relevantes? ¿Por qué?

Se abordan los aspectos de la problemática planteados en el momento del primer encuentro: la consistencia en los resultados entregados, es decir, que no varíen de búsqueda a búsqueda y la necesidad de asegurar la calidad de los resultados. Se espera que estas preguntas hagan pasar al estudiante por el *momento exploratorio* y el *momento de construcción del entorno tecnológico-teórico*.

Para esto, se pide una primera propuesta de solución, para observar la técnica a la que recurre el estudiante. Se espera observar aquí una propuesta de ordenación, o en su defecto de clasificación, para posteriormente proponer una búsqueda sobre el conjunto ordenado/clasificado. Es posible que haya propuestas de otros tipos, como condiciones de búsqueda o almacenamiento de resultados previos. Para esto se le hace considerar la idea de que el recorrer el conjunto de información siempre en el mismo orden arrojará una lista de resultados consistente. Inmediatamente se cuestiona la eficiencia de esta técnica así como el tipo de resultados que ofrece, buscando que el estudiante plantee algunas de las tecnologías que motivan la necesidad de un criterio de ordenación adecuado a los objetivos.

6. ¿Hay alguna manera de que puedas evitar los sitios que hacen spam?
7. Si muchos de esos sitios no son enlazados por otros, ¿puedes plantear alguna solución?
8. Considerando todo lo anterior, ¿puedes proponer una solución que resuelva ambos problemas?
9. Si al hacer una búsqueda inicias siempre desde el primer sitio de la tabla, ¿habrá una forma de ordenarlos que te permita evitar el spam?

De manera similar se aborda el segundo aspecto: la necesidad de asegurar la calidad del sitio que se muestra en la lista de resultados, relacionando la calidad de los resultados con el evitar los sitios que practican el spam y, por extensión, aquellos que no contengan información relevante. Nuevamente, se busca observar la técnica de solución propuesta preliminarmente por el estudiante para averiguar si es posible que desde un inicio consideren la implementación de un criterio de ordenación como suficiente. Considerando posible que el estudiante plantee una técnica distinta a lo que se espera, se le plantea que los sitios que practican el spam no son enlazados por otros sitios de la

red. Esto permite presentar la noción de que los sitios con contenido de baja calidad tienden a no ser enlazados por otros, estableciendo la idea de que el número de enlaces puede servir para predecir la importancia de un sitio. Importancia que dependerá sólo del comportamiento y organización de la Web, visto como la estructura de los enlaces entre sitios, y no del contenido específico del sitio.

Se pide al estudiante que considere ambas problemáticas y plantee una nueva propuesta de solución. Se espera observar si el análisis de la problemática resulta suficiente para hacer que llegue a la aproximación del primer modelo de importancias, o al menos, observar si permite encaminarlo hacia ella, en caso de que haya planteado una propuesta preliminar muy alejada a lo que buscamos. En caso de que haya hecho alguna propuesta que considerara la importancia, confiabilidad o relevancia del sitio, se espera que ahora sea expresada en términos de un criterio de ordenación.

El grupo llegó a la conclusión que para que el buscador cumpla con su objetivo de rapidez y eficiencia se puede considerar lo siguiente:

Si ordenamos el conjunto de sitios, y los primeros lugares fueran asignados a aquellos con la información más confiable, y así de forma descendiente, podríamos hacer la búsqueda en ése orden. De esa manera la lista de resultados que obtengamos, estaría ordenada de los más “valiosos” a los menos importantes.

10. ¿Puedes proponer alguna manera de ordenar los sitios de manera que se cumpla lo anterior con la información que tienes hasta ahora?

Para finalizar esta sección, se le cuestiona directamente en términos de un criterio de ordenación, para averiguar si tras lo mencionado sobre los sitios que hacen spam, el estudiante puede establecer alguna relación entre la cantidad de enlaces recibidos y la importancia del sitio. Se le plantea que el uso de un criterio de ordenación que genere una lista de sitios importantes puede ser una vía de solución para la problemática. Establecido esto volvemos a pedir al estudiante que proponga un criterio de ordenación ahora en términos de "importancia" de los sitios.

### **Simulación de la situación**

En la simulación de la situación se plantea un ejemplo práctico que permite emular la estructura de la Web a pequeña escala. Con esto se pretende que el estudiante pase nuevamente por un *momento de primer encuentro*. Para ello se emplean algunos sitios con nombres muy conocidos como Wikipedia, MuyInteresante y NationalGeographic

en contraste con otros poco conocidos como Botanicojerez y Selwo. Se pretende que esto sirva para que el estudiante pueda identificar con mayor facilidad la noción de "popularidad" de un sitio, para derivarla después en "importancia".

Como parte de las pruebas de desempeño para el buscador que estás desarrollando, se seleccionó un conjunto de 11 sitios en cuyo contenido aparece el término: 'panda rojo'.

[nationalgeographic.es](http://nationalgeographic.es)

[bioenciclopedia.com](http://bioenciclopedia.com)

[herramientas.educa.madrid.org](http://herramientas.educa.madrid.org)

[botanical-online.com](http://botanical-online.com)

[wikipedia.com](http://wikipedia.com)

[botanicojerez.com](http://botanicojerez.com)

[selwo.es](http://selwo.es)

[denverzoo.org](http://denverzoo.org)

waza.org

lareserva.com

muyinteresante.es

Si un usuario buscara información sobre el panda rojo para elaborar algún trabajo y tuviera tiempo limitado no resultaría eficiente que se detuviera a consultar el contenido de los 11 sitios. Se decidió entonces que consultando 5 de los 11 sitios el usuario dispondría ya de suficiente información para sus fines.

11. ¿Cómo elegirías esos 5 sitios de manera que, sin consultarlos previamente, puedas confiar en que la información que contienen es relevante y confiable?

Buscando una forma de determinar los sitios a mostrar, se encontró que los 11 sitios están conectados entre sí de la siguiente manera:

nationalgeographic.es, enlaza a:

Bioenciclopedia

Denverzoo

bioenciclopedia.com enlaza a:

Wikipedia

botanical-online

Selwo

herramientas.educa.madrid.org enlaza a:

Wikipedia

Nationalgeographic

Waza

Muyinteresante

botanical-online.com enlaza a:

Wikipedia

Lareserva

Selwo

wikipedia.com enlaza a:

Nationalgeographic

Muyinteresante

botanicojerez.com enlaza a:

Botanical-online

Wikipedia

Muyinteresante

Selwo

selwo.es enlaza a:

Waza

Denverzoo

Lareserva

denverzoo.org enlaza a:

Nationalgeographic

Waza

waza.org enlaza a:

Nationalgeographic

Wikipedia

lareserva.com enlaza a:

12. ¿Esta información (sitios/enlaces) te dice algo? ¿Por qué?

13. ¿Te es posible elegir los 5 sitios en este momento? ¿Por qué?

Nuevamente se busca que el estudiante pase por el momento *exploratorio* y de *construcción del entorno tecnológico-teórico*. Se cuestiona al estudiante, estableciendo la condición de seleccionar sólo un cierto número de sitios que contengan la "mejor" información, esto sin acceder previamente los sitios, por lo tanto, sin conocer su contenido. Se espera que la experiencia previa del estudiante le haga plantear que es factible buscar en primer lugar en los sitios más conocidos, y por ende, los más populares. La lista de sitios y los enlaces entre ellos, tiene por objetivo preparar al estudiante para la siguiente sección, donde trabajará las representaciones de esta estructura. Para esto las preguntas 12 y 13 buscan provocar que el estudiante identifique la dificultad de extraer información de esta representación, y por lo tanto, establecer la dificultad de utilizar esta representación para elegir los sitios.

### **Representaciones de la estructura de la Web**

El grupo llegó a la conclusión de que resulta complicado hacer una interpretación de esta información, así que se propuso buscar otras formas de plasmarla.

14. ¿Cómo representarías de manera gráfica los sitios y sus enlaces?
15. ¿Cómo representarías los sitios y enlaces en una tabla?
16. ¿Qué representación te parece más útil? ¿Por qué?

Partiendo del ejemplo, se busca abordar el uso del grafo dirigido y su matriz de incidencia como representaciones de la estructura de la Web. Se considera la lista de los enlaces entre cada sitio, mostrada en el ejemplo. Se le cuestiona entonces sobre esta información, con el objetivo de plantear la necesidad de utilizar una representación más comprensible, o útil, para el análisis de esta información, constituyendo esto otro *momento de primer encuentro*, al que seguirán nuevamente los momentos *exploratorio* y de *construcción del entorno tecnológico-teórico*.

Se presenta la tarea de representar la estructura de sitios y enlaces, gráfica y tabularmente, esperando observar nociones del grafo dirigido y de la matriz de incidencia. No se dan especificaciones sobre la representación a utilizar, para observar a que técnica recurre el estudiante en primer lugar. Se le cuestiona entonces sobre la

utilidad y comprensibilidad de las dos representaciones, como herramienta para analizar el conjunto de información. Para esto se espera identificar si el estudiante responde con tecnologías provenientes de la disciplina o en términos de preferencia personal.

### **Análisis de la estructura de la Web**

17. ¿Observas algún patrón o comportamiento en la gráfica o en la tabla?
18. ¿Te sirve esto para elegir los sitios a mostrar? ¿Por qué?
19. Tras observar ambas representaciones, ¿qué dirías que significa la cantidad de enlaces que recibe cada sitio?
20. Regresando al problema original, ¿podrías utilizar esta información para ordenar los

Posteriormente, y basado en las representaciones previas, en esta sección se pide al estudiante identificar algún comportamiento o patrón en la representación que le haya resultado más útil (preguntas 17-19). Se pretende entonces que pase por el *momento de construcción del entorno tecnológico-teórico*. Se retoma la noción de que la organización interna de la Web puede servir para establecer un criterio de importancia para cada sitio. En caso de que el estudiante observara o planteara algún patrón distinto al buscado, se le cuestiona directamente sobre el posible significado de la abundancia o escasez de enlaces recibidos por un sitio. Después, en lo que se espera sea un *momento exploratorio*, se le pide que relacione esta información con la situación original, esperando que pueda surgir una aproximación al primer modelo de importancias (pregunta 20). Con esto finaliza la primera actividad de la organización didáctica.

### **Actividad 2**

Como parte de la segunda actividad, se pretendía guiar al estudiante a la matematización del modelo, y a la evaluación de sus resultados como herramienta para deducir el segundo modelo y la matematización del mismo para nuevamente evaluar resultados y usarlos como validación del modelo. En este proceso, resultaría posible observar los momentos de *trabajo de la técnica y evaluación de la praxeología construida*, dentro de la *construcción del entorno tecnológico-teórico*, a la vez que la deducción del segundo modelo corresponde a un *momento exploratorio*. Sin embargo, tras evaluar la implementación de la primera actividad, se consideró que la naturaleza abierta de las preguntas y la posibilidad de que emplearan representaciones distintas al grafo y la matriz causaron que el estudiante no llegara a deducir el primer modelo, y que no observara el papel del grafo y su matriz como herramientas de análisis. Esto se debe a que, como se identificó en el análisis praxeológico, la representación de la estructura de la Web es necesaria para la deducción de las aproximaciones al criterio de importancias.

