

# **CURRICULUM VITAE**

del *Dr. Grigori SIDOROV*



**SNI 3**

**Miembro de la Academia Mexicana de Ciencias**

**Publicaciones: 149**

**ISI-JCR: 20**

**Libros: 7**

**Citas a sus publicaciones: 159**

**Patentes: 1 (más 1 en proceso)**

**Proyectos CONACYT: 4**

**Editor-en-jefe, asociado o invitado de revistas: 5**

# Índice

1.	Condición actual .....	2
2.	Idiomas hablados.....	3
3.	Escolaridad .....	¡Error! Marcador no definido.
4.	Las citas profesionales durante el periodo anterior .....	¡Error! Marcador no definido.
5.	SNI (Sistema nacional de investigadores de México) y Academia de Ciencias.....	4
6.	Experiencia de trabajo lingüístico en el campo.....	4
7.	Desarrollo de las herramientas lingüísticas .....	4
8.	Habilidades y conocimiento profesional profundo.....	4
9.	Interés actual en la investigación.....	5
10.	Director de proyectos de investigación .....	5
	Proyectos CONACyT (apoyo de gobierno de México) .....	5
	Proyectos científicos aplicados .....	5
	Proyectos institucionales (IPN) .....	5
11.	Asistencia a congresos .....	6
12.	Eventos de divulgación y difusión .....	6
13.	Cursos impartidos .....	7
14.	Distinciones .....	8
15.	Asociaciones profesionales.....	8
16.	Puestos en asociaciones profesionales .....	8
17.	Tesistas .....	8
18.	Coordinador de alumnos PIFI (programa institucional de fomento de investigadores) .....	9
19.	Sinodal en exámenes de grado o predoctoral.....	10
20.	Formación de grupos de investigación .....	11
1.	Preparación del libro de texto fundamental en el área de la Inteligencia Artificial "Métodos modernos de Inteligencia Artificial" .....	11
2.	Programa de investigación apoyado por el IPN, reg. 1008 (2009) "Herramientas para el desarrollo de robots móviles: interfaces en lenguaje natural y orientación espacial" .....	11
3.	Creación y utilización de los textos paralelos.....	12
4.	Búsqueda automática de las primitivas semánticas en el sistema léxico del español .....	12
5.	Parser sintáctico para el español .....	12
6.	Analizador morfológico automático para el español.....	12
7.	Análisis semántico comparativo del campo semántico "no-adulto" en español, inglés y ruso. ....	12
8.	Preparación de libros con los problemas lingüísticos y organización de las olimpiadas en lingüística. ....	13
21.	Organización de eventos, árbitro de revistas, miembro del comité del programa .....	13
22.	Publicaciones.....	14
23.	Referencias a mis trabajos .....	14

## 1. Condición actual

Profesor e Investigador Titular "C" (el puesto máximo de un investigador en el IPN),  
Laboratorio de Lenguaje Natural y Procesamiento de Texto,  
Centro de Investigación en Computación del Instituto Politécnico Nacional (CIC-IPN),  
Av. Juan de Dios Bátiz, s/n, esq. Othón de Mendizabal, Nueva Industrial Vallejo, C.P. 07738, México D. F., México.  
E-mail: [sidorov@cic.ipn.mx](mailto:sidorov@cic.ipn.mx)

## 2. Puestos académico-administrativos

- Coordinador del programa de doctorado en ciencias de la computación del CIC-IPN (agosto 2006-agosto 2009). El programa pertenece al PNPC (padrón de excelencia) de Conacyt. El programa tiene alrededor de 60 estudiantes en promedio por semestre (en total) y 35 profesores. El puesto no es remunerado, es un cargo por elección de los profesores del programa. Se trata de liderazgo académico en el doctorado, de definir las metas del programa a largo plazo.

- Miembro del Colegio Académico de Posgrado del IPN (agosto 2008-agosto 2009). Es un órgano académico que resuelve los asuntos académicos de posgrado del todo IPN (más de 100,000 estudiantes, más de 10,000 profesores).
- Tesorero de la Mesa Directiva de la Sociedad Mexicana de la Inteligencia Artificial (SMIA), 2008 a la fecha.
- Secretario de la Mesa Directiva de la Asociación Mexicana del Procesamiento de Lenguaje Natural (AMPLN), 2004–2009.
- Presidente del Comité de Programa, Mexican International Conference on Artificial Intelligence (MICA) 2010. Publicación de LNCS Springer-Verlag, Alemania. Este congreso es el congreso principal en México sobre la Inteligencia Artificial y uno tradicionalmente alrededor de 200 autores y 600 asistentes. Tasa de aceptación menor de 26%. Está en ISI Proceedings.
- Editor-en-jefe de la revista de investigación y desarrollo tecnológico en computación “Polibits” editada en el CIDETEC-IPN, enero 2008 hasta la fecha, ISSN 1870-9044.
- Editor Asociado, Revista “International Journal of Advanced Intelligence” ISSN 1883-3918, Japan, <http://aiia-i.com/ijai/>
- Editor Asociado, Revista “The Journal of China Universities of Posts and telecommunications” ISSN 1005-8885, Elsevier, China, [www.buptjournal.cn/xben/dqml.asp](http://www.buptjournal.cn/xben/dqml.asp)

### 3. Idiomas hablados

- Ruso (nativo, 100%),
- Español (100%),
- Inglés (100%),
- Japonés (lectura),
- Chino (lectura).
- Idiomas “antiguas” (lectura): latín, griego antiguo, eslavo antiguo, ruso antiguo.

### 4. Grados académicos y cursos tomados

- *Doctorado*: Doctor (Ph. D.) en Ciencias Filológicas (Universidad Estatal “Lomonosov” de Moscú, Rusia), 28/02/1996, tesis en la lingüística computacional. Título: “Diseño e implementación de los modelos, algoritmos y datos lingüísticos para los sistemas con análisis y generación morfológicos, para el ruso”.
- *Maestría*: Universidad Estatal “Lomonosov” de Moscú (Rusia), Facultad de Filología, Departamento de Lingüística Estructural y Aplicada, junio 1988. (Tesis de maestría “Análisis tipológico de la semántica de los modificadores de verbos”).
- *Licenciatura* Universidad Estatal “Lomonosov” de Moscú (Rusia), Facultad de Filología, Departamento de Lingüística Estructural y Aplicada, junio 1986

Cursos de actualización:

- Curso del Dr. Igor Melchuk (Montreal) “La morfología en el modelo lingüístico Sentido-Texto”, octubre 2005, UNAM, posgrado de lingüística, 20 horas.
- Curso, CIC, “Advanced topics in Distributed computing: From clusters to Grids”, 24 horas, 8-11 de mayo del 2007, impartido por Isaac Scherson (Universidad de California) y Peter Schulthess (Alemania)

### 5. Experiencia profesional

- 1988/09-1991/03 Centro de producción de sistemas de semiconductores (Moscú), ingeniero de software. Desarrollo de los sistemas computacionales usando técnicas de inteligencia artificial.
- 1991/04-1996/07 Centro de Tecnologías Nuevas (Moscú), investigador. Desarrollo de los sistemas computacionales usando técnicas de inteligencia artificial.
- 1996/08-1998/01 Instituto de Investigación en nanotecnologías de Fondo Internacional para Conversión (Moscú), investigador. Desarrollo de los sistemas computacionales usando técnicas de inteligencia artificial.
- 1992/08-1998/07 Universidad Estatal Lingüística de Moscú, Departamento de Semántica Lingüística (tiempo parcial), profesor-investigador. Clases en la lingüística computacional, semántica, traducción automática, programación en la lingüística computacional, entre otras; investigación en lingüística computacional.
- 1998/03-1998/07 Compañía de consultoría económica (Project Analysis Consulting Company (PACC)) (Moscú), programador. Desarrollo de sistemas computacionales.

## 6. SNI (Sistema nacional de investigadores de México) y Academia de Ciencias

- SNI, Nivel 3 (nivel máximo) (2010-20014),
- SNI, Nivel 2 (2006-2009),
- SNI, Nivel 1 (2002-2005),
- SNI, Nivel C (1999-2002).

