

CARRERA DEL INVESTIGADOR CIENTÍFICO Y TECNOLÓGICO

Informe Científico¹

PERIODO ²: 2015

Legajo N°:

1. DATOS PERSONALES

APELLIDO: Fernandez

NOMBRES: Alejandro

Dirección Particular: Calle:

Localidad: La Plata CP: 1900 Tel:

Dirección electrónica (donde desea recibir información): alejandro.fernandez@lifia.info.unlp.edu.ar

2. TEMA DE INVESTIGACION

Asistencia Semantica para aprendizaje colaborativo y gestión del conocimiento

3. DATOS RELATIVOS A INGRESO Y PROMOCIONES EN LA CARRERA

INGRESO: Categoría: Inv. Asistente Fecha: 4/Mayo/2012

ACTUAL: Categoría: Inv. Asistente desde fecha: 4/Mayo/2012

4. INSTITUCION DONDE DESARROLLA LA TAREA

Universidad y/o Centro: LIFIA (Centro Asociado CIC)

Facultad: Informatica, UNLP

Cátedra: Orientación a Objetos II; Groupware y Software Social

Dirección: Calle: 50 y 120 N°: S/N

Localidad: La Plata CP: 1900 Tel: 4228252

Cargo que ocupa: Profesor Adjunto

5. DIRECTOR DE TRABAJOS. (En el caso que corresponda)

Apellido y Nombres: Gabriel Baum

Dirección Particular: Calle:

Localidad: City Bell CP: Tel:

Dirección electrónica: gbaum@lifia.info.unlp.edu.ar

.....
Firma del Director (si corresponde)

.....
Firma del Investigador

¹ Art. 11; Inc. "e"; Ley 9688 (Carrera del Investigador Científico y Tecnológico).

² El informe deberá referenciar a años calendarios completos. Ej.: en el año 2008 deberá informar sobre la actividad del período 1°-01-2006 al 31-12-2007, para las presentaciones bianuales.

6. EXPOSICION SINTETICA DE LA LABOR DESARROLLADA EN EL PERIODO.

Durante el período informado, el foco de mi trabajo en el Centro Asociado CIC LIFIA, ha sido la gestión del conocimiento en la etapa de recuperación de conocimiento. En ese contexto, he explorado puntualmente el rol de los sistemas de recomendación y la utilización de datos semánticos, interconectados, abiertos (Linked Open Data).

Desarrollamos, en el marco del proyecto de colaboración “A recommender system for the Moodle learning management system and the WikiLOR learning object repository” con la Università Telematica Unitelma Sapienza (Italia) y la Universidad De Tampere (Finlandia), componentes para asistir en la recuperación de mensajes relevantes en foros técnicos y de entornos de aprendizaje. Dicho trabajo combina sistemas de recomendación y herramientas de visualización. Del proyecto resultaron tres artículos científicos:

- D. Distante, A. Fernandez, L. Cerulo, and A. Visaggio, “Enhancing Online Discussion Forums with Topic-Driven Content Search and Assisted Posting,” 2015, pp. 161–180.
- Kuosa, K., Distante, D., Tervakari, A., Cerulo, L., Fernández, A., Koro, J., & Kailanto, M. (2016). Interactive Visualization Tools to Improve Learning and Teaching in Online Learning Environments. *International Journal of Distance Education Technologies*, 14(1), 1–21. doi:10.4018/IJDET.2016010101
- M. Erdt, A. Fernandez, and C. Rensing, “Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey,” *IEEE Trans. Learn. Technol.*, vol. PP, no. 99, pp. 1–1, 2015.