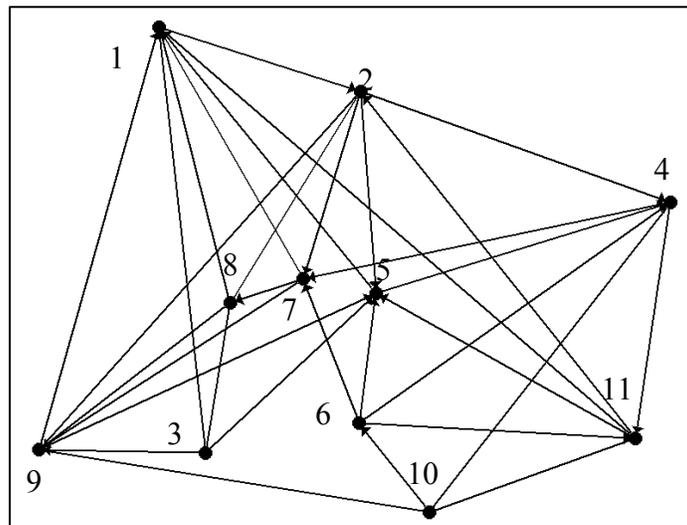
Debido a esto la actividad tuvo que ser modificada, retomando en la primera sección la lista de sitios del ejemplo, y presentando a continuación su grafo y matriz de incidencia correspondientes. Se pretendió asegurar con esto que el estudiante disponga de las representaciones adecuadas para que pueda realizar el análisis de la estructura de la Web, esto constituye un *momento de primer encuentro*, con el que se iniciará la actividad 2.

### **Análisis de la estructura de la Web**

Retomando la lista de sitios de la actividad anterior:

nationalgeographic.es  
bioenciclopedia.com  
herramientas.educa.madrid.org  
botanical-online.com  
wikipedia.com  
botanicojerez.com  
selwo.es

1. Uno de los miembros del equipo presentó la siguiente representación de los sitios y sus enlaces:



Para tratar de analizar mejor esta estructura, la representó a través de una matriz:

	A	B	C	D	E	F	G	H	I	J	K
A	0	1	0	1	0	1	1	1	0	1	
B	1	0	0	0	0	0	0	1	1	0	1
C	0	0	0	0	0	0	0	0	0	0	0
D	0	1	0	0	1	1	0	0	0	1	0
E	0	1	1	0	0	1	0	0	1	0	1
F	0	0	0	0	0	0	0	0	0	1	0
G	1	1	0	1	0	1	0	0	0	0	0
H	1	1	1	0	0	0	1	0	0	0	0
I	0	1	1	0	0	0	1	1	0	1	0
J	0	0	0	1	0	0	0	0	1	0	1
K	1	0	0	1	0	1	0	0	0	1	0

2. ¿Observas algún patrón o comportamiento en la estructura de los sitios y sus enlaces?
3. ¿Cuál te parece más fácil de analizar? ¿Por qué?
4. ¿Qué significa que un sitio reciba una gran cantidad de enlaces de otros sitios?
5. ¿Qué significa que no reciba ninguno como en el caso del sitio 3?
6. Si trataras de ordenar los sitios como se propuso en la actividad anterior, sin consultar cada uno para verificar su contenido, ¿crees que la cantidad de enlaces recibidos pueda servirte como criterio de ordenación? ¿Por qué?

Las preguntas planteadas tienen como objetivo que el estudiante identifique la existencia de un patrón en la estructura sitios/enlaces que permita establecer un criterio independiente del análisis del contenido de los sitios, tratándose de un *momento exploratorio*. Nuevamente se espera que el estudiante argumente la utilidad de cualquiera de las dos representaciones, ya sea desde el punto de vista de la práctica, como desde la preferencia personal. Para finalmente cuestionar directamente al estudiante sobre la posible utilidad del número de enlaces como criterio de ordenación, buscando observar que tecnologías utiliza para dar respuesta. Nuevamente se pretende que con estos cuestionamientos el estudiante pase por el momento de *construcción del entorno tecnológico teórico*.

### **Primer criterio de importancias**

Tras analizar la estructura anterior, otro miembro del equipo propuso lo siguiente:

Si los sitios fueran personas, podríamos ver los enlaces como ‘recomendaciones’, cuando un sitio enlaza a otro nos está invitando a consultarlo. Si un sitio es ‘recomendado’ muchas veces nos indica cierto grado de confiabilidad o importancia de su contenido, sin tener que consultarlo previamente.

7. De acuerdo a esta propuesta, escribe una expresión que te permita determinar la ‘importancia’ de un sitio.
8. Aplicando el criterio, ¿cuál es el orden resultante para los 11 sitios?

En esta sección se utiliza el análisis previo del grafo y la matriz, así como una pequeña analogía con un contexto tomado de la sociometría que le permite al estudiante observar la problemática desde un punto de vista distinto, analogía que podría ser vista como otro *momento de primer encuentro*. Se pide que escriba una expresión que le permita determinar la importancia, esperando que utilice alguna forma de sumatoria para plantear el criterio, o que al menos resuma verbalmente un criterio, siendo esto un *momento exploratorio*. Después se le solicita escribir el orden resultante al usar el criterio, para cuestionar si es que resuelve la problemática, así como sus justificaciones. Se espera observar si es que el estudiante considera que más de un sitio puede tener el mismo número de enlaces, y por lo tanto, que plantee que el criterio no es suficiente para resolver la problemática. En caso de que identifiquen tal situación, se espera que

planteen la necesidad de modificar el criterio para considerar esta situación, es decir, que el modelo esta "incompleto". Todo esto conforma un *momento de construcción del entorno tecnológico-teórico*, en el que puede observarse también el *momento de evaluación de la praxeología construida*.

### Segundo criterio de importancias

10. Si dos o más sitios reciben el mismo número de enlaces, ¿quién se colocaría en primer lugar en la lista sin tener que consultarlos?

11. ¿Tienen el mismo valor todos los enlaces?

Un miembro del equipo menciona lo siguiente:

No es lo mismo recibir un enlace desde un sitio grande como NationalGeographic, que desde un sitio casi desconocido como botanical-online o Laereserva, un enlace desde un sitio grande debería ser mucho más valioso.

12. Integra esta consideración al criterio anterior. Ahora, ¿cómo calcularías la importancia?

13. Escribe la expresión resultante.

En esta última sección buscamos que el estudiante deduzca el segundo criterio de importancias, como una adaptación del primero. Para esto, se expone directamente la situación en que más de un sitio reciba el mismo número de enlaces, para cuestionarle sobre el "valor" de los mismos. Se espera que la experiencia propia, el conocimiento sobre los sitios de la lista y la analogía de recomendación lo haga considerar que el valor de un enlace depende de quien lo emita. Se le pide entonces que integre esta consideración a su criterio anterior y lo reescriba, esperando encontrar aquí alguna aproximación, por lo menos verbal, al segundo modelo. En estas preguntas, se espera observar el *momento de construcción del entorno tecnológico-teórico*, en el que, debido a que se espera una adaptación del primer criterio, también se espera que aparezcan los momentos de *trabajo de la técnica y evaluación de la praxeología construida*.

### 5.4 Conclusión

Como puede observarse en las actividades, el enfoque del diseño recae en la construcción del entorno tecnológico-teórico, identificado en el análisis praxeológico. A su vez, cabe resaltar que no por tratarse de una AEI, el proceso de estudio se encuentra limitado a los primeros tres momentos didácticos. Consideramos que es posible observar los momentos de *trabajo de la técnica y evaluación de la praxeología construida*, al causar que el estudiante cuestione la efectividad de las técnicas y proponga o plantee modificaciones a las mismas. Pudiendo observar en estos procesos, las funciones tecnológicas de evaluación y facilitación de la técnica. Podemos observar que cuando aparecen estos dos momentos, particularmente la *evaluación de la praxeología construida*, se encuentran conectados con el *momento de construcción del entorno tecnológico-teórico*. Esto parece verse influenciado por el diseño de la organización didáctica, el cual busca que el estudiante proponga técnicas propias, para tomarlas como punto de partida y guiarlo a través de su adaptación o sustitución. Por lo que los dos últimos momentos, a pesar de no ser considerados originalmente como parte del proceso de estudio, son favorecidos.

Consideramos que la implementación es de tipo exploratorio, buscando que los resultados obtenidos al ponerlas a prueba en un ambiente escolar permitan plantear un diseño más robusto, que pueda ser implementado posteriormente. Esto a través de la identificación de problemáticas durante la aplicación, pero de manera más importante aún, con la identificación de tecnologías expuestas por los estudiantes, que pudieran ayudar a refinar el diseño.

Con estas consideraciones en mente, se llevó a cabo la implementación de la organización didáctica en dos sesiones, aplicadas a un grupo de tercer semestre, de la carrera de ingeniería en sistemas computacionales del Instituto Tecnológico de Chilpancingo. Los resultados obtenidos en la implementación de la actividad serán expuestos en el capítulo siguiente.

## 6 RESULTADOS

### 6.1 Introducción

En este capítulo se presenta el análisis de la implementación de la organización didáctica, “La problemática de los buscadores Web”, en un grupo de 20 estudiantes de la carrera de ingeniería en sistemas del ITCH. El grupo seleccionado es de tercer semestre de la asignatura de Investigación de operaciones, que ya ha cursado Matemáticas discretas y Álgebra Lineal. Dicha implementación tuvo como objetivo principal, que el estudiante viviera un proceso de estudio conformado por las praxeologías representación matemática de la estructura de la Web y determinación del criterio de ordenación.

Se muestran en primer lugar, las condiciones en las que la implementación fue realizada, condiciones de partida, grado y asignatura que cursan, organización del grupo durante la actividad, así como condiciones institucionales que influenciaron la implementación. Se describen también las partes que componen las actividades, así como los objetivos de cada una y se aborda la reorganización que la organización didáctica experimentó a raíz de la implementación. En segundo lugar se aborda el análisis de la primera actividad, dividiéndola en las partes mostradas, analizando la actividad de los estudiantes ilustrándola con las técnicas y tecnologías expresadas en sus respuestas. Finalmente, se aborda la implementación de la segunda actividad y el respectivo análisis de las partes que la conforman, así como los resultados del análisis de la implementación de la primera actividad que motivaron su reorganización.

### 6.2 Condiciones de la implementación y reorganización de la organización didáctica

“La problemática de los buscadores Web” se conformó por dos actividades implementadas en dos sesiones de 50 minutos, con un día de diferencia entre ellas, lo que permitió analizar rápidamente la realización de la actividad 1 y modificar en base a ello la actividad 2, como se precisa más adelante. Recordemos que la primera actividad tiene cinco partes, con objetivos asociados al entorno tecnológico-teórico que se busca construir (Capítulo 5):

- 1) *Planteamiento de la problemática*, busca hacer surgir la noción de importancia así como la necesidad de una técnica de ordenación.
- 2) *Análisis de la problemática*, reafirmar la noción de importancia como medida de confiabilidad y como posible criterio de ordenación.
- 3) *Simulación de la situación*, busca provocar que los estudiantes generen un primer criterio y que comiencen a identificar la utilidad del grafo dirigido y la matriz de incidencia como herramientas de análisis.
- 4) *Representación de la estructura de la Web*, busca que los estudiantes representen la lista de sitios y enlaces de manera gráfica y tabular, con el objetivo de que aparezcan nociones del grafo y la matriz de incidencia, así como que identifiquen su utilidad como herramienta de análisis.

- 5) *Análisis de la estructura de la Web*, busca que los estudiantes reconozcan la concentración de enlaces como patrón que puede derivarse en el criterio de importancias.