Miembro regular de la Academia Mexicana de Ciencias desde 2006.

## 7. Experiencia de trabajo lingüístico en el campo

En las expediciones lingüísticas de la facultad de filología de la Universidad Estatal "Lomonosov" de Moscú.

- 1984, julio-agosto, duración 6 semanas, lengua: dungano, población Irdyk, Kirgizstan.
- 1985, julio-agosto, duración 6 semanas, lengua: dungano, población Alexandrovka, Kirgizstan.
- 1986, julio-agosto, duración 6 semanas, lengua: dungano, población Alexandrovka, Kirgizstan.

## 8. Desarrollo de las herramientas lingüísticas

- Analizador morfológico automático para el ruso (1994-1999).
- Sistema de acentuación automática de palabras para el ruso (1996-1998).
- Sistema para generar concordancias con la posibilidad de usar las características morfológicas para el ruso (1996-1998).
- Analizador morfológico automático para el español AGME (2002-2006), (con Dr. A. Gelbukh, M.C. F. Velásquez, M.C Liliana Chanona). Registro de derechos del programa registrados.
- Analizador sintáctico automático (1999-2004) (con Dr. A. Gelbukh, Dr. I. Bolshakov, Dra S. Galicia).
- Clasificador de textos basado en el tesoro (1999-2005) (con Dr. A. Gelbukh).
- Sistema para marcación semiautomática de sentidos de palabras (2002-2003), (con M.C. Y. Ledo, Dr. A. Gelbukh).
- Herramienta de lexicógrafo con verificación de los círculos viciosos, vocabulario definidor, etc. (en preparación) (2003-2004) (con Dr. A. Gelbukh, M.C. W. Olan).
- Herramienta de procesamiento semiautomático de textos en español (2005-2006) (con Raúl Ávila).
- Herramienta de marcación de textos paralelos (con A. Gelbukh) (2006).
- Herramienta de división de palabras en morfemas con algoritmo genético (con D. Lara) (2008).
- Interfaz de búsqueda de los glifos maya (con O. Pichardo) (2008).
- Herramienta de búsqueda inteligente en un diccionario zapoteco-español (junto con Thomas Smith Stark y A. Gelbukh) (2008).

## 9. Habilidades y conocimiento profesional profundo

Las áreas científicas:

- Procesamiento automático de lenguaje natural,
- Bases de datos lingüísticas, diccionarios y tesauros,
- Morfología, sintaxis, semántica de lenguajes naturales.

Las áreas de aplicación:

- Procesamiento automático de textos (clasificación, generación de resúmenes, etc.)
- Automatización de oficinas,
- Interfaces en lenguajes naturales para bases de datos,
- Sistemas automatizados de lenguajes naturales, para la extracción de información objetiva desde textos,
- Recursos lingüísticos grandes (bases de datos, tesauros) para el aprendizaje de lenguaje y revisión de la gramática,
- Programación aplicada y sistémica.

Las herramientas de programación:

- Los lenguajes de programación: Pascal (con Delphi), C, C++, Java, Ensamblador, SQL.
- Los sistemas operativos: MS DOS, Windows 3.1x, 95, NT, 98, 2000, XP, Vista.
- DBMS: dBASE, Paradox, MS-SQL, InterBase.

## 10. Interés actual en la investigación

Procesamiento automático de textos y diccionarios grandes, tesauros, procesamiento de lenguajes naturales, programación, análisis automático sintáctico, semántico y morfológico.

## 11. Director de proyectos de investigación

### *Proyectos CONACyT (apoyo de gobierno de México)*

- Responsable del proyecto I31989-A: “La extracción automática de información semántica de los diccionarios bilingües y explicativos de español” (1999-2000)
- Responsable del proyecto 38435-A “Construcción automática de la base del sistema léxico del español a través del diccionario explicativo y su aplicación a las tareas de recuperación de información” (2002-2004)
- Responsable del proyecto 50206-H “Desarrollo del corpus paralelo alineado español-inglés usando desambiguación automática global con diccionarios bilingües” (2006-2007). (Solicitante y director por 1,5 años Dr. Grigori Sidorov, después se hizo el cambio del director del proyecto debido a la necesidad que yo sea el responsable del otro proyecto de una red de investigación).
- Responsable del proyecto 83270 “Análisis semiautomático comparativo de patrones de manejo sintáctico en español e inglés y su aplicación en el analizador sintáctico para el español” (2009-2011).

### *Proyectos científicos aplicados*

- Desarrollador principal. Sistema de análisis semiautomático de los documentos legales (para el Abogado General del IPN) (1999)
- Desarrollador principal. Sistema de búsqueda inteligente de documentos en la base de datos del **Senado de la República Mexicana**. (1999-2000)

### *Proyectos institucionales (IPN)*

- Responsable del proyecto SIP-IPN, 20100668 “Extracción automática de patrones sintácticos”
- Responsable del proyecto SIP-IPN, 20090772 “Desarrollo de gramáticas formales para reconocimiento de texto y voz” (2009)
- Responsable del proyecto SIP-IPN, 20080787 “Representación estructural de datos visuales y búsqueda inteligente en diccionarios de idiomas con escritura ideográfica” (2008)
- Responsable del proyecto SIP-IPN, 20060735 y 20071579 “Métodos de aprendizaje automático para detección de estructura morfémica de palabras en español”, (2006 y 2007)
- Responsable del proyecto CGEPI-IPN 20050459 “Desarrollo del sistema para alineación semiautomática de textos paralelos español – inglés y su aplicación para creación de un corpus paralelo” (2005).
- Responsable del proyecto CGEPI-IPN 20040374 “El método y la herramienta para compilación automática de los diccionarios grandes de colocaciones para el español” (2004-2005).
- Responsable del proyecto CGEPI-IPN 20031322 “Desarrollo del método mejorado de desambiguación de sentidos de palabras basado en el diccionario de colocaciones y en el reagrupamiento automático de sentidos de palabras en el diccionario explicativo” (2003-2004).
- Responsable del proyecto CGEPI-IPN 20020235 “Desambiguación de sentidos de palabras en recuperación de información” (2002-2003).
- Responsable del proyecto CGEPI-IPN 20010769 “Sistema de resolución de ambigüedades sintácticas para el español, basado en un diccionario de colocaciones” (2001-2002).
- Responsable del proyecto CGEPI-IPN 200107 “Compilación Automática del Diccionario de las Primitivas Semánticas a través de Extracción de los Diccionarios Bilingües y Explicativos de Español y su Aplicación para la Resolución de Ambigüedad” (2000-2001).
- Responsable del proyecto CGEPI-IPN 990078 “Sistema inteligente para la clasificación de documentos basado en un diccionario de español con una jerarquía ponderada de temas” (1999-2000).

## 12. Participante de los proyectos de investigación

Varios. Alrededor de 30.