En el marco del proyecto de incentivos “Adaptación de Aplicaciones Web, Web Semántica Social, Software Conducido por Modelos” se comenzó a explorar el fenómeno de la Ciencia Ciudadana como mecanismo para involucrar a voluntarios ciudadanos. De esta tarea nace la plataforma de ciencia ciudadana Cientopolis (<http://cientopolis.lifia.info.unlp.edu.ar>). La Plataforma, que continuará su desarrollo en el bienio 2016-2017, ofrece herramientas y asistencia para el desarrollo de proyectos científicos con la participación de ciudadanos. Experimentos iniciales con voluntarios dan cuenta del potencial de esta metodología de trabajo. Se trata de un proyecto de colaboración multidisciplinar, que involucra profesionales otras disciplinas, en particular de las ciencias de la astronomía, las ciencias sociales, y la arquitectura.

De este trabajo surgieron dos proyectos, que han recibido financiación, y que enmarcarán mis actividades en el próximo período: El proyecto del programa de incentivos “Ingeniería Ágil y Conducida por Modelos, De Software Adaptable para la Web del Conocimiento” y el proyecto “RUC-APS: Enhancing and implementing Knowledge based ICT solutions within high Risk and Uncertain Conditions for Agriculture Production Systems”. Este último, presentado al programa “H2020-MSCA-RISE-2015” de Horizonte 2020 de la unión Europea, enfoca el requerimiento de soluciones integradas que sea capaces de hacer frente a los retos de colaboración entre los diferentes agentes involucrados, desde de las primeras fases de la agricultura

como el diseño de productos, a la cosecha, hasta las fases posteriores a la cosecha y procesamiento final.

Los resultados de estas actividades de investigación se reflejan a su vez en mis actividades académicas, en el curso de grado Orientación a Objetos II de Licenciatura en informática y en el curso Groupware y Software Social del Doctorado en Ciencias Informáticas de la UNLP.

7. TRABAJOS DE INVESTIGACION REALIZADOS O PUBLICADOS EN ESTE PERIODO. 7.1 PUBLICACIONES.

Distante, D., Fernandez, A., Cerulo, L., & Visaggio, A. (2015). Enhancing Online Discussion Forums with Topic-Driven Content Search and Assisted Posting. In Knowledge Discovery, Knowledge Engineering and Knowledge Management (pp. 161–180). doi:10.1007/978-3-319-25840-9_11

Online forums represent nowadays one of the most popular and rich repository of user generated information over the Internet. Searching information of interest in an online forum may be substantially improved by a proper organization of the forum content. With this aim, in this paper we propose an approach that enhances an existing forum by introducing a navigation structure that enables searching and navigating the forum content by topics of discussion. Topics and hierarchical relations between them are semi-automatically extracted from the forum content by applying Information Retrieval techniques, specifically Topic Models and Formal Concept Analysis. Then, forum posts and discussion threads are associated to discussion topics on a similarity score basis. Moreover, to support automatic moderation in websites that host several forums, we propose a strategy to assist a user writing a new post in choosing the most appropriate forum into which it should be added. An implementation of the topic-driven content search and navigation and assisted posting forum enhancement approaches for the Moodle learning management system is also presented in the paper, opening to the application of these approaches to several real distance learning contexts. Finally, we also report on two case studies that we have conducted to validate the two approaches and evaluate their benefits.

S. Jaramillo Valbuena, S. A. Cardona, and A. Fernandez, “Minería de datos sobre streams de redes sociales, una herramienta al servicio de la Bibliotecología,” Inf. Cult. y Soc., no. 33, 2015.

Los sistemas de soporte al trabajo colaborativo, Groupware, son una herramienta valiosa en contextos en los cuales se requiere la participación de un grupo de personas para llevar a cabo una tarea. Las interacciones entre las personas que los utilizan generan grandes flujos de datos (streams) no estructurados. Estos streams pueden analizarse para estudiar aspectos tales como influencia, relaciones de cercanía, opinión y para la generación de recomendaciones. Desde la perspectiva de la minería de datos, el procesamiento de estos streams plantea importantes desafíos. Los algoritmos de minería a utilizar deben adaptarse a la alta velocidad en que llegan los datos, a la diversidad de las fuentes de datos y su estructura, a variabilidad de los

datos en el tiempo y a trabajar sin restricciones de memoria. Este artículo revisa el estado del arte en lo referente a algoritmos de minería de datos sobre streams originados en sistemas groupware. Se presenta una revisión de las técnicas más representativas y de cómo cada una de ellas aporta al descubrimiento de conocimiento. Específicamente, se analiza la gestión de la información proveniente de redes sociales. Para concluir se presentan algunos de los problemas que son objeto de investigación activa.