Un primer análisis del trabajo de los estudiantes, permitió identificar dificultades con las representaciones de la estructura de la Web, ya que algunos equipos utilizaron tablas de datos, árboles, histogramas, mapas conceptuales en lugar de grafos o matrices. Esto motivó, como señalado anteriormente, una reorganización de la segunda actividad, que se dividió en dos partes con objetivos asociados, que se describen a continuación:

- 1) *Análisis de la estructura de la Web*, se retoma de la primera actividad, pero dando al estudiante las representaciones que se esperaba produjeran en la parte 4, buscando nuevamente que el estudiante reconozca la concentración de enlaces como patrón que puede derivarse en el criterio de importancias.
- 2) *Aproximación a los criterios de importancias*, donde se espera que los estudiantes exprese el primer criterio, y deduzcan el segundo.

Esta nueva organización, particularmente en la segunda actividad, pretende que los estudiantes puedan superar las dificultades asociadas a una representación adecuada de la Web. Es importante señalar que se disponía de poco tiempo para realizar las actividades, lo que obligaba a “apresurar” este proceso. Es posible, que en otras condiciones esta reorganización no se requiera, gestionando de manera distinta el proceso de estudio, por ejemplo a través de nuevas cuestiones que permitan confrontar a los estudiantes con las limitantes de las representaciones utilizadas (árboles, histogramas, etc.). Además del tiempo, otras condiciones institucionales influyeron fuertemente en la implementación de la organización didáctica en cuanto al trabajo en equipo. A pesar de que esta forma de trabajo es característica de la práctica profesional de ingeniero en sistemas computacionales (como se expuso en el diseño del planteamiento de la problemática, en el Capítulo 5), no es la forma de trabajo usual en las asignaturas de  $E(M)$ , siendo que se introduce y es utilizada como forma de trabajo estándar en las asignaturas de  $E(DI)$  en semestres posteriores. Debido a esto, a pesar de que se hizo que los estudiantes trabajaran en equipo, el profesor solicitó que todos los estudiantes respondieran sus actividades de manera individual, como evidencia de trabajo. Considerando esto, se recurrió a identificar al estudiante que respondiera de manera más clara dentro de su equipo y que expresara las respuestas consensadas en el mismo, como medio para ilustrar el análisis.

En esta implementación se hicieron cuatro equipos de cinco estudiantes, aunque es importante mencionar que no todos los participantes realizaron las dos sesiones, sí se mantuvo un trabajo en equipo. El análisis que aquí se presenta se hace sobre la actividad entregada por los cuatro equipos y se busca mostrar de manera general la forma en que enfrentaron la problemática de los buscadores Web.

### **6.3 Análisis de la actividad 1**

En esta sección se analizan las cinco partes que conforman la actividad 1: planteamiento de la problemática, análisis de la problemática, simulación de la situación, representación de la estructura de la Web y análisis de la estructura de la Web. Para

ilustrar el análisis se consideran las respuestas dadas por algunos estudiantes de los diferentes equipos, que representan los consensos alcanzados en ellos, así como las técnicas y tecnologías expuestas con mayor frecuencia o aquellas que consideramos importante evidenciar.

### 6.3.1 Planteamiento de la problemática

La problemática es planteada a los estudiantes con el objetivo de que identifiquen la necesidad de un método de ordenación como medio para satisfacer la necesidad del buscador: eficiencia y rapidez, así como el establecimiento de la importancia de un sitio como medida de confiabilidad, para posteriormente utilizarla como criterio de ordenación.

#### **La problemática de los buscadores Web**

Eres parte de un grupo de desarrolladores contratados por una empresa para trabajar en el diseño y desarrollo de un buscador Web para uso propio que sea rápido y eficiente. La preocupación principal es la eficiencia, para ello desean reducir al máximo la cantidad de sitios que deben consultar para tener información suficiente, y además, la seguridad de que dicha información sea realmente valiosa y relevante. Durante las primeras etapas de desarrollo, la principal preocupación fue resolver la manera de buscar el término introducido por el usuario, y analizar cómo las características del término en un sitio pueden ayudar a identificar un resultado valioso. Para esto se tomaron en cuenta las siguientes características:

- La cantidad de veces que el término aparece dentro de un sitio.
- Si se encuentra dentro de algún título, o dentro del texto normal.
- El tamaño de fuente, así como su comparativa con respecto al texto que le rodea.
- El tipo de fuente.
- Características del texto: negrita, cursiva, subrayado.

Se analizó el significado que cada una de ellas puede tener:

- Si el término se repite varias veces, es posible que sea parte importante del contenido del sitio.
- Su aparición dentro de un título indica que el contenido del sitio gira alrededor del término.
- El tamaño de fuente, un mayor tamaño “realza” el término, puede indicar que es importante para el sitio.
- El tipo de fuente, puede hacer más llamativo el término, también puede indicar cierta importancia.
- Así como los dos anteriores, la negrita, cursiva y subrayado, realzan y pueden indicar una cierta importancia del término.

Para cada característica se hizo un análisis para darle cierta prioridad a cada una, y así se estableció un “criterio de búsqueda”.

- Sin embargo, al momento de realizar pruebas de desempeño para este primer prototipo de buscador se encontraron algunos problemas: Al realizar la búsqueda del mismo término más de una vez, las listas de resultados variaban de búsqueda a búsqueda.
- Los resultados encontrados no siempre eran de valor, apareciendo sitios “basura”, de contenido irrelevante.
- Dentro de estos sitios basura, se detectó una alta ocurrencia de “spam”. Sitios que aparecían dentro de los resultados no tenían contenido real sobre el término a buscar, sino que contenían grandes cantidades de palabras repetidas múltiples veces, engañando con esto a nuestro buscador. Como parte del diseño, la información de la que disponemos son las direcciones de los sitios, así como los enlaces existentes entre ellos.

3. Con la información disponible, ¿puedes proponer alguna solución?
4. ¿Cómo podrías asegurar que el buscador entregue siempre la misma lista de resultados, sin importar cuantas veces hagas la búsqueda?

El análisis a las actividades de los cuatro equipos permitió constatar que los estudiantes identifican la noción de “importancia de un sitio” como criterio para seleccionar los sitios a entregar al usuario, para ello utilizan términos como, relevancia, popularidad, información útil, sitios confiables. Sin embargo, la noción que evidencian de la importancia de un sitio, aún no está totalmente relacionada con la calidad sino que va dirigida a la idea de ser popular o famoso. Podemos notar que el conocimiento sobre el contexto les permite adentrarse rápidamente en el análisis de la problemática, aunque la mayor parte de las técnicas que proponen son procesos “manuales” de ordenación: establecer una lista indexada, hacer revisiones del contenido de los sitios, utilizar el número de fuentes citadas en su contenido, hacer encuestas de satisfacción a los usuarios, considerar el número de visitas. Estas técnicas están sustentadas por tecnologías tanto de la disciplina (indización de listas), como de la experiencia personal de los estudiantes con la Web (revisiones de contenido, encuestas de satisfacción, fuentes citadas, conteo de visitas, almacenamiento de resultados previos). Los estudiantes se encontraban cursando la asignatura de Estructura de datos, donde se trabajan métodos de ordenación y búsqueda, por lo que para identificar criterios recurrió a tecnologías estudiadas en  $E(DI)$  relacionadas con conteo: cantidad de coincidencias del término buscado, contadores de visitas/fuentes/actualizaciones. Para ilustrar estas técnicas, consideramos las respuestas de la estudiante Zulema (E1), perteneciente al equipo 1, quien señala en primer lugar, al responder la pregunta 1, que un criterio se vuelve necesario para eliminar los sitios basura: “*Sí, de manera personal propongo que se establezca una condición para el máximo de palabras que puedan coincidir con el término que se está buscando, esto para evitar que los sitios “basura”, engañen al buscador*”. Se espera que este reconocimiento la lleve a determinar el criterio de importancia en la parte 5 de la actividad. Mientras que su respuesta a la pregunta 2 es:

Se podría establecer que para cuando una palabra sea buscada de 1 a 5 veces (continuas), entonces automáticamente se mostrará los primeros resultados que se obtuvieron. Para esto, se tiene que asegurar que los primeros resultados cumplan con los requerimientos. (Respuesta estudiante E1)

Esta respuesta parece basarse en tecnologías provenientes de  $E(DI)$  e  $Ip$  como son, automatización de la solución, ahorro de tiempo y recursos, condiciones de búsqueda, asegurando que éstas garantizarán los resultados que buscamos. Asimismo, cuando se expresa que los primeros resultados cumplan con los requerimientos, se hace referencia a la revisión del contenido de los sitios que “verificará” que se cumplan los requerimientos de la búsqueda.

### **6.3.2 Análisis de la problemática**

Recordemos que en esta segunda parte de la actividad se plantean preguntas con el objetivo de que los estudiantes analicen las condiciones que aseguren eficiencia y rapidez del buscador:

a) Consistencia de resultados

6. Si tienes una tabla con la lista de sitios en un orden cualquiera, y siempre inicias la búsqueda desde el primero de la lista, ¿entregarás siempre la misma lista de resultados?
7. ¿Esta búsqueda te parece eficiente? ¿Por qué?
8. ¿Es seguro que obtendrías los resultados más relevantes? ¿Por qué?

b) Calidad de la información (relacionada con el evitar sitios spam o “basura”)

10. ¿Hay alguna manera de que puedas evitar los sitios que hacen spam?
11. Si muchos de esos sitios no son enlazados por otros, ¿puedes plantear alguna solución?
12. Considerando todo lo anterior, ¿puedes proponer una solución que resuelva ambos problemas?
13. Si al hacer una búsqueda inicias siempre desde el primer sitio de la tabla, ¿habrá una forma de ordenarlos que te permita evitar el spam?

Al responder estas preguntas, los estudiantes reforzaron la noción de “popularidad” de un sitio como posible criterio para su selección (equivalente a la importancia de un sitio). Aunque esta tecnología se va refinando, relacionándola ahora con la confiabilidad, no se llega a establecer una técnica para determinarla. Sin embargo, se identificó que las soluciones planteadas por los estudiantes hasta este punto se mantuvieron enfocadas a las técnicas mencionadas anteriormente: conteo de visitas, encuestas de satisfacción, análisis del contenido del sitio, etc., llegando a mezclarlas con el uso de la popularidad como medida. Y en el caso de los estudiantes que solo consideran la popularidad, no establecen un criterio para determinarla, para ilustrarlo mostramos las respuestas del estudiante E6, perteneciente al equipo 2, a las preguntas 9 y 10:

- 9) Si al hacer una búsqueda inicias siempre desde el primer sitio de la tabla, ¿habrá una forma de ordenarlos que te permita evitar el spam?

E6: Con el de mayor popularidad.

- 10) ¿Puedes proponer alguna manera de ordenar los sitios de manera que se cumpla lo anterior con la información que tienes hasta ahora?

E6: Sí, basándote en la popularidad de los sitios y a partir de ahí poner los más populares al principio y de ahí los secundarios.

En el caso de aquellos estudiantes que ya sugerían una técnica de ordenación (caso de estudiante E1), aseguraban que un criterio adecuado permitiría dar respuesta a ambas condiciones, consistencia y calidad de la información, aunque sin expresar el criterio que pudiera lograrlo más allá de algunas condiciones. Hasta este momento los estudiantes reconocieron el valor de la importancia o popularidad de un sitio, sin llegar a considerarlo como posible solución, recurriendo a plantear las técnicas mencionadas

anteriormente: conteo de visitas, encuestas de calidad, análisis de contenido y conteo de fuentes.