### 13. Asistencia a congresos

1. Congreso Internacional en Inteligencia Artificial, COMIA Tlaxcala, 13 de octubre 2010, Plática magistral "Algoritmos genéticos en PLN".
2. Seminario en la Universidad Politécnica de Valencia (Departamento de sistemas computacionales), "Dos casos de aplicación de algoritmos genéticos en las tareas de PLN: división de palabras en morfemas y detección automática de vocabulario definidor", 2010.
3. 5 taller de Procesamiento de lenguaje natural. Puebla, 2-4 de octubre del 2008 (Representación estructural de glifos para búsqueda en un diccionario de glifos maya).
4. Congreso Internacional EURALEX (lexicografía), Barcelona, España, 2008
5. 4 seminario de ingeniería lingüística, UNAM, 2008 (Diseño y construcción de un sistema inteligente de recuperación de textos políticos de la base de datos documental del Senado de la República Mexicana).
6. 4o Taller de Procesamiento de Lenguaje Natural, Puebla, 2007 (Bases para la alineación automática de textos paralelos a nivel de palabras utilizando información lingüística diversa).
7. Orizaba, 2007, 14th International Congress on Computer Science Research CIICC'07.
8. Dialog 2007, Computational linguistics and its applications, Moscú, Rusia, 2007
9. IX encuentro internacional de lingüística en el noroeste, México, 2006.
10. 7o encuentro internacional de computación ENC-2006, México, 2006
11. Dialog 2006, Computational linguistics and its applications, Moscú, Rusia, 2006
12. 6o encuentro internacional de computación ENC-2005, México, 2005
13. 4 Congreso mexicano internacional de inteligencia artificial, 2005.
14. Congreso internacional de computación Iberamia 2004, México, 2004.
15. VIII encuentro internacional de lingüística en el noroeste, México, 2004.
16. 5o encuentro internacional de computación ENC-2004, México, 2004.
17. Dialogue-2004, Computational linguistics and its applications, Moscú, Rusia, 2004.
18. "International conference on e-learning, e-education, e-business" L'Aquila, Italia, 2003.
19. Telemática-2002. Havana, Cuba, 2002.
20. ACH/ALLC International Conference (Association of computers and humanities / Association for Literary and Linguistic Computing) (Tübingen, Germany, 2002).
21. ICEIS, International conference (International Conference on Enterprise Information Systems, workshop NDDL - New developments in Digital Libraries) (Setúbal, Portugal, 2001).
22. COMPLEX-2001, Computational lexicography, (Birmingham, UK, 2001).
23. TALN-2001, Tratamiento automático de lenguaje natural (Tours, Francia, 2001).
24. 7th International Pragmatic Conference (Budapest, Hungary, 2000).
25. International conference CILing2000 (Computational Linguistics and Intelligent Text processing) (Mexico, 2000).
26. Congreso Internacional en Computación, CIC-2000 (México, 2000).
27. MICAI-2000, Mexican International Conference on Artificial Intelligence, (Mexico, 2000).
28. IJCAI-99, International Joint Conference on Artificial Intelligence (Stockholm, Sweden, 1999).
29. 2nd International Conference on Cognitive Science and 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99) (Tokyo, Japan, 1999).
30. Congreso Internacional en Computación, CIC-99, (México, 1999).
31. 2º Encuentro Nacional de Computación, ENC-99 (México, 1999).
32. VI Congreso internacional: Knowledge-Dialogue-Solution (KDS-97) (Yalta, Ucrania, 1997).
33. Congreso internacional: Lingüística funcional: Principios de formación de texto. (Symferopol, Ucrania, 1996).
34. II Simposium Internacional de AIMLIR (Asociación internacional de maestros de literatura e idioma rusos), Fonética en el sistema de lenguaje. (Moscú, Rusia, 1996)
35. Multimedia en enseñanza de idiomas. (Moscú, Rusia, 1995).
36. Sistemas y Dispositivos Optoelectrónicos. (Moscú, Rusia, 1990).
37. 5º Congreso de los orientistas jóvenes de toda la Unión Soviética (Rusia, Moscú, 1989).

### 14. Eventos de divulgación y difusión

1. Tutorial "Fundamentals of Information Retrieval", INAOE, 8 de octubre 2010.
2. Tutorial "Procesamiento de textos en el español", COMIA 2010, 12 de septiembre 2010.
3. Congreso Internacional en Computación, Apizaco, 28 de octubre 2010, Plática magistral "Aplicaciones de procesamiento de lenguaje natural".

4. Tutorial "Procesamiento automático de lenguaje natural", Congreso Internacional en Computación, Apizaco, 26 de octubre 2010.
5. Seminario de investigación de Sección de Investigación y estudios de posgrado, ESCOM, 2010, "Aplicaciones de procesamiento de lenguaje natural", agosto 2010.
6. CORE 2010, Plática magistral "Aplicaciones de procesamiento de lenguaje natural" mayo 2010.
7. Universidad Autónoma de Estado de México, Pachuca, "Procesamiento automático de lenguaje natural y sus aplicaciones", sept. 2009.
8. Centro Universitario Ixtlahuaca, Universidad Autónoma de Estado de México, "Procesamiento automático de lenguaje natural y sus aplicaciones" (plática invitada), 2008.
9. UAEM, facultad de ingeniería, "Oferta educativa en el CIC", 2008.
10. Escuela Primaria "Roman Saldaña Oropesa", Muestra de Robótica para alumnos, enero 2009
11. Escuela Primaria Club de los Leones 4, "Interacción con los robots" (Concurso de robótica para niños), 2008.
12. Escuela Secundaria 105, "Interacción con los robots" (Concurso de robótica para niños), 2008
13. Semana de Ingeniería (plática invitada), Universidad Autónoma de Estado de México, campus Tianguistenco, 2008. "Lingüística computacional y sus aplicaciones"
14. Congreso Neurohumanidades, U. Iberoamericana, conferencia magistral, "Análisis computacional de lenguaje", 2007.
15. Seminario de Investigación, plática invitada, IPN, "Métodos estadísticos en desciframiento de escrituras antiguas" 2007.
16. 3 Coloquio de la lingüística computacional, conferencia magistral, UNAM, "Corpus paralelo alineado español-inglés de los textos literarios", 2007.
17. Mesa redonda de lingüística computacional "Lingüística, más allá de ciencias exactas y naturales", conferencia magistral, UNAM, México, "Métodos computacionales en la investigación lingüística", 2006
18. Plática invitada, 2do encuentro de investigación de la UVM, México, Universidad del Valle de México, "Procesamiento automático de lenguaje natural", 2006
19. UNAM, Facultad de ciencias, México, seminario de investigación, plática invitada, "Procesamiento automático de lenguaje natural y sus aplicaciones", 2005.
20. Colegio de Bachilleres, México, XV semana de investigación científica, conferencia magistral, 2004.
21. Universidad de las Américas, México, X semana académica, conferencia magistral, 2002.
22. Semana cultural, Universidad de Chilpancingo, México, conferencia magistral, 2001.
23. Seminario de Investigación, CIC, 2001.
24. 1er Congreso Internacional de Ingeniería Electromecánica, Instituto Tecnológico de San Andrés Tuxtla, Veracruz, conferencia magistral, 2000.

## 15. Cursos impartidos

1. Lingüística teórica y computacional, Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1992-1998 (7 semestres). A92-A98, licenciatura, 74 horas por semestre.
2. Fundamentos de traducción automática, Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1994, A94. licenciatura, 108 horas.
3. Tecnologías modernas de información. Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1994, A94, licenciatura, 18 h.
4. Programación para tareas lingüísticas, Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1993-1998 (6 semestres). B93-B98, licenciatura, 32 horas por semestre.
5. Programación para tareas lingüísticas, Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1994-1998 (5 semestres). A94-A98, maestría, 38 horas por semestre.
6. Modelos y algoritmos lingüísticos para los sistemas automatizados. Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1993-1998 (6 semestres). licenciatura, A93-A98, 18 horas por semestre.
7. Modelos lingüísticos aplicados. Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1996, B96, maestría, 36 h
8. Semántica aplicada, Universidad Estatal Lingüística de Moscú, Departamento de semántica lingüística, 1998, A98, licenciatura, 36 h.
9. Tratamiento de lenguaje natural, CIC, maestría, 80 horas, B04, B07.
10. Lingüística aplicada I (Introducción a lingüística aplicada), CIC, maestría, 80 horas, B99, B03, A05, A06, B06.
11. Lingüística aplicada II (Sistemas de aplicación lingüística), CIC, doctorado, 80 horas, A03, A04, A07, B08.
12. Lingüística aplicada II (Sistemas de aplicación lingüística), CIC, maestría, 80 horas, A00, A02, B03, A04, B04.
13. Lingüística aplicada III (Procesamiento inteligente de textos), CIC, doctorado, 80 horas, A01, B01.