M. Erdt, A. Fernandez, and C. Rensing, "Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey," *IEEE Trans. Learn. Technol.*, vol. PP, no. 99, pp. 1–1, 2015.

The increasing number of publications on recommender systems for Technology Enhanced Learning (TEL) evidence a growing interest in their development and deployment. In order to support learning, recommender systems for TEL need to consider specific requirements, which differ from the requirements for recommender systems in other domains like e-commerce. Consequently, these particular requirements motivate the incorporation of specific goals and methods in the evaluation process for TEL recommender systems. In this article, the diverse evaluation methods that have been applied to evaluate TEL recommender systems are investigated. A total of 235 articles are selected from major conferences, workshops, journals, and books where relevant work have been published between 2000 and 2014. These articles are quantitatively analysed and classified according to the following criteria: type of evaluation methodology, subject of evaluation, and effects measured by the evaluation. Results from the survey suggest that there is a growing awareness in the research community of the necessity for more elaborate evaluations. At the same time, there is still substantial potential for further improvements. This survey highlights trends and discusses strengths and shortcomings of the evaluation of TEL recommender systems thus far, thereby aiming to stimulate researchers to contemplate novel evaluation approaches.

A. Diaz, A. Fernandez, and M. Chiaradia, "WikiLOR: a Collaborative Approach for Improving Learning Object Metadata," in *The proceedings of LACLO 2015*, 2015.

The quality of metadata in learning object repositories is frequently low; metadata records are incomplete, inaccurate, and out of date. This paper introduces a collaborative, metadata editing approach to improve metadata quality and completion by the community's effort. WikiLOR is a wiki that acts as both: a learning object repository and a collaborative metadata editor. WikiLOR proposes a wiki-way for editing both a learning object's content and its metadata. WikiLOR encourages collaborative editing, promotes discussion, helps in conflict resolution and supports version management. The paper additionally documents an experiment that reveals the qualities of WikiLOR and compares it to other benchmark applications.

7.2 TRABAJOS EN PRENSA Y/O ACEPTADOS PARA SU PUBLICACIÓN.

Kuosa, K., Distante, D., Tervakari, A., Cerulo, L., Fernández, A., Koro, J., & Kailanto, M. (2016). Interactive Visualization Tools to Improve Learning and Teaching in Online Learning Environments. *International Journal of Distance Education Technologies*, 14(1), 1–21. doi:10.4018/IJDET.2016010101

This paper presents two visualization tools that introduce new functionalities into learning management systems in order to improve learning and teaching processes in online courses. The first tool implements learning analytics functionalities by exploiting automatically recorded user logs. The tool analyzes user's activity logs and builds interactive views that provide valuable insights into the students' learning process and participation in a course, both to teachers and students themselves. The second tool extends navigation and search functionalities usually available in a discussion forum of an e-learning environment. It uses information retrieval techniques to analyze forum threads and messages and automatically identifies discussion topics. Then it enhances the original forum with a topic-driven navigation structure and an interactive search graph. Both tools have been developed as plugins for the Moodle learning management system, but their base analysis processes and techniques can be adopted into any learning management system.

Jurado, J. L., Fernandez, A., & Collazos, A. C. (2015). Applying gamification in the context of knowledge management. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*. Graz, Austria. doi:10.1145/2809563.2809606

In this paper we present the preliminary results of an ongoing project called (KMM+G). It explores an approach to improve participation in knowledge management processes, in software development companies, through gamification. The paper documents a case study in which we compared two approaches to knowledge construction and refinement. One approach is based on email communication, document templates and the guidance of a project leader. The other approach incorporates gamification mechanics known as PBL (points, medal, and leaderboard). We observed that gamification mechanics substantially improve aspects like: participation, contribution, collaboration and knowledge refinement in the process of software development.