Los estudiantes que proponen tales técnicas, que consideramos de tipo “manual”, comenzaron a identificar debilidades en sus propuestas, particularmente en el aspecto de la calidad de la información contenida en los sitios, dándose cuenta que sus técnicas no pueden garantizarla. Esto los llevó a plantear técnicas más robustas, la mayor parte de las propuestas se enfocaron en una técnica que ordenara los sitios de manera que pudieran identificarse los más confiables. Sin embargo, encontramos estudiantes que al tiempo que reconocen la necesidad de un método de ordenación, no profundizan en el funcionamiento y en los criterios asociados a los mismos métodos. Como se ilustra en las respuestas de la estudiante E2 a las preguntas 3, 4 y 5.

- 3) Si tienes una tabla con la lista de sitios en un orden cualquiera y siempre inicias la búsqueda desde el primero de la lista, ¿entregará siempre la misma lista de resultados?

E2: Lo hará, mientras no haya una condición o algún método de ordenamiento, con el cual la información pueda ser manipulada y procesada para cada vez que se busque.

- 4) ¿Esta búsqueda te parece eficiente? ¿Por qué?

E2: No, porque no existe ningún método o condición que repercuta en cuanto se realice una búsqueda.

- 5) ¿Es seguro que obtendrás los resultados más relevantes? ¿Por qué?

E2: No, porque siempre empezaría por el primero, pero puede que el primer sitio no contenga la información que estamos buscando

Los estudiantes muestran tecnologías correspondientes a los sitios que practican spam para tratar de evitarlos: grandes cantidades de enlaces salientes, que en la práctica se considerarían fuentes citadas, además de que se caracterizan por utilizar medios deshonestos de promoción. Estas tecnologías, particularmente la primera, no fueron proporcionadas por la organización didáctica sino que provienen de la experiencia y conocimiento propio del estudiante. Sin embargo, a pesar de que existen menciones al uso de un posible algoritmo de ordenación, en esta parte no se propusieron “verdaderas” técnicas de ordenación, sino técnicas basadas en la selección y aprobación del usuario o análisis del contenido del sitio, que sirviera para indicar la calidad del mismo y generar listas (con el objetivo de evitar los sitios “basura”). Estas técnicas, como se mencionó previamente provienen de los conocimientos de la asignatura de Estructura de datos *E(DI)* y de su experiencia personal con la Web. Cabe recordar que estas técnicas son motivadas por su facilidad de implementación, sin requerir el diseño de un algoritmo, pero que solo funcionarían para cantidades relativamente pequeñas de sitios y no para el volumen real, que contempla billones.

- 6) ¿Hay alguna manera de que puedas evitar los sitios que hacen spam?

E7: Marcar cada sitio que contenga spam como dañino, identificarlos y evitar por completo estos sitios.

E1: Con un algoritmo, filtros, bloqueando las páginas maliciosas.

A pesar de que dentro de las preguntas se planteó directamente a los estudiantes una relación entre la mala calidad de un sitio que practica spam con una escasez de enlaces entrantes, las tecnologías expuestas en sus respuestas continuaron girando fuertemente alrededor del análisis del contenido del sitio como forma de identificar y “marcar” los sitios de baja calidad. Aunque comienzan a hacerse comunes las propuestas de “enlistar los sitios”, los estudiantes siguen sin proponer técnicas definidas para hacerlo, continúan nombrando algunos criterios que a su consideración resultan útiles (visitas, fuentes, recomendación del usuario, análisis de contenido del sitio) como se ilustra a continuación.

6) Si muchos de esos sitios no son enlazados por otros, ¿puedes plantear alguna solución?

E7: En el protocolo de búsqueda, implementar la solución anterior (haciendo referencia a la identificación de los sitios que hacen spam para luego evitarlos por completo)

7) Considerando todo lo anterior, ¿puedes proponer una solución que resuelva ambos problemas?

E7: Mover estos sitios fuera del alcance de nuestro motor de búsqueda para evitar datos irrelevantes.

8) Si al hacer una búsqueda inicias siempre, desde el primer sitio de la tabla, ¿habrá una forma de ordenarlos que te permita evitar el spam?

E7: Por nivel de desempeño.

Por nivel de seguridad.

Por nivel de confiabilidad.

Finalmente, una vez que el estudiante respondió esta parte de la actividad, se le planteó directamente la noción de sitios ordenados por importancia como solución a la problemática.

El grupo llegó a la conclusión que para que el buscador cumpla con su objetivo de rapidez y eficiencia se puede considerar lo siguiente:

Si ordenamos el conjunto de sitios, y los primeros lugares fueran asignados a aquellos con la información más confiable, y así de forma descendiente, podríamos hacer la búsqueda en ése orden. De esa manera la lista de resultados que obtengamos, estaría ordenada de los más “valiosos” a los menos importantes.

11. ¿Puedes proponer alguna manera de ordenar los sitios de manera que se cumpla lo anterior con la información que tienes hasta ahora?

Sin embargo, aún con esta información, la mayoría de los estudiantes continuó proponiendo técnicas de ordenación en función del contenido y volviendo a considerar la frecuencia de uso (número de visitas). De acuerdo a su experiencia las visitas son muy importantes para determinar la calidad de un sitio: “mientras más visitado sea un sitio, más confiable es” toma el papel de criterio de ordenación. Esto causa que aun cuando consideren la noción de importancia, ésta viene a colocarse como criterio secundario. Esto se ilustra con lo expresado por el estudiante E11 al responder la pregunta 10: “*Indizar y que mi buscador tenga la opción de mostrar un top 10 de lo más relevante a lo que estoy buscando o páginas más visitadas y confiables*”. Observamos que propone el uso de una lista indexada, tecnología de  $E(DI)$ , y de manera secundaria una ordenación que considere tanto la confiabilidad como el número de visitas. Esta situación nos permitió dar cuenta de que sin el grafo dirigido como herramienta de análisis los estudiantes tienen dificultades para reconocer la cantidad de enlaces como posible criterio, a pesar de que la noción se les plantee directamente, sujetándose fuertemente a técnicas propias.

### 6.3.3 Simulación de la situación: ordenación empírica vs criterio de ordenación

Como parte de las pruebas de desempeño para el buscador que estás desarrollando, se seleccionó un conjunto de 11 sitios en cuyo contenido aparece el término: 'panda rojo'.

[nationalgeographic.es](http://nationalgeographic.es)

[bioenciclopedia.com](http://bioenciclopedia.com)

[herramientas.educa.madrid.org](http://herramientas.educa.madrid.org)

[botanical-online.com](http://botanical-online.com)

[wikipedia.com](http://wikipedia.com)

[botanicojerez.com](http://botanicojerez.com)

[selwo.es](http://selwo.es)

[denverzoo.org](http://denverzoo.org)

[waza.org](http://waza.org)

[lareserva.com](http://lareserva.com)

wikipedia.com enlaza a:  
Nationalgeographic  
Muyinteresante  
botanicojerez.com enlaza a:  
Botanical-online  
Wikipedia  
Muyinteresante  
Selwo  
selwo.es enlaza a:  
Waza  
Denverzoo  
Lareserva  
denverzoo.org enlaza a:  
Nationalgeographic  
Waza  
waza.org enlaza a:  
Nationalgeographic  
Wikipedia  
lareserva.com enlaza a:  
Botanical-online  
Botanicajerez  
muyinteresante.es enlaza a:  
Wikipedia  
Nationalgeographic  
Bioenciclopedia

21. ¿Esta información (sitios/enlaces) te dice algo? ¿Por qué?
22. ¿Te es posible elegir los 5 sitios en este momento? ¿Por qué?

En la simulación de la situación, el ejemplo práctico que emula la situación real pretende que el estudiante llegue a plantear una aproximación al criterio de ordenación. Para esto se le presenta una lista de sitios (y sus enlaces), cuyos nombres, conocidos por el estudiante, le permitirían asociar la noción de importancia a la cantidad de enlaces recibidos. Sin embargo, encontramos que la pregunta, “¿cómo elegirías esos 5 sitios de manera que, sin consultarlos previamente, puedas confiar en que la información que

contienen es relevante y confiable?” les dio la libertad de presentar una lista de 5 sitios como respuesta empírica, sin que fuera necesario para ellos proponer una técnica de solución, es decir, una aproximación al criterio de ordenación.

Sin embargo, encontramos respuestas como la siguiente: “1. *Por el nombre: si es famoso o reconocido (NatGeo, Muy Interesante)*. 2. *Por la extensión: es más confiable si se trata de un .org*”. En ésta se exponen criterios de selección, que se corresponden a tecnologías que emergen de su conocimiento y experiencia, que consideran les permiten juzgar ciertos sitios como confiables: identificamos la noción de popularidad o reconocimiento, éste se establece de manera subjetiva, decidiendo que un nombre conocido es suficiente, así como la extensión del sitio, reconociendo que un .org, tiene más peso que un .com, por ejemplo. Esta última tecnología proviene de su experiencia con la Web. Para finalizar esta parte se plantearon las siguientes dos preguntas:

23. ¿Esta información (sitios/enlaces) te dice algo? ¿Por qué?

24. ¿Te es posible elegir los 5 sitios en este momento? ¿Por qué?

La primera pregunta tiene el objetivo de que los estudiantes comiencen a identificar la dificultad de analizar la información en la presentación dada (lista de sitios/enlaces), sentando la base para la siguiente parte de la actividad, coincidiendo los estudiantes en que la información mostrada no les es de utilidad. La segunda pregunta tiene como función continuar con este proceso, reafirmando que la información dada por la lista les es insuficiente para responder, aunque encontramos nuevamente que los estudiantes utilizan sus propios criterios y proponen listas de sitios sustentadas por tecnologías como las presentadas previamente. Esto se ilustra con la respuesta del estudiante E12 a la pregunta, “¿te es posible elegir los 5 sitios en este momento? ¿Por qué?”: *“nationalgeographic.es, muyinteresante.es, son reconocidos, y denverzoo.org, waza.org y herramientas.educa.madrid.org, tienen una extensión que brinda confianza en la información”*.

Se observa nuevamente que al considerar que la información de la que disponen no les resulta significativa, los estudiantes recurren a tecnologías propias para dar respuesta, de la misma manera en que anteriormente sus propuestas de ordenación se enfocaban a tecnologías que les resultaban familiares como el índice de visitas, constituyendo esto un posible freno para que los estudiantes lleguen a proponer la técnica de ordenación esperada.

#### **6.3.4 Representaciones de la estructura de la Web: grafo y matriz**

Una vez planteada la lista de sitios y sus enlaces se pide a los estudiantes que representen la información gráfica y tabularmente, esperando que surjan representaciones que asemejen un grafo dirigido y su matriz de incidencia y que utilicen ambos para analizar el comportamiento de la estructura. Con el objetivo de que identifiquen las concentraciones de enlaces a ciertos sitios como relación entre número de enlaces e importancia, para después convertirla en un criterio de ordenación.

El grupo llegó a la conclusión de que resulta complicado hacer una interpretación de esta información, así que se propuso buscar otras formas de plasmarla.

25. ¿Cómo representarías de manera gráfica los sitios y sus enlaces?
26. ¿Cómo representarías los sitios y enlaces en una tabla?
27. ¿Qué representación te parece más útil? ¿Por qué?

Se encontró que muchos estudiantes demuestran confusión al solicitarles una representación gráfica y una tabla de la información de sitios y enlaces de manera abierta. A pesar de esto, fue posible encontrar las representaciones esperadas, que asemejaran un grafo y que por lo tanto permitieran analizar la organización de los sitios. Encontrando incluso un estudiante que planteó directamente una relación con teoría de grafos.