14. Lingüística aplicada III (Procesamiento inteligente de textos), CIC, maestría, 80 horas, B02, B05.
15. Tecnologías de Lenguaje Natural, CIC, maestría, 80 horas, A04, A06, A07.
16. Temas selectos de computación I. CIC, Doctorado, 80 horas, A08.
17. Temas selectos de computación II. CIC, Doctorado, 80 horas, A08.
18. Seminario I (Lingüística computacional), CIC, doctorado, 60 horas, B04, A06, B06, B08.
19. Seminario I (Lingüística computacional), CIC, maestría, 60 horas, A01, A03, A04, A07.
20. Seminario II (Lingüística computacional), CIC, doctorado, 60 horas, B01, B06, B07.
21. Seminario II (Lingüística computacional), CIC, maestría, 60 horas, A02, B02, B03, A05, B05, B08.
22. Seminario III (Lingüística computacional), CIC, doctorado, 60 horas, A04, B05, A06, B07, A08.
23. Seminario III (Lingüística computacional), CIC, maestría, 60 horas, A02, A03, A04, A07.
24. Tema de tesis, CIC, doctorado, 60 horas, A04, B04, A05.
25. Tema de tesis, CIC, maestría, 60 horas, B03, A04, B04, A05.

## 16. Distinciones

1. **Miembro regular de la Academia Mexicana de Ciencias** en el área de ciencias sociales y humanidades, desde 2006.
2. **Editor-en-jefe** de la revista "Polibits" editada en el CIDETEC-IPN, desde enero del 2008 hasta la fecha, ISSN 1870-9044.
3. **Miembro del sistema nacional de investigadores (SNI)** de México, nivel 3.
4. Editor invitado, revista "Informatica. International Journal of Computing and Informatics" (Slovenia, ISSN 0350-5596) <http://www.informatica.si/>. Número especial en lingüística computacional, 2009.
5. Editor Asociado, Revista "International Journal of Advanced Intelligence" ISSN 1883-3918, Japan, <http://aia-i.com/ijai/>
6. Editor Asociado, Revista "The Journal of China Universities of Posts and telecommunications" ISSN 1005-8885, Elsevier, China, [www.buptjournal.cn/xben/dqml.asp](http://www.buptjournal.cn/xben/dqml.asp)
7. Publicación de mis datos en Marquis "Who's Who in the World", 2006, 2007, 2008, 2009 (Inclusión en este catálogo no tiene costo ni condición alguna y es efectuada por el Comité Científico en base del único criterio del reconocimiento internacional e impacto internacional de las actividades profesionales del galardonado).
8. Publicación de mis datos en Marquis "Who's Who in Science and Engineering", 2006, 2007, 2008, 2009 (Inclusión en este catálogo no tiene costo ni condición alguna y es efectuada por el Comité Científico en base del único criterio del reconocimiento internacional e impacto internacional de las actividades profesionales del galardonado).
9. Reconocimiento del Abogado General del Instituto Politécnico Nacional (IPN) y director del CIC por desarrollo del programa de búsqueda inteligente jurídica para el Reglamento Interno del IPN, 1998.
10. Reconocimiento del Instituto Politécnico Nacional por destacada actividad de investigación 2000.
11. Presea Lázaro Cárdenas 2003 en la categoría de mejor alumno de doctorado del IPN que obtuvo el alumno de quien soy el director (premio anual que entrega personalmente el Presidente de la República Mexicana y después tiene una plática informal con los galardonados).

## 17. Asociaciones profesionales

Asociación Mexicana de Procesamiento de Lenguaje Natural 2001 – hasta la fecha.

Asociación de Aplicación de Metáfora, Gran Bretaña, 2007, 2008.

Sociedad mexicana de inteligencia artificial, desde 2008 - hasta la fecha.

## 18. Puestos en asociaciones profesionales

- Secretario de Asociación Mexicana de Procesamiento de Lenguaje Natural, 2004-...
- Tesorero de Sociedad Mexicana de Inteligencia Artificial, enero de 2008-...
- Presidente de la Comisión de Doctorado del H. Colegio de Profesores del Centro de Investigación en Computación, IPN, 2004-2006.
- Coordinador del programa de Doctorado en ciencias del CIC (agosto 2006-agosto 2009)

## 19. Tesistas

1. Concluido. 12.07.1996. Ekaterina Bazova. "Development of linguistic models and data for the Russian language tutoring system for foreigners (semantic field "Elections and voting")." Moscow State Linguistic University. Master degree.

2. Concluido. 01.07.2003. Francisco A. Castillo Velásquez. Maestría. CIC. "Sistema de análisis morfológico para el español.", mención honorífica.
3. Concluido. 04.06.2003. Gabriela Rivera Loza. Maestría (codirector). CIC. "Selección automática de un vocabulario definidor para un diccionario semántico del idioma español "
4. Concluido. 19.08.2004. Erika Hernández Rubio. Maestría. CIC. "Compilación automática del diccionario de colocaciones para el español usando estructuras sintácticas."
5. Concluido. 16.12.2005. Noé Alejandro Castro Sánchez. Maestría. CIC. "Sistema para detección automática de los perfiles lingüísticos.", mención honorífica.
6. Concluido. 16.12.2005. José Angel Vera Félix. Maestría. CIC. "Sistema de Detección Semiautomática de Frasemas usando Análisis de Recursos Léxicos Bilingües".
7. Concluido. José Ernesto Gómez Balderas, Licenciatura, BUAP "Etiquetación de partes de oración con redes neuronales", 28/06/2006.
8. Concluido, 23/06/2006. Yoel Ledo Mezquita. Doctorado. CIC. "Recuperación de información con desambiguación de sentidos de palabras para el español." Presea Lázaro Cárdenas 2003 (premio anual al mejor alumno de doctorado del IPN que entrega personalmente el Presidente de la República Mexicana y después tiene una plática informal con los galardonados en todas las categorías).
9. Concluido, 23/06/2006. Omar A. Olivas Zazueta. Maestría. CIC. "Sistema de construcción de redes semánticas con detección de anáfora."
10. Concluido, 8/12/2006, Ma de los Ángeles Alonso Lavernia. Doctorado. CIC. "Representación y manejo de la información semántica heterogénea en interacción humano-máquina."
11. Concluido, 22/05/2007. José Ernesto Gómez Balderas. Maestría. CIC. "Resolución automática de la homonimia morfológica para el español".
12. Concluido. 18/06/2008. Diego Lara Reyes. Maestría. CIC. "Sistema de segmentación automática de palabras en morfemas para el español"
13. Concluido. 17/12/2008. Obdulia Pichardo Lagunas. Maestría. CIC. "Representación computacional de la escritura maya".
14. Concluido. 14/01/2009. Vicente Cubells Nonell. Doctorado. CIC. "Análisis sintáctico del español utilizando gramáticas de adjunción de árboles."
15. Concluido. 20/01/2010. Eduardo Moises Lavín Villa. Maestría, CIC. "Extracción automática de diccionarios semánticos de los textos".
16. Concluido 15/05/2010. Grettel Barceló Alonso. Doctorado. CIC. "Desambiguación de los sentidos de las palabras en español usando los textos paralelos".
17. Eduardo Antonio Cendejas Castro. Doctorado. CIC. "Alineación de palabras en textos paralelos en español usando información lingüística diversa". (en preparación, registro de tema).
18. Noé Alejandro Castro Sánchez. Doctorado. CIC. "Un método automático para la extracción de los patrones de reción en el español basado en diccionarios explicativos y relaciones léxicas". (en preparación, predoctoral).
19. Obdulia Pichardo Lagunas. Doctorado. CIC. "Detección automática de primitivos semánticos con algoritmos bio-inspirados" (en preparación, registro de tema).
20. Ismael Díaz Rangel. Doctorado. CIC. "" (en preparación, registro de tema).
21. Sabino Miranda Jiménez. Doctorado. CIC. "GENERACIÓN AUTOMÁTICA DE RESÚMENES ABSTRACTIVOS CON MÉTODOS SIMBÓLICOS" (en preparación, registro de tema).
22. Francisco Antonio Castillo Velásquez. Doctorado. CIC. "Un método de detección automática de autoría de textos basado en la información sintáctica" (en preparación, registro de tema).
23. Juve Andrea Herrera de la Cruz. Maestría, CIC. "SISTEMA DE EXTRACCIÓN AUTOMÁTICA DE INFORMACIÓN SEMÁNTICA DE LOS LIBROS DE TEXTO ESTRUCTURADOS" (en preparación).
24. Juan Pablo Posadas Durán. Maestría, CIC. "Sistema de alineación de textos paralelos a nivel de oraciones" (en preparación).
25. Analhy Ochoa Vargues. Maestría, CIC. "Sistema de detección automática de clases gramaticales de palabra en el español" (en preparación).
26. Carlos Ignacio Reséndiz Juárez. Maestría. CIC. "Traducción automática en el dominio médico". (en preparación)
27. Adalberto Robles Valádez. Maestría. CIC. "Identificación automática de modismos en el español." (en preparación).