7.3 TRABAJOS ENVIADOS Y AUN NO ACEPTADOS PARA SU PUBLICACION.
Incluir un resumen de no más de 200 palabras de cada trabajo, indicando el lugar al que han sido enviados. Adjuntar copia de los manuscritos.

7.4 TRABAJOS TERMINADOS Y AUN NO ENVIADOS PARA SU PUBLICACION.
Incluir un resumen de no más de 200 palabras de cada trabajo.

7.5 COMUNICACIONES. *Incluir únicamente un listado y acompañar copia en papel de cada una. (No consignar los trabajos anotados en los subtítulos anteriores).*

Zuccarelli, G., Lanzarini, L., & Fernandez, A. (2015). *Instancias de ontologías de la Web Semántica mejoradas a través de Minería de Texto. III Jornadas de Cloud Computing y Big Data*. La Plata.

7.6 INFORMES Y MEMORIAS TECNICAS. *Incluir un listado y acompañar copia en papel de cada uno o referencia de la labor y del lugar de consulta cuando corresponda.*

8. TRABAJOS DE DESARROLLO DE TECNOLOGÍAS.

8.1 DESARROLLOS TECNOLÓGICOS. *Describir la naturaleza de la innovación o mejora alcanzada, si se trata de una innovación a nivel regional, nacional o internacional, con qué financiamiento se ha realizado, su utilización potencial o actual por parte de empresas u otras entidades, incidencia en el mercado y niveles de facturación del respectivo producto o servicio y toda otra información conducente a demostrar la relevancia de la tecnología desarrollada.*

8.2 PATENTES O EQUIVALENTES. *Indicar los datos del registro, si han sido vendidos o licenciados los derechos y todo otro dato que permita evaluar su relevancia.*

8.3 PROYECTOS POTENCIALMENTE TRANSFERIBLES, NO CONCLUIDOS Y QUE ESTAN EN DESARROLLO. *Describir objetivos perseguidos, breve reseña de la labor realizada y grado de avance. Detallar instituciones, empresas y/o organismos solicitantes.*

8.4 OTRAS ACTIVIDADES TECNOLÓGICAS CUYOS RESULTADOS NO SEAN PUBLICABLES *(desarrollo de equipamientos, montajes de laboratorios, etc.).*

8.5 *Sugiera nombres (e informe las direcciones) de las personas de la actividad privada y/o pública que conocen su trabajo y que pueden opinar sobre la relevancia y el impacto económico y/o social de la/s tecnología/s desarrollada/s.*

9. SERVICIOS TECNOLÓGICOS. *Indicar qué tipo de servicios ha realizado, el grado de complejidad de los mismos, qué porcentaje aproximado de su tiempo le demandan y los montos de facturación.*

10. PUBLICACIONES Y DESARROLLOS EN:

10.1 DOCENCIA

10.2 DIVULGACIÓN

11. DIRECCION DE BECARIOS Y/O INVESTIGADORES. *Indicar nombres de los dirigidos, Instituciones de dependencia, temas de investigación y períodos.*

Guido Zuccarelli. "Curado de datos provenientes de Linked Open Data". Becario de experiencia laboral. SPU 2674. Lugar de trabajo: LIFIA. Inicio 1/7/2012

Matías Celasco, LIFIA, SPU 2674. "Crowdsourcing y Gamification". Lugar de trabajo: LIFIA. Inicio 1/5/2014

Juan Ignacio Yañez, LIFIA. Pasante LIFIA. "Crowdsourcing y Gamification".
Lugar de trabajo: LIFIA. Inicio 1/5/2014

12. DIRECCION DE TESIS. *Indicar nombres de los dirigidos y temas desarrollados y aclarar si las tesis son de maestría o de doctorado y si están en ejecución o han sido defendidas; en este último caso citar fecha.*

Dos alumnos con mi dirección, Geraldine Yarasca y Jhonny Barrera se encuentran completando créditos del Doctorado en Ciencias Informáticas de la UNLP.