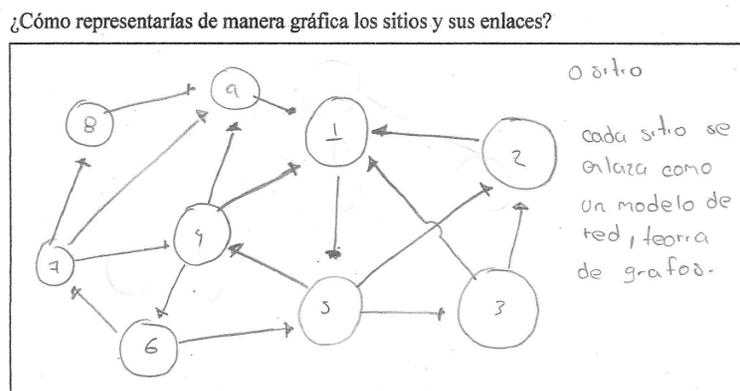


Figura 1. Aproximación al grafo, relación con Teoría de Grafos (Equipo 1)

¿Cómo representarías de manera gráfica los sitios y sus enlaces?

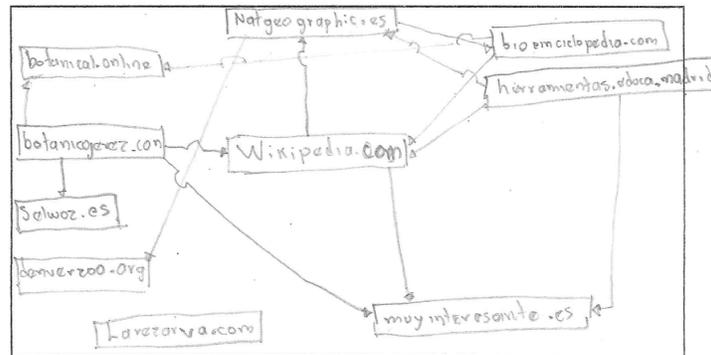


Figura 2. Aproximación al grafo (Equipo 2)

Sin embargo, debido a la confusión mencionada previamente, también encontramos representaciones de otros tipos: desde gráficas de barras, hasta diagramas de tipo conceptual, árboles e histogramas. Estas representaciones parecen estar influenciadas por la asignatura en que se implementó la organización didáctica (Investigación de operaciones), los estudiantes parecen haber considerado que la actividad tenía que estar relacionada con los contenidos estudiados en la asignatura. También encontramos representaciones influenciadas por la ingeniería en sistemas computacionales, siendo el caso de los árboles y diagramas de base de datos. Esto ocurrió con mayor frecuencia al pedir una tabla, confundiendo los estudiantes con “tablas” de base de datos. Esto causó que no pudieran llegar a una representación aproximada a la matriz de incidencia como se esperaba.

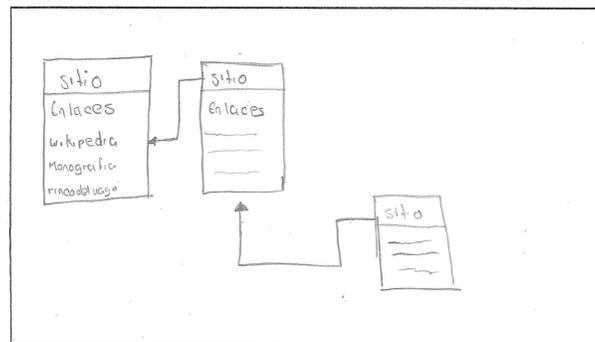


Figura 3. Representación tabular, tablas de Base de Datos (Equipo 1)

Por lo anterior, la intención de comparar ambas representaciones, tratando de que el estudiante planteara ventajas y desventajas no pudo lograrse. Los estudiantes que trataron de responder la pregunta respecto a esta comparación, debido a que no pudieron plantear una aproximación a la matriz de incidencia, coincidieron en afirmar de manera intuitiva que la representación gráfica resulta útil debido a que presenta la información de manera visualmente comprensible, por ejemplo el estudiante E2: “La manera gráfica, es más entendible y da una mejor visión”.

### 6.3.5 Análisis de la estructura de la Web

Una vez planteadas las representaciones, se cuestionaba al estudiante esperando que pudiera observar el comportamiento de la estructura de la Web e identificar el patrón de

comportamiento que nos interesa: la concentración de enlaces recibidos. De manera que esto lo llevara a determinar el número de enlaces como medida de importancia.

28. ¿Observas algún patrón o comportamiento en la gráfica o en la tabla?
29. ¿Te sirve esto para elegir los sitios a mostrar? ¿Por qué?
30. Tras observar ambas representaciones, ¿qué dirías que significa la cantidad de enlaces que recibe cada sitio?
31. Regresando al problema original, ¿podrías utilizar esta información para ordenar los

Sin embargo, debido a las dificultades que encontraron con las representaciones (grafo y matriz), es decir, al no disponer de ellas como herramientas de análisis, muchos estudiantes no pudieron observar la estructura de la Web con la claridad suficiente. Debido a esto, resultó difícil que determinaran la relación de importancia esperada y por lo tanto no llegaron a deducir la primera aproximación al modelo. Sin embargo, encontramos que algunos estudiantes, en los equipos que plantearon aproximaciones cercanas al grafo, lograron identificar las concentraciones de enlaces como significativas, proponiendo verbalmente un modelo como técnica para resolver la problemática de ordenación que considerando la cantidad de enlaces como medida de importancia.

18) ¿Te sirve esto para elegir los sitios a mostrar? ¿Por qué?

E1: Sí, porque a mayor concurrencia de enlaces mayor importancia tendrá la página.

Es importante mencionar, que uno de los estudiantes expresó estar abiertamente en contra de la noción de que la cantidad de enlaces recibidos haga importante a un sitio. Manifestó que Wikipedia, sitio utilizado en el ejemplo, es enlazado constantemente por muchos sitios, pero su información no es realmente confiable debido a un alto nivel de manipulación. Esto nos dice que la elección de sitios para la situación ejemplo deberá manejar sitios más “serios”, evitando aquellos como Wikipedia, RincondelVago, Monografías, por citar algunos ejemplos. Estos sitios son presentados a lo largo de su formación como poco confiables por sus profesores, idea que lo acompañó en el resto del proceso de estudio, recurriendo entonces a una técnica basada en el análisis del contenido del sitio. Esto deja entrever que el estudiante recurrió a la tecnología de que un sitio de baja calidad, tiende a enlazar una gran cantidad de sitios, esto significa que no distingue la diferencia entre enlaces recibidos y enlaces salientes. A pesar de que el resto del equipo no compartía su opinión el estudiante mantuvo su propuesta, influenciando a uno de sus compañeros de equipo, efecto que pudo observarse en la segunda actividad. Esta situación pudo identificarse al responder el estudiante las preguntas 12 y 13:

12) ¿Esta información (sitios/enlaces) te dice algo? ¿Por qué?

E13: Sí, que todos se enlazan a Wikipedia, una página en la que no se puede confiar porque todos pueden hacer modificaciones en su información, por lo que no es confiable.

13) ¿Te es posible elegir los 5 sitios en este momento? ¿Por qué?

E13: No, porque primero necesito analizarlos.

Durante el análisis pudimos identificar que, a pesar de que se les presente la noción de la concentración de enlaces, los estudiantes continúan recurriendo principalmente a tecnologías propias: conteo de visitas, encuestas de satisfacción, frecuencia de actualización, extensión del sitio, como medidas de popularidad, proponiéndolos como criterios. Tecnologías que, al tiempo que facilitan la comprensión del contexto de la situación, demuestran que también pueden convertirse en un freno. Consideramos que de haber planteado a los estudiantes el volumen real de sitios y enlaces, en el orden de los billones (tecnología perteneciente al análisis, pero que no se consideró en el diseño), éstos habría desechado las tecnologías evidenciadas. Esto se debe a que no sería eficiente llevar el control de la información de una cantidad tan grande de sitios, siendo estas últimas (eficiencia y ahorro de espacio de memoria) tecnologías de gran importancia para Ip, como se identificó en el análisis praxeológico.

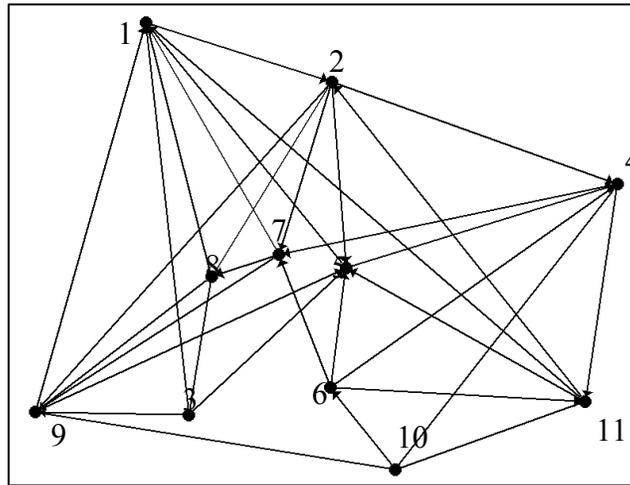
#### **6.4 Segunda Actividad, reorganización del proceso de estudio**

Como se ha mencionado previamente, esta segunda actividad tuvo que ser reorganizada tras un primer análisis de la implementación de la primera actividad, debido a las dificultades encontradas al momento de pedir las representaciones a los estudiantes. Por ello la actividad tuvo que ser modificada, partiendo de la idea de que la razón principal por la cual los estudiantes no dedujeron la primera aproximación era por la falta de una representación que les permitiera analizar la estructura de los enlaces y determinar el patrón en ellos. Esto fue propuesto debido a que se trata de la misma relación identificada entre ambas praxeologías (representación de la estructura de la Web y determinación del criterio de importancias) durante el análisis praxeológico. Esta reorganización de la actividad y el tiempo del que se disponía para realizar la implementación, provocó que fuera necesario dejar de lado el trabajo matemático con los criterios de importancia, recurriendo solo al planteamiento de cuestiones que se consideró, provocarían en el estudiante la evaluación y refinación del primer criterio. Se retomó la parte cuatro y cinco de la primera actividad: Representaciones de la estructura de la Web y Análisis de la estructura de la Web. Se retomó la lista de sitios/enlaces ejemplificada, pero en lugar de solicitar las representaciones se presentó a los estudiantes con el grafo dirigido y la matriz de incidencia, ignorando los nombres de los sitios del ejemplo para reducir la complejidad. Se buscó con esto que los estudiantes poseyeran de manera previa las herramientas necesarias para realizar su análisis, de manera que estuvieran en condiciones de plantear las aproximaciones al criterio de importancias, así como reconocer la utilidad de las representaciones para analizar la información.

##### **6.4.1 Análisis de la estructura de la Web**

En esta primera parte de la actividad, se les cuestionó nuevamente de manera similar, como se hizo en la primera actividad.