## **20. Coordinador de alumnos PIFI (programa institucional de fomento de investigadores)**

1. Yulia Nikolaevna Ledeneva (2008)
2. Horacio Alberto García Salas (2008)

3. Javier Leandro Tejada Cárcamo (2008)
4. Analhy Ochoa Vargues (2007)
5. Diego Lara Reyes (2007)
6. José Ernesto Gómez Balderas (2006)
7. Vicente Cubells Nonell (2006)
8. Diego Lara Reyes (2006)
9. agosto 2005 – diciembre 2005 Omar Olivas Zazueta (0.5 año)
10. agosto 2005 – diciembre 2005 José Ernesto Gómez Balderas (0.5 año)
11. agosto 2004 – junio 2005 Omar Olivas Zazueta (1 año)
12. agosto 2002 – junio 2003 Francisco Antonio Castillo Velázquez (1 año)
13. agosto 2002 – junio 2003 Sabino Miranda Jiménez (1 año)
14. agosto 2002 – junio 2003 Yoel Ledo Mezquita (1 año)
15. enero 2002 – junio 2002 Francisco Antonio Castillo Velázquez (0.5 año)
16. agosto 2000 – junio 2001 Liliana Chanona Hernández (1 año)
17. agosto 2000 – junio 2001 Leticia Hernández Luque (1 año)
18. agosto 2000 – junio 2001 Jorge Alberto Virgen Montesinos (1 año)
19. agosto 1999 – junio 2000 Liliana Chanona Hernández (1 año)
20. agosto 1999 – junio 2000 Roberto A. Morales Pacheco (1 año)
21. agosto 1999 – junio 2000 Jorge Alberto Virgen Montesinos (1 año)

## 21. Sinodal en exámenes de grado o predoctoral

1. 07/1996. Ekaterina Bazova, Maestría, Universidad estatal Lingüística de Mocsú, Rusia. Development of linguistic support for the Russian language tutoring system (semantic field “Elections and voting”)
2. 07/2003. Francisco A. Castillo Velázquez. Maestría, CIC. Sistema de análisis morfológico para el español.
3. 06/2003 Gabriela Rivera Loza. Maestría, CIC. Selección automática de un vocabulario definidor para un diccionario semántico del idioma español.
4. 13.11.2003 Ma. Julia Calderón Sambarino. Maestría, CIC. Sistema para el estudio, medición y evaluación del conducto auditivo externo del oído mediante estímulos sonoros.
5. 26/02/2002. M. Montes y Gomez. Doctoral, CIC. Minería de texto empleando la semejanza entre estructuras semánticas
6. 03/2002 J. G. González Serna. Predoctoral, CIC. Esquema adaptivo para la gestión de movilidad en sistemas cliente/servidor a través de Internet.
7. 12/2002. Salvador Godoy Calderón. Predoctoral, CIC. Ambiente para evaluaciones de cubrimientos y algoritmos de clasificación, según el enfoque lógico-combinatorio del reconocimiento de patrones.
8. 07/2002. M. Reyes Alfonso. Predoctoral, CIC. Un modelo de componentes de software modelados por roles para aplicaciones distribuidas.
9. 04/2005. M. Reyes Alfonso. Doctoral, CIC. Un modelo de componentes de software modelados por roles para aplicaciones distribuidas.
10. 01/2004. Raúl Morales Carrasco. Doctoral, CIC. Resolución automática de anáfora indirecta en español
11. 08/2004. Erika Hernández Rubio. Maestría, CIC. Compilación automática del diccionario de colocaciones para el español usando estructuras sintácticas.
12. 14/12/2004, Edgar Moyotl Hernández, licenciatura, Puebla, BUAP.
13. 01/2005, Edgar Moyotl Hernández, maestría, Puebla, BUAP.
14. 04/2005, Sofía Paniagua Rivera, maestría, Puebla, BUAP.
15. 2006 Salvador Godoy Calderón. Doctoral, CIC. Ambiente para evaluaciones de cubrimientos y algoritmos de clasificación, según el enfoque lógico-combinatorio del reconocimiento de patrones.
16. 2005, Juan Javier Gonzalez Barbosa, PhD thesis (intermediate exam), CENIDET (external examiner).
17. 2005, Juan Javier Gonzalez Barbosa, PhD thesis (final exam), CENIDET (external examiner).
18. 2005 Hiram Calvo. Predoctoral, CIC.
19. 2006 Hiram Calvo. Doctoral, CIC.
20. 2006 J. G. González Serna. Doctoral. CIC. Esquema adaptivo para la gestión de movilidad en sistemas cliente/servidor a través de Internet.
21. 2006 Ricardo Vargas de Basterra. Predoctoral, CIC.
22. 2006 Ricardo Vargas de Basterra. Doctoral, CIC.
23. 2006, Francisco Rojas, Maestría, BUAP.
24. 2007. Predoctoral, CIC. Alejandro Canales Cruz

25. 2007. Doctoral, CIC. Alejandro Canales Cruz
26. 2007. Predoctoral, CIC. Juan Carlos Rojas, CENIDET, Cuernavaca.
27. 2007. Maestría, CIC. Benjamin Luna Benjoso.
28. 2007. Doctoral. Carlos Mario Zapata Jaramillo, Universidad Nacional de Colombia, Medellín, Colombia.
29. 2006 Predoctoral, Alma Delia Cuevas Rasgado. CIC.
30. 2006 Doctoral, Alma Delia Cuevas Rasgado. CIC.
31. 2007. Predoctoral. Jose Zárate Marcelleño, CENIDET, Cuernavaca.
32. 2007. Doctoral. Jose Zárate Marcelleño, CENIDET, Cuernavaca.
33. 2008. Doctoral. Universidad de Greenwich, UK.
34. 2008. Predoctoral. Miguel Felix Mata Rivera. CIC
35. 2009. Doctoral. Miguel Felix Mata Rivera. CIC
36. 2008. Predoctoral. Juan Carlos Rojas. CENIDET.
37. 2009. Doctoral. Juan Carlos Rojas. CENIDET.
38. 2009. Maestría. Ismael Rodriguez, CIC.
39. 2009. Maestría. Israel Toledo Alvarado, CIC.
40. 2008. Predoctoral. Yulia Ledeneva. CIC
41. 2008. Doctoral. Yulia Ledeneva. CIC
42. 2008. Predoctoral. Sulema Torres. CIC
43. 2008. Predoctoral. Grettel Barceló. CIC
44. 2008. Predoctoral. Eduardo Cendejas. CIC
45. 2008. Predoctoral. Javier Tejada. CIC
46. 2009. Doctoral. Javier Tejada. CIC
47. 2008. Maestría. Mayra Paredes. CIC
48. 2009. Maestría. Alfredo Lopez Monroy. CIC
49. 2009. Doctorado. Rafael Guzmán, UPV, España.

## 22. Formación de grupos de investigación

### 1. *Preparación del libro de texto fundamental en el área de la Inteligencia Artificial “Métodos modernos de Inteligencia Artificial”*

Participantes: Soy el único EDITOR de este libro. Participaron con contribuciones por capítulo varios investigadores de reconocido prestigio, mexicanos e extranjeros (total 15, de INAOE; Lania, IPN, IMP, U. Michoacana, U. Veracruzana, U. de Greenwich).