Completaron su tesina de grado con mi dirección, durante el 2015, los alumnos: Guido Zucarrelli, Francisco Apesteguía, Máximo Zarza y Tomás Campodónico.

Se encuentran trabajando en sus tesinas de maestría con mi dirección: Matías Banchoff y Francisco Lizarralde (co-director).

Se encuentran trabajando en sus tesinas de grado con mi dirección: Matías Celasco, Juan Ignacio Yañez, Santiago Pravisani, Agustín Marisi y Ezequiel Claramunt.

13. PARTICIPACION EN REUNIONES CIENTIFICAS. *Indicar la denominación, lugar y fecha de realización, tipo de participación que le cupo, títulos de los trabajos o comunicaciones presentadas y autores de los mismos.*

14. CURSOS DE PERFECCIONAMIENTO, VIAJES DE ESTUDIO, ETC. *Señalar características del curso o motivo del viaje, período, instituciones visitadas, etc.*

15. SUBSIDIOS RECIBIDOS EN EL PERIODO. *Indicar institución otorgante, fines de los mismos y montos recibidos.*

Recibí un subsidio personal de la CIC, por \$8750, el cual fue completamente ejecutado y rendido.

16. OTRAS FUENTES DE FINANCIAMIENTO. *Describir la naturaleza de los contratos con empresas y/o organismos públicos.*

Se hizo una presentación al programa H2020, para acciones MC-RISE el cual fue adjudicado para iniciar en Octubre 2016.

17. DISTINCIONES O PREMIOS OBTENIDOS EN EL PERIODO.

18. ACTUACION EN ORGANISMOS DE PLANEAMIENTO, PROMOCION O EJECUCION CIENTIFICA Y TECNOLÓGICA. *Indicar las principales gestiones realizadas durante el período y porcentaje aproximado de su tiempo que ha utilizado.*

19. TAREAS DOCENTES DESARROLLADAS EN EL PERIODO. *Indicar el porcentaje aproximado de su tiempo que le han demandado.*

El tiempo dedicado a tareas docentes es de aproximadamente el 30%.

Objetos II, Materia de Grado de Licenciatura en Informática, UNLP
Groupware y Software Social, Materia del Doctorado en Ciencias Informáticas de la UNLP.
Dirección de trabajos finales y tesis.

20. OTROS ELEMENTOS DE JUICIO NO CONTEMPLADOS EN LOS TÍTULOS ANTERIORES. *Bajo este punto se indicará todo lo que se considere de interés para la evaluación de la tarea cumplida en el período.*

21. TÍTULO Y PLAN DE TRABAJO A REALIZAR EN EL PROXIMO PERIODO. *Desarrollar en no más de 3 páginas. Si corresponde, explicita la importancia de sus trabajos con relación a los intereses de la Provincia.*

Título: Inteligencia Colectiva contra el Riesgo y la Incertidumbre en el Sector Agrícola

La incertidumbre (en relación a aspectos como el medio-ambiente, los recursos, la economía, los mercados) es una causa frecuente de asignación ineficiente de recursos y de pérdidas económicas a lo largo de todos los eslabones de la cadena de valor. Si bien este es un tema estudiado desde hace muchos años, nuevas formas de incertidumbre originadas en el cambio climático, la globalización, la preocupación por el cuidado de los recursos naturales y el crecimiento constante de una demanda por productos más saludables, no han sido estudiadas en profundidad en la literatura de apoyo a las decisiones.

La disponibilidad de datos adecuados es de vital importancia para el manejo del riesgo en agricultura. Dado que los procesos biológicos y económicos son en gran medida aleatorios, la administración de riesgos agropecuarios se basa mayormente en métodos probabilísticos. En este sentido, la disponibilidad de datos fiables es crítica para la estimación precisa de la probabilidad de ocurrencia de los distintos fenómenos de interés (ej: niveles de rendimientos y precios, entre otros).