Uno de los miembros del equipo presentó la siguiente representación de los sitios y sus enlaces:



	A	B	C	D	E	F	G	H	I	J	K
A	0	0	1	0	1	0	1	1	1	0	1
B	1	0	0	0	0	0	0	1	1	0	1
C	0	0	0	0	0	0	0	0	0	0	0
D	0	1	0	0	1	1	0	0	0	1	0
E	0	1	1	0	0	1	0	0	1	0	1
F	0	0	0	0	0	0	0	0	0	1	0
G	1	1	0	1	0	1	0	0	0	0	0
H	1	1	1	0	0	0	1	0	0	0	0
I	0	1	1	0	0	0	1	1	0	1	0
J	0	0	0	1	0	0	0	0	1	0	1
K	1	0	0	1	0	1	0	0	0	1	0

14. ¿Observas algún patrón o comportamiento en la estructura de los sitios y sus enlaces?

En un principio no muchos estudiantes detectaron el patrón o relación esperada entre la concentración de enlaces y la importancia. Aunque si fue posible encontrar estudiantes que identificaran la concentración de enlaces como medida de importancia, que más adelante causaron que otros estudiantes en sus respectivos equipos pudieran identificarlo también. Al cuestionarles sobre la utilidad de las representaciones, varios estudiantes dieron respuestas similares a la primera actividad: que la representación gráfica les permite comprender mejor la estructura. Sin embargo, en esta ocasión también nos encontramos con que algunos prefirieron la matriz, encontrando sustentos tecnológicos como la “precisión, sencillez y rapidez” con que permite la lectura de la información. Se ilustra con la siguiente pregunta:

2) ¿Cuál te parece más fácil de analizar? ¿Por qué?

E2: La matriz, porque se puede leer con más precisión, sencillez y rapidez.

E14: La matriz, porque es más sencillo verificar que sitios tienen enlace y cuáles sitios no lo tienen, ya que las líneas del primer diagrama podrían tener errores al comparar sus enlaces.

Al hablar de posibles “errores al comparar sus enlaces”, los estudiantes manifiestan de fondo las dificultades que aparecen al analizar el grafo, el cual va adquiriendo mayor complejidad al considerar mayores cantidades de sitios y enlaces. De tal manera, que aunque al principio consideraban que el grafo era una herramienta clara, esto se debía a que en las representaciones propuestas no consideraban el total de enlaces. El observar el grafo completo los hace reconocer dicha complejidad y decantarse por la representación matricial, que les resulta ahora más organizada y comprensible. Esto podría sentar las bases para plantear el número total, real, de sitios y enlaces que conforman la Web (en el orden de los billones), que haría necesario el uso de una matriz. El uso de esta tecnología, identificada en el análisis, no se consideró para la actividad debido a que su utilidad se observa fuertemente como motivación para adaptaciones a los criterios de importancia, particularmente haciendo énfasis en la eficiencia al determinar una técnica para calcular la importancia, parte que no fue considerada en el diseño.

3) ¿Qué significa que un sitio reciba una gran cantidad de enlaces de otros sitios?

E3: Eso significa que ese sitio es de confianza o que ese sitio al tener información de otros sitios hace pensar que tendrá información más completa.

4) ¿Qué significa que no reciba ninguno como en el caso del sitio 3?

E3: Quiere decir que este sitio no tiene otros sitios conectados por lo tanto es muy baja su fiabilidad

5) Si trataras de ordenar los sitios como se propuso en la actividad anterior, sin consultar cada uno para verificar su contenido, ¿crees que la cantidad de enlaces recibidos pueda servirte como criterio de ordenación? ¿Por qué?

E3: Sí, porque ordenaría el que tenga mayor cantidad de enlaces al que tenga menos, así consecutivamente, yo sé que estos a mi criterio, serán de mayor importancia y entre más abajo queden de la lista menor será la importancia del sitio.

Como se mencionó previamente encontramos que la mayor parte de los estudiantes, apoyados por las representaciones dadas, identifican la relación entre concentración de enlaces y la importancia de un sitio. Sin embargo, las tecnologías que utilizan nos indican que algunos no necesariamente están identificando la relación esperada. Por ejemplo, el estudiante E menciona que un sitio que recibe una gran cantidad de enlaces tendrá “información más completa” que proviene de los sitios enlazados, lo que puede ser equiparado a tener una gran cantidad de fuentes. Esto muestra que no hace diferencia entre enlaces salientes (cantidad de fuentes) y enlaces entrantes (referencias de otros sitios), por lo que aunque llega a un criterio de importancia que es correcto, lo hace utilizando tecnologías incompletas.

Sin embargo, aún en los estudiantes que utilizan la matriz como herramienta de análisis e identifican la concentración de enlaces como patrón de comportamiento relacionado a

la importancia, encontramos aproximaciones que se apartan de lo buscado. Identificamos como causa la misma tecnología identificada previamente: los estudiantes no están diferenciando entre enlaces salientes o entrantes. Encontramos también otra tecnología evidenciada por los estudiantes: un sitio que hace spam, tiende a enlazar a una gran cantidad de sitios para hacerse promoción. Estas dos tecnologías causan que en lugar de identificar a un sitio como confiable por la gran cantidad de enlaces, se le asocie con un sitio poco confiable, o productor de spam. Esto se relaciona también con la noción presentada anteriormente, del estudiante que identifica a Wikipedia como sitio poco confiable, que lleva a que él y otro estudiante de su equipo, propongan un criterio de ordenación inverso, como se ilustra a continuación. Consideramos que estas tecnologías podrían provocar dificultades en una organización didáctica que considere la matematización y adaptación de los criterios de importancia.

3) ¿Qué significa que un sitio reciba una gran cantidad de enlaces de otros sitios?

E14: Que crea spam y la información no puede ser relevante.

4) ¿Qué significa que no reciba ninguno como en el caso del sitio 3?

E14: La información presentada en el sitio 3 es de mayor confianza.

#### **6.4.2 Aproximaciones a los criterios de importancias**

Algunos estudiantes demostraron problemas al momento de analizar la estructura de los sitios y sus enlaces, sin embargo esto parece ser fuertemente influenciado por la noción persistente de que el contenido de un sitio y las visitas recibidas deben ser un criterio de ordenación suficiente. Algunos estudiantes lo consideraron, pero sin aceptarlo como criterio principal, proponiendo técnicas que lo combinan con el análisis de contenido o contadores de visitas. Previendo que una situación así podría ocurrir, se propone directamente al estudiante que la cantidad de enlaces servirá para establecer la confiabilidad de un sitio, haciendo una analogía con “recomendaciones”, contexto tomado de la sociometría.

Tras analizar la estructura anterior, otro miembro del equipo propuso lo siguiente:

Si los sitios fueran personas, podríamos ver los enlaces como ‘recomendaciones’, cuando un sitio enlaza a otro nos está invitando a consultarlo. Si un sitio es ‘recomendado’ muchas veces nos indica cierto grado de confiabilidad o importancia de su contenido, sin tener que consultarlo previamente.

6. De acuerdo a esta propuesta, escribe una expresión que te permita determinar la ‘importancia’ de un sitio.

7. Aplicando el criterio, ¿cuál es el orden resultante para los 11 sitios?

Este contexto tiene como objetivo mostrar directamente la relación importancia-cantidad de enlaces, para los estudiantes que no la hubieran identificado previamente. Esto tuvo un efecto positivo, aunque sin poder evitar que algunos estudiantes propusieran criterios combinados con las características mencionadas anteriormente.

Partiendo de que en este punto los estudiantes tendrían mucho más clara la relación enlaces-importancia, se planteó el caso de que dos sitios tuvieran el mismo número de enlaces (situación que aparece en el ejemplo), pidiendo al estudiante que considerara un criterio para ordenarlos en ese caso.

6) De acuerdo a esta propuesta, escribe una expresión que te permita determinar la 'importancia' de un sitio.

E4: "Recomendaciones = fiabilidad"

7) Aplicando el criterio, ¿cuál es el orden resultante para los 11 sitios?

E: A,I,B,D,E,G,H,J,K,F.

8) ¿Crees que este criterio es suficiente para resolver la problemática original? ¿Por qué?

E4: Sí, porque ya los 11 sitios estás ordenados del más al menos fiable.

9) Si dos o más sitios reciben el mismo número de enlaces, ¿quién se colocaría en primer lugar en la lista sin tener que consultarlos?

E4: El sitio de mayor ordenación. Ej. Si el A y el C tienen el mismo número de enlaces como primera instancia se toma el A.

10) ¿Tienen el mismo valor todos los enlaces?

E4: No, hay sitios con más prestigio que otros

Podemos encontrar que aun cuando los estudiantes identifican la relación de manera más clara, utilizando términos como "prestigio" tras haberles planteado el contexto sociométrico. Sin embargo, a pesar de que reconocen que los enlaces tendrán valores distintos dependiendo del sitio del que provengan, encontramos, por ejemplo, que recurren a tecnologías que les facilitan la resolución de la tarea, como la ilustrada previamente donde se da un orden "alfabético" a la lista de sitios, que determina quien se colocará primero en caso de que múltiples sitios compartan el mismo número de enlaces. Consideramos que esta búsqueda de los estudiantes por evadir la complejidad deberá ser considerada en el diseño, planteando al estudiante ejemplos que muestren que esto no garantiza la calidad de los resultados, recurriendo al contexto original del proceso de estudio: el proyecto de diseño de un buscador para una empresa.

Una vez que el estudiante hubiera respondido y partiendo de que es necesario reconocer la diferencia en el valor de los enlaces para determinar el segundo criterio, se le presentó esta noción de manera directa, para determinar si pueden llegar a su aproximación.

11) Integra esta consideración al criterio anterior. Ahora, ¿cómo calcularías la importancia?

E7: En cantidad de visitas y actualizaciones que recibe el sitio y los sitios a los que enlaza

E: Después de mirar cuales son los sitios más enlazados o recomendados, calcularía la importancia consultando el contenido, para mirar y analizar cuál es el más fiable.

12) Escribe la expresión resultante.

E7: Un sitio con enlaces a otros sitios que están constantemente actualizándose y que reciben mayor número de visitas son más seguros y confiables.

A pesar de las dificultades identificadas en cuanto al criterio para ordenar los sitios con mismo número de enlaces, encontramos que la mayor parte de los estudiantes expresaron la noción de que un enlace proveniente de un sitio más importante, “con más enlaces”, tendría más valor que los demás, de manera que podrían ordenar sitios con la misma cantidad de enlaces. Esta situación les permite comenzar el proceso de refinamiento de la técnica de ordenación, que llevaría a plantear el segundo criterio de importancias. Sin embargo, a pesar de que los estudiantes identificaron el criterio de valor de los enlaces, la mayoría expresa técnicas que tratan de mezclar criterios basados en el contenido del sitio, con el número de enlaces. Finalmente, se observa que al solicitar una expresión del criterio de ordenación, los estudiantes solo los expresaron de manera verbal, misma situación que al pedirles “calcular la importancia”, donde respondieron con una lista de los sitios ordenados, sin mostrar intentos de expresar alguna expresión matemática o algoritmo.