Objetivo: Desarrollo de un libro de texto que presente los conceptos básicos y los métodos modernos con ejemplos de uso de la Inteligencia Artificial.

Impacto: Se desarrolló el libro de texto fundamental que sirve como libro de texto y libro de consulta en el área de la Inteligencia Artificial, que hasta el momento no existía en el español. 345 páginas. Está en proceso de revisión por la casa editorial **Pearson**.

Permanencia: 2007-2009.

### 2. *Programa de investigación apoyado por el IPN, reg. 1008 (2009) “Herramientas para el desarrollo de robots móviles: interfaces en lenguaje natural y orientación espacial”*

Participantes: Soy el único COORDINADOR de este programa. Participaron varios investigadores del CIC. El programa une 4 proyectos de investigación (número máximo de un programa en el IPN): Dr. Grigori Sidorov, Dra Nareli Cruz Cortés, Dr. Sergio Suárez Guerra, Dr. José Luis Oropeza Rodríguez.

Objetivo: El programa une varios proyectos orientados a desarrollo de las herramientas para los robots móviles, incluyendo las herramientas para la interacción con los robots (análisis de texto de entrada y de reconocimiento y síntesis de voz) y orientación espacial de los robots. La meta principal es el desarrollo de las interfases en lenguaje natural con los robots móviles y determinación de la ubicación espacial de los robots.

Impacto: Se están desarrollando varias herramientas para la interacción con los robots (análisis de texto de entrada y de reconocimiento y síntesis de voz) y orientación espacial de los robots.

Permanencia: 2008-2009.

### **3. Creación y utilización de los textos paralelos**

(derivado de proyecto CONACyT 50206H, solicitante y director por 1,5 años Dr. Grigori Sidorov, después se hizo el cambio del director del proyecto debido a la necesidad de encabezar otro proyecto de la red de investigación)

Participantes: Dr. Grigori Sidorov, Dr. Alexander Gelbukh, Dr. Igor Bolshakov, M en C J. Ángel Vera Félix, M en C. Noé Alejandro Castro Sánchez, M en C. Grettel Barceló Alonso, M. en C. Eduardo Cendejas Castro (México, CIC-IPN)

Objetivo: Desarrollo de los métodos de construcción y utilización de los textos paralelos, principalmente español-inglés.

Impacto: Se desarrollaron varios métodos y herramientas de construcción de textos paralelos (alineación) y varios métodos de su utilización (desambiguación de los sentidos de palabras, extracción de patrones de manejo sintáctico), se publicó varios artículos en revistas nacionales e internacionales, se presentó una tesis de maestría, están en desarrollo tres tesis de doctorado.

Permanencia: 2004-2009

### **4. Búsqueda automática de las primitivas semánticas en el sistema léxico del español**

(derivado de proyecto CONACyT 38435A, director Dr. Grigori Sidorov)

Participantes: Dr. Grigori Sidorov, Dr. Alexander Gelbukh, Dr. Igor Bolshakov, Dra Nareli Cruz Cortés, M en C. Yoel Ledo Mezquita, M. en C. Vicente Cubells Nonell, M. en C. Gabriela Rivera Loza, M en C. Francisco A. Velázquez, Ing. Wilbert A. Olán Cristóbal, Ing. J. Ángel Vera Félix, Ing. Noé Alejandro Castro Sánchez, M. en C. Obdulía Pichardo Lagunas (México, CIC-IPN)

Objetivo: Desarrollo de los métodos de análisis automático de los diccionarios utilizando métodos de optimización.

Impacto: Se desarrollaron varios métodos y herramientas de análisis automático de los diccionarios, se publicó varios artículos en revistas nacionales e internacionales, se presentó una tesis de maestría, se desarrolla una tesis de doctorado.

Permanencia: 2001-2008

### **5. Parser sintáctico para el español**

(junto con Dr. Alexander Gelbukh)

Participantes: Dr. Grigori Sidorov, Dr. Alexander Gelbukh, Dr. Igor Bolshakov (México, CIC-IPN), Dra. Sofía Galicia Haro (UNAM), Dr. Vicente Cubells Nonell (ITESM).

Objetivo: Desarrollo de un parser sintáctico para el español basado en una gramática libre de contexto y en un diccionario de patrones de manejo sintáctico.

Impacto: Se desarrolló el parser sintáctico para el español; se publicó varios artículos, se presentó dos tesis doctorales.

Permanencia: 1999-2008

### **6. Analizador morfológico automático para el español**

Participantes: Dr. Grigori Sidorov, Dr. Alexander Gelbukh, Dr. Igor Bolshakov, M. en C. Francisco Velázquez, M. en C. Liliانا Chanona Hernández (México, CIC-IPN),

Objetivo: Desarrollo del analizador morfológico para el español. Yo fui el coordinador del grupo y director de la tesis de maestría "Sistema de análisis morfológico para el español" de M. en C. Francisco Velázquez. El sistema usa las ideas de mi tesis doctoral, desarrolladas primero para el idioma ruso y aplicadas al español.

Impacto: Se desarrolló el analizador morfológico para el español, el cual es una herramienta muy útil y se usa en todos los grupos que se dedican al procesamiento automático de lenguaje natural en México (UNAM, INAOE, IPN, UAM) y en algunos grupos en España (Valencia, Murcia, Barcelona); se publicó varios artículos en revistas nacionales e internacionales, se presentó una tesis de maestría.

Permanencia: 2002-2006

### **7. Análisis semántico comparativo del campo semántico "no-adulto" en español, inglés y ruso.**

Participantes: Dr. Grigori Sidorov, Dr. Alexander Gelbukh, Dra. Sofía Galicia Haro, Dr. Patrick Cassidy (EEUU), Dr. Igor Bolshakov (México, CIC-IPN)

Objetivo: Análisis semántico comparativo del campo semántico "no-adulto" en español, inglés y ruso.

Impacto: Se hizo análisis semántico comparativo del campo semántico "no-adulto" en español, inglés y ruso. Parte del español hizo Dra. Sofía Galicia Haro, parte de inglés hizo Dr. Patrick Cassidy (EEUU), parte de ruso hice yo en

su mayoría. También yo fui el responsable de hacer el análisis comparativo. Se publicó dos artículos en los congresos internacionales.

Permanencia: 1998-2000

### **8. Preparación de libros con los problemas lingüísticos y organización de las olimpiadas en lingüística.**

Participantes: Dr. Grigori Sidorov, Dr. Boris Gorodetsky, Dra. Tatiana Zevakhina, M. en C. Elena Kuz'mina, Dr. Pavel Parshin, Dra. Yulia Pirogova (Rusia, Universidad Estatal Lingüística de Moscú)

Objetivo: Se trata de algunos miembros del departamento de semántica lingüística, quienes se organizaron para preparación de los libros con los problemas lingüísticos, es decir, micromodelos en diferentes idiomas. Yo fui el coordinador del grupo.

Impacto: Se publicó dos libros con los problemas lingüísticos y se organizó dos olimpiadas de lingüística (basados en los libros) entre los estudiantes de preparatorias y secundarias de Moscú en 1993 y 1994.