La gestión del conocimiento para la toma de decisión en la agricultura requiere de datos de distintos dominios y con una especificación muy detallada. La complejidad de problema al que nos enfrentamos se debe a la insuficiencia estructural de los datos. Focalizándonos en la gestión de metadatos de alta calidad, la iniciativa de "Linked Open Data (LOD)" es la clave para enfrentar este problema. Con el objeto de reducir el gap entre los proveedores y los consumidores de datos en el área agrícola es necesario: mejorar la disponibilidad de los datos, su estandarización y posibilidad de ser

descubiertos; ser útiles para los tomadores de decisiones y los procesos de toma de decisión, permitiendo hacer procesamiento y análisis de datos de la forma más automática posible; establecer LOD como plataforma para proveer y consumir datos masivos, definir, adoptar ontologías como estándares para la interoperabilidad semántica, y definir metodologías que permitan evaluar la calidad de los datos desde distintas dimensiones de la calidad semántica: completitud, exactitud, actualidad, etc. Inteligencia colectiva se refiere a la capacidad de grupos interconectados de personas y computadoras, que actúan de manera colectiva en formas que parecen inteligentes.

En ese contexto, me resultan particularmente interesante las iniciativas de ciencia ciudadana, en las que el público general se involucra en actividades de investigación asistiendo en la colección de datos, resolución de tareas, y aportando a la formulación de preguntas de investigación relevantes para la sociedad. El potencial de este tipo de iniciativas ha quedado demostrado con proyectos como FoldIt que ha hecho importantes descubrimientos en el plegado de proteínas. En nuestra región las iniciativas de ciencia ciudadana se encuentran en un estado incipiente destacándose, por ejemplo, el proyecto Chile Científico, que implementa proyectos enfocados en la conservación de especies y de captura de datos hidrológicos. La ciencia ciudadana es, en si misma, un desafío científico que postula interrogantes en relación a la formulación de proyectos, a la motivación de los participantes, en la evaluación de los resultados, en relación a las tecnologías que la hacen posible y en cuanto a aspectos éticos.

El foco de mi trabajo para el próximo período será investigar la aplicabilidad de estrategias de inteligencia colectiva para resolver los desafíos de la incertidumbre y el riesgo en la agricultura, enfocando tareas como la obtención de información, la evaluación de calidad de las fuentes de datos, y la generación de alertas tempranas.

Condiciones de la presentación:

- A. El Informe Científico deberá presentarse dentro de una carpeta, con la documentación abrochada y en cuyo rótulo figure el Apellido y Nombre del Investigador, la que deberá incluir:
 - a. Una copia en papel A-4 (puntos 1 al 21).
 - b. Las copias de publicaciones y toda otra documentación respaldatoria, en otra carpeta o caja, en cuyo rótulo se consignará el apellido y nombres del investigador y la leyenda "Informe Científico Período".
 - c. Informe del Director de tareas (en los casos que corresponda), en sobre cerrado.
- B. Envío por correo electrónico:
 - a. Se deberá remitir por correo electrónico a la siguiente dirección: infinvest@cic.gba.gov.ar (puntos 1 al 21), en formato .doc zipeado, configurado para papel A-4 y libre de virus.

- b. En el mismo correo electrónico referido en el punto a), se deberá incluir como un segundo documento un currículum resumido (no más de dos páginas A4), consignando apellido y nombres, disciplina de investigación, trabajos publicados en el período informado (con las direcciones de Internet de las respectivas revistas) y un resumen del proyecto de investigación en no más de 250 palabras, incluyendo palabras clave.

Nota: El Investigador que desee ser considerado a los fines de una promoción, deberá solicitarlo en el formulario correspondiente, en los períodos que se establezcan en los cronogramas anuales.