## 6.5 CONCLUSIÓN

Este análisis muestra primeramente, que el contexto propuesto permite a los estudiantes reconocer la tarea principal sin mucha dificultad, determinar un método de ordenación para una búsqueda. A lo largo del proceso de estudio los estudiantes evidencian múltiples tecnologías: sitios spam enlazando gran cantidad de sitios, visitas como indicador de popularidad, encuestas de satisfacción y cantidad de fuentes como medida de calidad de un sitio, por ejemplo, que le sirven a los estudiantes como posibles criterios, y que emplean para proponer listas de sitios. Esto viene influenciado por la asignatura de  $E(DI)$ , Estructura de datos, que se encuentran cursando, aunque las tecnologías también provienen del conocimiento propio. Pudimos identificar que esto llega a causar ciertas dificultades que no pudieron ser consideradas durante el diseño. Un punto importante es que no se abordó dentro del proceso de estudio el volumen de sitios de la Web, esperando que los estudiantes lo consideraran por sí mismos. Aunque esta decisión también se debió a que esta tecnología resultaría mucho más útil al momento de estudiar la matematización del modelo. Esta situación causó que el estudiante arraigara fuertemente técnicas dependientes del análisis del contenido de cada sitio, así como diferentes técnicas de conteo, las cuales se hubieran invalidado de haber considerado la cantidad real de sitios que componen la red. Por lo tanto, para un diseño más sólido resultará necesario exponer al estudiante rápidamente a esta tecnología. A pesar de que la situación anterior fue un freno para que los estudiantes llegaran a deducir el modelo, al menos de manera parcial, algunos pudieron llegar a plantearlo en la primera actividad.

De manera similar, la problemática encontrada al trabajar las representaciones de la Web: el grafo dirigido y la matriz de incidencia, supusieron un freno para el estudiante. Sin embargo, debido a la estrecha relación entre estas representaciones y los criterios de importancia, como herramientas de análisis necesarias para su aproximación, estas dificultades tuvieron que ser atendidas sobre la marcha con la reorganización expuesta en el capítulo. Para mitigar esta situación y de manera que la importancia de las representaciones como herramientas de análisis quedara clara para los estudiantes, se considera que sería posible utilizar la problemática de los buscadores Web para que los

estudiantes analicen y construyan el concepto de grafo dirigido y matriz de incidencia dentro del curso de Matemáticas discretas. Para esto se utilizarían la praxeología representación de la estructura de la Web como única base para el diseño de una nueva organización didáctica. Consideramos posible que tal proceso de estudio utilice también contextos análogos (abordados en el Capítulo 4), provenientes de ciencias sociales, por ejemplo. Esto con el objetivo de que los estudiantes puedan construir el concepto de grafo dirigido y matriz de incidencia, al tiempo que estudia su utilidad como herramienta para analizar estructuras de información. Esto lo prepararía para enfrentar una organización didáctica diseñada sobre la praxeología determinación del criterio de importancias, donde podría abordar el proceso completo de deducción y matematización del modelo.

También se considera necesario, estudiar brevemente la manera en que están enlazados los sitios en la Web, haciendo énfasis en lo que significa un enlace entrante (recomendación) y uno saliente (señalando la fuente que sirve de respaldo), de manera que se puedan evitar los modelos invertidos que se observaron durante la implementación, así como aquellas tecnologías que relacionan la alta concentración de enlaces con una baja confiabilidad. Además de utilizar sitios de nombre conocido para los estudiantes, pero que tengan una imagen de “seriedad” para él, de manera que se evite que las ideas negativas sobre la confiabilidad de los sitios puedan entorpecer el proceso de estudio, o llevar a los estudiantes a plantear modelos erróneos. Al mismo tiempo, sería de mucha utilidad que la implementación se llevara a cabo dentro de un curso de álgebra lineal. Idealmente después de haber aplicado la actividad propuesta para matemáticas discretas, como se mencionó previamente, de manera que los estudiantes tengan una noción más clara del funcionamiento de la Web, y puedan no solo deducir los criterios, sino plantearlos y trabajarlos matemáticamente. Consideramos también, que esto permitiría que el diseño de la nueva organización didáctica (praxeología determinación del criterio de importancias) involucrara el estudio de los valores y vectores propios, como modelo de la situación. Esto sería muy importante, debido a que es un tema de gran utilidad dentro de la disciplina, pero que solo es abordado en la materia de álgebra lineal si el profesor en turno lo considera posible (dentro del Instituto Tecnológico de Chilpancingo).

Finalmente, consideramos que implementar dos organizaciones didácticas de la manera propuesta, cobra mayor importancia debido a que la asignatura de matemáticas discretas se imparte en el primer semestre. Se observó que la falta de familiaridad de los estudiantes con este tipo de actividades, que requieren analizar un problema para determinar una solución probable los lleva a proponer soluciones de tipo empírico (con apenas un par de menciones sobre el uso de un algoritmo), mostrando una cierta reticencia al empleo de técnicas matemáticas. Prefiriendo mantener, por ejemplo, las técnicas y el modelo de importancias deducido en términos verbales. Se espera que esta situación mejore al presentar actividades de modelación desde el primer semestre, en las asignaturas de  $E(M)$ . Actividades que les permitan abordar la utilidad de las herramientas matemáticas que estudian (en este caso grafos y sus representaciones matriciales, sistemas de ecuaciones lineales y cálculos con matrices). De manera que

más adelante puedan recurrir a la matemática de una manera mucho más fluida y natural al enfrentar problemas de mayor complejidad, en los que el uso de la matemática no sea evidentemente requerido.

## 7 CONCLUSIONES

Como se planteó al comienzo de este trabajo de investigación, existe una necesidad, documentada al menos desde la década de los 80's, de profesionistas que sean capaces de utilizar sus conocimientos matemáticos en la resolución de problemáticas “reales” en su ambiente de trabajo. Esta necesidad, manifestada en el campo laboral se ha venido traduciendo en una necesidad de profesionistas “matemáticamente competentes”. Esto ha repercutido fuertemente en las formaciones profesionales, particularmente en las de futuros ingenieros, las cuales tienen la responsabilidad de preparar adecuadamente al estudiante para su futura vida laboral. Esta situación ha ido adquiriendo un papel cada vez más apremiante debido al rápido avance industrial y científico, que demanda atender problemáticas cada vez más complejas y requiere que los profesionistas sean preparados durante su formación para afrontarlas.

Para atender esta problemática, la modelación se muestra como una herramienta fundamental, como una interfaz que permite llevar los conocimientos teóricos a la práctica profesional. Por lo que actualmente se ha recalcado su papel como “puente” entre la teoría matemática y sus aplicaciones en la industria, permitiendo al estudiante reconocer que es posible entablar relaciones entre la matemática y el “mundo real”. Sin embargo, como se muestra en el Capítulo 1, integrar la modelación a la formación posee una gran complejidad. Consideramos que el análisis realizado, ilustra el papel fundamental que tiene el identificar contextos laborales y analizar el papel que la modelación matemática desempeña en ellos para lograr esta integración. Esto con el objetivo de posibilitar y sustentar el diseño de actividades de modelación que permitan abordar dichos contextos en el aula, mostrando al estudiante cómo es que la modelación se lleva a cabo en lo que será su campo laboral. Aunque nuestro interés es la formación de ingenieros, el integrar actividades de modelación a la misma constituye una problemática demasiado amplia, por lo que es necesario enfocarnos a una formación en particular, eligiendo en esta investigación la ingeniería en sistemas computacionales, así como una institución que imparta tal formación y por lo tanto nos permita analizarla, eligiendo para ello al Instituto Tecnológico de Chilpancingo.

Para el fuerte interés en el análisis de la actividad ingenieril, las herramientas tomadas de la Teoría Antropológica de lo Didáctico (TAD) resultaron adecuadas a las necesidades planteadas. Permitiendo analizar una actividad particular de la práctica profesional del ingeniero en sistemas computacionales a través del uso de sus nociones de institución y de praxeología. Estas nociones fungieron como herramientas que permitieron analizar, no solo la actividad de modelación, sino la formación misma del ingeniero en sistemas computacionales en términos de instituciones. Particularmente dos: la práctica profesional  $I_p$  de la ingeniería en sistemas computacionales (en este caso, la empresa “Google” que desarrolló el método PageRank) y la institución de enseñanza de matemáticas discretas y de álgebra lineal (pertenecientes ambas a  $E(M)$ ), aunque dentro del análisis resultó posible identificar también la posible intervención de

la institución de enseñanza de estructura de datos, perteneciente a la formación de especialidad  $E(DI)$ . Esto permitió evidenciar que un análisis de contextos profesionales, en particular, de una actividad de modelación puede facilitar la generación de relaciones entre estas 3 instituciones.

Esto nos permitió traducir nuestro objetivo en: el diseño de una organización didáctica que generara una relación entre estas instituciones,  $E(M)$ ,  $E(DI)$  e  $Ip$ , la cual obtendría su sustento del análisis previo.

Para alcanzar este objetivo elegimos utilizar una propuesta metodológica planteada originalmente por Macías (2012), la cual no solo está orientada al diseño de actividades de modelación para formaciones de ingenieros, sino que está elaborada en el marco de la TAD. Esta propuesta, que sigue en proceso de refinamiento, propone el uso de cuatro fases: *Elección de un contexto extra-matemático*, *Análisis praxeológico e identificación de un modelo matemático*, *Análisis del modelo en  $E(M)$* , *Diseño de la actividad didáctica*. La flexibilidad de esta propuesta metodológica permitió la adaptación de sus fases de acuerdo a las necesidades y condiciones que rodearon la investigación, adaptaciones que son abordadas en el Capítulo 3.

Siguiendo esta metodología, se realizó un análisis praxeológico a la actividad de modelación: construcción del método PageRank, perteneciente al contexto de la programación Web, de la práctica profesional  $Ip$  del ingeniero en sistemas computacionales. Como se indica en la propuesta y debido a que se trata del análisis a una actividad práctica, se recurrió al uso del modelo praxeológico extendido, que permitió analizar más finamente las tecnologías provenientes de la práctica  $Ip$ , que guían y condicionan la actividad. Estos análisis permitieron identificar la existencia de una arborescencia praxeológica, dentro de la cual se identificaron a su vez las siguientes praxeologías que conforman la construcción del método PageRank (la arborescencia praxeológica completa y la descripción de las praxeologías a continuación presentadas se abordan en el Capítulo 4):

1. Representación matemática de la estructura de la Web

Esta praxeología encuentra correspondencia en las asignaturas de matemáticas discretas (grafo dirigido y matriz de incidencia) y álgebra lineal (nociones de matriz) pertenecientes a  $E(M)$ , así como en estructura de datos  $E(DI)$ , donde se estudian ambos modelos (grafo dirigido y matriz de incidencia) desde el punto de vista de la especialidad.

2. Determinación de un criterio de ordenación (importancias)

Esta praxeología encuentra correspondencia en la asignatura de álgebra lineal (sistema de ecuaciones lineales), sin embargo, los modelos más complejos identificados en la misma no encuentran un referente claro, debido a que emplean elementos que no se encuentran presentes en  $E(M)$ . La excepción a esto es el modelo de valores y vectores propios que, de acuerdo al docente con el que se realizó la implementación de la actividad exploratoria, es estudiado

dependiendo de la disponibilidad de tiempo, y a decisión del profesor a cargo de la materia.

### 3. Determinar/calcular las importancias

El modelo “método de las potencias” utiliza conceptos de Álgebra Lineal, y podría ser enseñado en la asignatura de Estructura de Datos como herramienta para el análisis de estructuras de información, debido a que se identificó que se estudia solo parcialmente, como herramienta para determinar caminos dentro de un grafo.