Permanencia: 1993-1994

## **23. Organización de eventos, árbitro de revistas, miembro del comité del programa**

1. **Editor-en-jefe** de la revista "Polibits" editada en el CIDETEC-IPN.
2. **Program Chair**, Mexican International Conference on Artificial Intelligence (MICAI) 2010 (Springer LNCS, aceptación alrededor de 26%).
3. **Editor invitado**, revista "Informatica. International Journal of Computing and Informatics" (Slovenia), Número especial en lingüística computacional, 2009.
4. **Editor Asociado**, Revista "International Journal of Advanced Intelligence" ISSN 1883-3918, Japan, <http://aia-i.com/ijai/>
5. **Editor Asociado**, Revista "The Journal of China Universities of Posts and telecommunications" ISSN 1005-8885, Elsevier, China, [www.buptjournal.cn/xben/dqml.asp](http://www.buptjournal.cn/xben/dqml.asp)
6. Presidente del comité de programa de los talleres del Congreso Internacional MICAI-2008.
7. Presidente del comité de programa de los tutoriales del Congreso Internacional MICAI-2006.
8. Presidente del comité de programa del congreso CORE-2009 (México), congreso internacional en computación.
9. Presidente del comité de programa del congreso CORE-2008 (México), congreso internacional en computación.
10. Presidente del comité de programa del congreso CORE-2004 (México), congreso internacional en computación.
11. Evaluador de los proyectos Conacyt.
12. Co-organizador de la mesa especial "Herramientas computacionales para lingüistas" en el IX Congreso Internacional del Lingüística en Noroeste, Hermosillo, Sonora, 2006 (con Alfonso Medina, UNAM).
13. Árbitro de revistas "Computación y sistemas", "Advances in Artificial Intelligence" (IASTED), "Research in computing science".
14. Miembro de "LACCIR technical committee", 2007, 2008. (evaluación de proyectos de investigación a nivel de América Latina).
15. Miembro del comité de programa:
  - CICLing (2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009) (Congreso Internacional de lingüística computacional y procesamiento de texto), México,
  - CIC (2003, 2004, 2005, 2006) (Congreso Internacional de Computación), México
  - ENC, taller "Tecnologías de lenguaje humano" (2004, 2005, 2006) (Congreso internacional de Computación), México,
  - NLDB (Natural Language and Data bases) (2008 en Inglaterra, 2007 en Francia, 2009 en Alemania)
  - MICAI (2005, 2006, 2007, 2008) (Mexican International Conference on Artificial Intelligence), México
  - INTERFACES AND HUMAN COMPUTER INTERACTION – IHCI 2007, 2008, 2009 (se lleva a cabo en Europa)
  - Organizador de Track "Natural language processing", MICAI 2008, <http://www.micai2008.net/>, 27-31 octubre.
  - Natural Language Processing and Knowledge Engineering (2003, 2008), China, Japón.
  - Web 2.0 and Natural Language Engineering tasks at the 4th Indian International Conference on Artificial Intelligence (IICAI-09) December 16-18, 2009, Tumkur (Bangalore), India.
  - Iberamia, taller de lenguaje natural (2004) (Congreso internacional en Computación), México.

## **24. Publicaciones**

Más de 150 trabajos científicos, véase la lista de publicaciones.

## **25. Referencias a mis trabajos**

Total 159 citas (sin autocitas –citas tipo A; datos diciembre 2008):

- 64 por el índice ISI
- 27 en otras revistas
- 46 en congresos
- 7 en tesis de Doctorado
- 6 en informes técnicos
- 4 otras publicaciones
- 2 en tesis de Maestría
- 2 en libros
- 1 en patentes.

## Publicaciones del Dr. Grigori Sidorov

### Índice

Libros de investigación.....	15
Libros editados .....	15
Revistas editadas .....	16
Libros de docencia.....	16
Capítulos en libros.....	16
Revistas internacionales ISI-JCR .....	17
Revistas de padrón de CONACYT.....	19
Otras revistas de investigación con arbitraje .....	19
Revistas de divulgación.....	21
Patentes.....	21
Software con derechos de autor registrado.....	21
Trabajos inéditos .....	22
Informes técnicos .....	22
Congresos nacionales .....	23
Resúmenes en congresos internacionales .....	24
Congresos internacionales publicados por IEEE.....	24
Otros congresos internacionales in extenso.....	25

### Libros de investigación

1. Alexander Gelbukh, Grigori Sidorov. *Procesamiento automático del español con enfoque en recursos léxicos grandes*. IPN, ISBN 970-36-0264-9, 2006, 258 p.
2. Alexander Gelbukh, Grigori Sidorov. (Segunda edición ampliada y revisada) *Procesamiento automático del español con enfoque en recursos léxicos grandes*. IPN, 2010, 978-607-414-171-9, 307 p.
3. Grigori Sidorov. *Investigaciones en lingüística computacional*. (aceptado, IPN) 2009, 237 p.
4. Daniel Cazés Menache, Nelson del Castillo, Ricardo Mansilla, Luis. A. Pineda, Grigori Sidorov, Gerardo Sierra Martínez. *El dominio de la lingüística: más allá de las ciencias exactas y naturales*, 978-607-02-0875-1, UNAM, 2009, 213 p.

### Libros editados

5. A. Gelbukh, G. Sidorov, W. Olán Cristobal, J. Ángel Vera Félix (Eds) *Recientes avances en la ciencia de la computación en México*, Serie “Research on computing science”, vol. 7, CIC-IPN, ISBN 970-36-0149-9, 2004, 207 p.
6. G. Sidorov (Ed.) *Artificial intelligence for humans: Service robots and social modeling*, Sociedad mexicana de inteligencia artificial, Mexico, ISBN 978-607-00-0478-0, 2008, 161 p.
7. G. Sidorov (Ed.) *Métodos modernos de inteligencia artificial*, (preparado) México, 2010, 345 p.

8. G. Sidorov, A. Hernandez-Aguirre, C. Reyes-García (Eds) *Advances in Artificial Intelligence and Applications*, Proceedings of MICA I 2010 Special Session, IEEE Computer Society CPS, 2010, ISBN 978-0-7695-4284-3, 178 p.
9. G. Sidorov, A. Hernandez-Aguirre, C. Reyes-García (Eds) *Advances in Artificial Intelligence*, Lecture Notes in Artificial Intelligence 6438, Springer, 2010, 978-3-642-16760-7, 479 p.
10. G. Sidorov, A. Hernandez-Aguirre, C. Reyes-García (Eds) *Advances in Soft Computing*, Lecture Notes in Artificial Intelligence 6438, Springer, 2010, 978-3-642-16772-0, 515 p.

### **Revistas editadas**

11. **Editor en jefe**, revista de investigación y desarrollo tecnológico en computación “Polibits”, editada por el CIDETEC-IPN, enero 2008-hasta la fecha. ISSN 1870-9044.
12. Editor invitado: G. Sidorov, B. Cruz, M. Martínez, S. Torres (eds.) Special issue “Advances in computer science and engineering”, Journal “Research in computing science”, ISSN 1665-9899, vol. 34, 2008, 327 p.
13. Editor invitado: G. Sidorov (Ed.) Special issue “Advances in artificial intelligence and applications”, Journal “Research in computing science”, ISSN 1665-9899, vol. 40, 2008, 145 p.
14. Editor invitado: Yu. Ledeneva, G. Sidorov (eds.) revista “Informatica. International Journal of Computing and Informatics” (Slovenia, ISSN 0350-5596) <http://www.informatica.si/>. Número especial en lingüística computacional, vol. 34, 2010, 132 p.
15. Editor invitado: G. Sidorov, A. Hernandez-Aguirre, C. Reyes-García (Eds), Special issue “Advances in Soft Computing Algorithms”, Journal “Research in computing science”, ISSN 1665-9899, vol. 49, 2010, 183 p.

### **Libros de docencia**

16. B. Yu. Gorodetsky, T. S. Zevakhina, I. A. Kirillov, E. S. Kuz'mina, M. N. Mikhailov, E. E. Oleinikova, P. B. Parshin, Yu. K. Pirogova, L. P. Semenenko, G. O. Sidorov, L. P. Smirnova, M. V. Chubukova, I. V. Tsverava. *Problemas Lingüísticos – 93*. (en ruso) - Moscú, Rusia, “Rema”, 1993. 22 p.
17. V. M. Alpatov, A. A. Amelenkov, B. Yu. Gorodetsky, A. V. Gorshkov, E. Yu. Goriachev, T. S. Zevakhina, I. A. Kirillov, E. S. Kuz'mina, P. B. Parshin, Yu. K. Pirogova, V. A. Proskurkina, and G. O. Sidorov. *Problemas Lingüísticos – 94*. (en ruso) - Moscú, Rusia, “Rema”, 1994. 32 p.

### **Capítulos en libros**

18. G. O. Sidorov. *Lematización en el sistema automatizado para la formación del diccionario de estilo personal de los autores de literatura*. – Capítulo en: Palabra de Dostoyevski (en ruso), ISBN 5-88744-007-4. - Moscú, Rusia, Academia de Ciencias de Rusia, 1996. pág. 266-300.