Estas praxeologías y las relaciones identificadas fungieron como base para el diseño de una organización didáctica. Tales relaciones con asignaturas pertenecientes a E(M) y E(DI) fueron establecidas a su vez en un análisis a estas dos instituciones, que sustenta la elección de las dos primeras praxeologías para el diseño de la organización didáctica. Dicha organización se estructuró como un proceso de estudio, que llevó a los estudiantes a enfrentar tareas que emergen de la práctica profesional, pero que son abordadas ahora bajo una lógica escolar. Esto con el objetivo de que los estudiantes reconstruyan las praxeologías seleccionadas. Dentro de la actividad, los estudiantes se enfrentan a tareas que demandan de comprender un problema real, el cual requiere poner en juego tecnologías que la enseñanza, tanto matemática como de la disciplina, ha puesto a su disposición pero también de utilizar tecnologías propias, así como refinarlas y generar nuevas. Proponer, interpretar y validar una técnica o modelo solución es requerido por la actividad, para esto se consideraron los momentos didácticos: momento del primer encuentro, momento exploratorio y momento de construcción del entorno tecnológico-teórico. Estos estructuraron el proceso de estudio, el cual fue diseñado de manera que se hiciera pasar al estudiante por estas etapas. Debido a esto consideramos que la organización didáctica resultante es una Actividad de Estudio e Investigación (AEI), y que, aunque se considera que con los primeros tres momentos es posible hacer pasar al estudiante por un proceso de estudio, no se descarta que los otros tres: *momento de trabajo de la técnica*, *momento de evaluación de la técnica* y *momento de institucionalización*, puedan hacerse presentes durante el mismo. Esto fue observado durante la implementación, mostrando que, sin contar el *momento de institucionalización*, estos momentos se encuentran fuertemente ligados al *momento de construcción del entorno tecnológico-teórico*, el cual puede motivar o provocar su aparición.

Durante la implementación de la organización didáctica, los estudiantes demostraron la capacidad de enfrentar un problema real, proponiendo técnicas de solución (de manera verbal) y justificar las decisiones que toman y las técnicas que proponen con variadas tecnologías, provenientes de la ingeniería *Ip* así como de los conocimientos y experiencia personal. Mostrando que la oportunidad de trabajar una problemática perteneciente a un contexto familiar les permite poner en juego conocimientos (tecnologías teóricas y prácticas) que normalmente no emplean en la clase de matemáticas. Sin embargo, también resultó posible observar dificultades provenientes, principalmente, de la falta de familiaridad de los estudiantes con el tratamiento de problemáticas reales, así como cierta reticencia al momento de emplear sus

conocimientos matemáticos en tales problemáticas, sin que se les hubiera requerido expresamente. Se observa que, al no trabajar actividades de modelación matemática que familiaricen al estudiante con el uso de sus herramientas matemáticas en contextos “reales”, los estudiantes no intentan recurrir a ellas como primera opción y en el momento en que lo hacen, recurren a las herramientas más recientemente estudiadas (es decir, aquellas que pertenecen a la asignatura matemática más reciente que hayan o se encuentren cursando) como primera herramienta de solución. Esto los lleva a intentar utilizar modelos que no necesariamente aportan a la comprensión o resolución de la problemática, mostrando que el estudiante no ha desarrollado una comprensión sobre el funcionamiento y uso de los modelos matemáticos que ha estudiado. A su vez, la implementación mostró que ciertas tecnologías provenientes de la experiencia de los estudiantes pueden afectar negativamente el desarrollo de la actividad, “atrapando” al estudiante en propuestas de técnicas de solución, que mantiene debido a su familiaridad y a las tecnologías mencionadas (estas tecnologías son abordadas en el Capítulo 6). Esto a pesar de que en la práctica no arrojen resultados congruentes o no puedan dar argumentos tecnológicos suficientes o que la sustenten ampliamente.

Estas dificultades requieren que la actividad mantenga su diseño abierto en un primer momento, de forma que se permita a los estudiantes la exploración y comprensión de la problemática. Esto con el propósito de brindarle la oportunidad de presentar sus posibles propuestas de técnicas de solución y las tecnologías que utilizan para sustentarlas y/o validarlas. Esto para que en un segundo momento y utilizando las tecnologías identificadas en el análisis praxeológico como sustento, la actividad se vuelva cerrada, de manera que permita depurar aquellas propuestas que no permiten al estudiante abordar las aproximaciones al modelo que se busca estudiar, evitando el “encasillamiento” identificado. Los resultados del análisis de la implementación de la organización didáctica, así como las características observadas en la primera praxeología identificada (representación matemática de la Web), nos permiten considerar la posibilidad de utilizarla como base para el diseño de una organización didáctica que pudiera implementarse de manera independiente. Y cuyo enfoque sea la construcción de los conceptos de grafo dirigido y matriz de incidencia, a la vez que permite a los alumnos estudiar la utilidad de ambos modelos para el análisis de información. Se propone que esta organización podría ser aplicada en el curso de matemáticas discretas, que se imparte en el primer semestre de la carrera, en donde se abordan ambos conceptos. De manera que podríamos comenzar a familiarizar a los estudiantes con actividades de modelación de manera temprana, así como con el uso de la matemática para la resolución de problemáticas reales. Creemos que una vez expuestos a esta actividad los estudiantes se encontrarían en mejores condiciones para abordar una organización didáctica diseñada sobre la segunda praxeología (definición de un criterio de importancias). Este proceso de estudio, al existir ya el análisis previo sobre la estructura de la Web, podría estar en condiciones de trabajar el modelo de valores y vectores propios, convirtiendo a la organización didáctica en una herramienta para el profesor, que le auxilie en la enseñanza de un modelo que resulta de gran utilidad a la especialidad, pero que no siempre puede ser abordado en el aula.

Este trabajo permite mostrar la complejidad asociada al análisis de un contexto ingenieril y el rol que dentro del mismo, toman los modelos matemáticos. Es necesario introducirnos a las lógicas de estos contextos y tener un repertorio de fuentes donde las elecciones que guían el uso de los modelos se expliciten. Como se plantea en la metodología este análisis también debe ser complementado por uno de las asignaturas que se reconoce pueden relacionarse con las praxeologías identificadas en el análisis de la práctica. De manera que se identifiquen aquellas áreas en las que es susceptible la generación de organizaciones didácticas que relacionen ambas instituciones, Ip y E(M).

Consideramos también que este trabajo **permite** evidenciar que la modelación matemática puede favorecer la vinculación entre la práctica y la formación a partir de análisis praxeológicos finos que sustenten diseños didácticos para la clase de matemáticas. Esta vinculación, **permitirá** que el estudiante no solo **obtenga** las herramientas matemáticas que dicta el plan de estudios, sino que desarrolle las habilidades necesarias para utilizarlos en la resolución de problemáticas “reales”, similares a las que enfrentará dentro de su actividad profesional, cobrando conciencia de la importancia que la matemática tiene para su formación, motivándolo a sacar el mayor provecho posible de su formación matemática. Sin embargo, esto pone también en evidencia la necesidad de continuar realizando estos análisis de la práctica profesional, de manera que puedan seguirse diseñando organizaciones didácticas que impacten en la formación de futuros ingenieros matemáticamente competentes. Esto con el objetivo de que se encuentren adecuadamente preparados para las necesidades actuales del campo laboral.

## 8 REFERENCIAS

Artaud, M. (2007). Some conditions for modelling to exist in mathematics classrooms. In *Modelling and applications in mathematics education* (pp. 371-378). Springer US.

Bissell, C. & Dillon, C. (2000). Telling tales: models, stories and meanings. *For the Learning of Mathematics*, 20(3), 3-11.

Bosch, M., García, F., Gascón, J., & Ruiz Higuera, L. (2006). La modelización matemática y el problema de la articulación de la matemática escolar. Una propuesta desde la teoría antropológica de lo didáctico. *Educación Matemática*, 18(2), 37-74.

Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*, 56(18), 3825-3833.

Niss, M., Blum, W., & Galbraith, P. (2007). *Modelling and applications in mathematics education*. New York: Springer.

Broder, A. (1999). Graph structure in the web. Recuperado de: <http://www9.org/w9cdrom/160/160.html>

Chevallard, Y. (1999). El análisis de las prácticas docentes en la Teoría Antropológica de lo Didáctico. *Recherches en Didactique des Mathématiques*, 19 (2), 221- 266.

Chevallard, Y. (2002). Organiser l'étude. 3. Écologie et régulation [Organize the study 1. Ecology and regulation]. Grenoble, France: La Pensée Sauvage.

D'Andrea, C. (2012). El álgebra detrás de los buscadores de internet. Recuperado de: [http://cms.dm.uba.ar/académico/materias/2docuat2012/algebra\\_lineal/](http://cms.dm.uba.ar/académico/materias/2docuat2012/algebra_lineal/)

Fernández, P. (2004). El secreto de Google y el Álgebra Lineal. *Bol. Soc. Esp. Mat. Apl.* Nro. 30, 115-141.

Franceschet, M. (2011). PageRank: Standing in the shoulders of giants. *Communications of the ACM*, 54, Issue 6, 92-101.

García, F. J., Gascón, J., Ruiz Higuera, L. & Bosch, M. (2006). Mathematical modelling as a tool for the connection of school mathematics, *ZDM The International Journal on Mathematics Education*, 38(3), 226-246.

Geiger, V. (2013). Mathematical Applications, Modelling and Technology as Windows into Industry Based Mathematical Practice. In *Educational Interfaces between Mathematics and Industry* (pp. 271-278). Springer International Publishing.

Henning, H., & Keune, M. (2007). Levels of modelling competencies. In *Modelling and applications in mathematics education* (pp. 225-232). Springer US.

Howson, A. G., J. P. Kahane, P. Lauginie y E. de Turckheim (eds.) (1988). *Mathematics as a Service Subject*. ICMI Study Series, Cambridge, Cambridge University Press.

Lay, D. C. (2001). *Álgebra lineal y sus aplicaciones* (2ª Ed). Prentice Hall Mexico.

Li, T. (2013). Mathematical Modeling Education is the Most Important Educational Interface Between Mathematics and Industry. In *Educational Interfaces between Mathematics and Industry* (pp. 51-58). Springer International Publishing.

Macias, C. (2012). *Uso de las nuevas tecnologías en la formación matemática de ingenieros*. Tesis de maestría no publicada, México, CICATA-IPN.

Martínez, M. (2007). *Detector de partes vocalizadas de la voz esofágica en dispositivos FPGA*. Tesis de maestría no publicada, México, CIDETEC-IPN.

Palmer, M. A., Amat, S., Busquier, S., Romero, P., & Tejada, J. (2013). Mathematics for Engineering and Engineering for Mathematics. In *Educational Interfaces between Mathematics and Industry* (pp. 185-198). Springer International Publishing.

Pollak H. O. (1988). Mathematics as a service subject – why? In A. G. Howson et al. (eds), *Mathematics as a service subject*. (pp. 28-34). Cambridge : Cambridge University Press (Series: ICMI study).

Romo-Vázquez, A. (2009). *Les mathématiques dans la formation d'ingénieurs*, Paris, IREM de Paris.

Castela, C. y Romo-Vázquez, A. (2011). Des mathématiques à l'automatique étude des effets de transposition sur la transformée de Laplace dans la formation des ingénieurs. *Recherches en didactique des mathématiques*, 31(1). 79-130

Romo-Vázquez, A. (2014). La modelización matemática en la formación de ingenieros. *Revista de Educación Matemática*, 26(3), pp. 314-338.

Soto, S. (2013). *Una secuencia didáctica basada en modelación matemática*. Tesis de maestría no publicada. México, CICATA-IPN.

Wake, G. (2007). Considering workplace activity from a mathematical modelling perspective. In *Modelling and applications in mathematics education* (pp. 395-402). Springer US.

Ye, Q. (2013). Incorporating the ideas and methods of mathematical modeling into calculus teaching. In *Educational Interfaces between Mathematics and Industry* (pp. 385-394). Springer International Publishing.