19. Yoel Ledo Mezquita, [Grigori Sidorov](#), [Alexander Gelbukh](#), y Caridad Anías Calderón. *Búsqueda en bibliotecas digitales indexadas con sentidos de palabras*. In: Caridad Anías Calderón, Yoel Ledo Mezquita (eds.) *La telemática y su aplicación en la educación a distancia y en la informatización de la sociedad*. 2002, Editorial "Félix Varela", Cuba, ISBN 959-261-042-8, Tomo I, p. 288-299.
20. W. Olán Cristobal, A. Gelbukh, y G. Sidorov. *Desarrollo de herramientas para lexicógrafo con detección automática de relaciones semánticas implícitas*. En: A. Gelbukh, G. Sidorov, W. Olán Cristobal, J. A. Vera Félix (Eds.). *Recientes avances en la ciencia de la computación en México*. Serie "Research on computing science", vol. 7, CIC-IPN, 2004, pp 99-104.
21. Noé Alejandro Castro Sánchez, J. Ángel Vera Felix, Igor Bolshakov, Grigori Sidorov. *Representation of name sequences in Spanish using context free grammar*, In: J. Figueroa Nazuno, A. Gelbukh, C. Yáñez Márquez, O. Camacho Nieto (Eds.) "Advances in: Artificial intelligence, computing science and computer engineering", Series Research on computing science, vol. 10, IPN, 2004, pp 157-166.
22. Grigori Sidorov. *Conceptos básicos de inteligencia artificial y sus relaciones*. En: G. Sidorov (Ed.) "Métodos modernos de inteligencia artificial", (preparado) Mexico, 2090, pp. 3-18. Se presenta un análisis lingüístico de relaciones de los conceptos básicos de la inteligencia artificial que aclara las relaciones existentes entre ellos.

### **Revistas internacionales ISI-JCR**

23. A. Gelbukh, G. Sidorov, A. Guzman-Arenas. *Use of a weighted topic hierarchy for text retrieval and classification*. Lecture Notes in Artificial Intelligence, No. 1692, 1999, ISSN 0302-9743, Springer (<http://www-kiv.zcu.cz/events/tsd99/abstract.html>), pp. 130–135.
24. Alexander Gelbukh, Grigori Sidorov, and Igor A. Bolshakov. *Dictionary-based Method for Coherence Maintenance in Man-Machine Dialogue with Indirect Antecedents and Ellipses*. Lecture Notes in Artificial Intelligence N 1902, 2000, ISSN 0302-9743, Springer-Verlag, pp. 357–362.
25. [Alexander Gelbukh](#) and [Grigori Sidorov](#). *Zipf and Heaps Laws' Coefficients Depend on Language*. Lecture Notes in Computer Science N 2004, 2001, ISSN 0302-9743, Springer-Verlag, pp. 330–333.
26. [Alexander Gelbukh](#) and [Grigori Sidorov](#). *Automatic Selection of Defining Vocabulary in an Explanatory Dictionary*. Lecture Notes in Computer Science N 2276, 2002, ISSN 0302-9743, Springer-Verlag, pp. 300–303.
27. [Alexander Gelbukh](#), [Grigori Sidorov](#), and Liliana Chanona-Hernández. *Compilation of a Spanish representative corpus*. Lecture Notes in Computer Science N 2276, 2002, ISSN 0302-9743, Springer-Verlag, pp. 285–288.
28. Alexander Gelbukh, Grigori Sidorov, A. Guzmán-Arenas. *Relational Data Model in Document Hierarchical Indexing*. Lecture Notes in Computer Science, 2389, 2002, ISSN 0302-9743, Springer-Verlag, pp 259-262.
29. Alexander Gelbukh and Grigori Sidorov. *Approach to construction of automatic morphological analysis systems for inflective languages with little effort*. Lecture Notes in Computer Science, N 2588, 2003, ISSN 0302-9743, Springer-Verlag, pp. 215–220.

30. Yoel Ledo Mezquita, Grigori Sidorov, Alexander Gelbukh. *Tool for Computer-Aided Spanish Word Sense Disambiguation*. Lecture Notes in Computer Science, N 2588, 2003, ISSN 0302-9743, Springer-Verlag, pp. 277–280.
31. Alexander Gelbukh, Grigori Sidorov, SangYong Han, Liliana Chanona-Hernandez. *Automatic Evaluation of Quality of an Explanatory Dictionary by Comparison of Word Senses*. Lecture Notes in Computer Science, N 2890, Springer-Verlag, 2003, p. 555-561.
32. Alexander Gelbukh, Grigori Sidorov, San-Yong Han, and Erika Hernández-Rubio. *Automatic Enrichment of Very Large Dictionary of Word Combinations on the Basis of Dependency Formalism*. Lecture Notes in Artificial Intelligence N 2972, 2004, ISSN 0302-9743, Springer-Verlag, pp 430-437.
33. Alexander Gelbukh, Grigori Sidorov, San-Yong Han, and Erika Hernández-Rubio. *Automatic Syntactic Analysis for Detection of Word Combinations*. Lecture Notes in Computer Science N 2945, 2004, ISSN 0302-9743, Springer-Verlag, pp 243-247.
34. Alexander Gelbukh, Grigori Sidorov, SangYong Han. *On Some Optimization Heuristics for Lesk-Like WSD Algorithms*. Lecture Notes in Computer Science, N 3513, ISSN 0302-9743, Springer-Verlag, 2005, pp. 402–405.
35. R. Pazos, J. Perez, J. J. González, A. Gelbukh, G. Sidorov, and M. Rodríguez. *A domain independent natural language interface to databases capable of processing complex queries*. Lecture Notes in Artificial Intelligence, N 3789, ISSN 0302-9743, Springer-Verlag, 2005, pp. 833-842.
36. María de los Ángeles Alonso-Lavernia, Argelio Víctor De-la-Cruz-Rivera, and Grigori Sidorov. *Generation of Natural Language Explanations of Rules in an Expert System*. Lecture Notes in Computer Science N 3878, Springer-Verlag, ISSN 0302-9743, 2006, pp. 311-314.
37. Alexander Gelbukh, Grigori Sidorov, José Ángel Vera-Félix. *Paragraph-Level Alignment of an English-Spanish Parallel Corpus of Fiction Texts using Bilingual Dictionaries*. Lecture Notes in Artificial Intelligence, N 4188, ISSN 0302-9743, Springer-Verlag, 2006, pp. 61-67.
38. Alexander Gelbukh, Grigori Sidorov, José Ángel Vera-Félix. *A Bilingual Corpus of Novels Aligned at Paragraph Level*. Lecture Notes in Artificial Intelligence, N 4139, ISSN 0302-9743, Springer-Verlag, 2006, pp 16-23.
39. Alexander Gelbukh and Grigori Sidorov. Alignment of Paragraphs in Bilingual Texts using Bilingual Dictionaries and Dynamic Programming. *Lecture Notes in Computer Science*, N 4225, ISSN 0302-9743, Springer-Verlag, 2006, pp 824-833.
40. Alexander Gelbukh, Grigori Sidorov and Liliana Chanona-Hernandez. Lexical-based Alignment in the Task of Reconstruction of Parallel Text Structure. *Lecture Notes in Computer Science*, N 4592, ISSN 0302-9743, Springer-Verlag, 2007, pp 401-406
41. Juan Javier González Barbosa, Rodolfo A. Pazos, Alexander F. Gelbukh, Grigori Sidorov, Héctor J. Fraire H., I. Cristina Cruz C. Prepositions and Conjunctions in a Natural Language Interfaces to Databases. *Lecture Notes in Computer Science*, N 4743, ISSN 0302-9743, Springer-Verlag, 2007, pp 173-182

42. Grettel Barceló, Eduardo Cendejas, Igor Bolshakov y Grigori Sidorov. Ambigüedad en nombres hispanos. *Revista Signos. Estudios de Lingüística* 42 (70), ISSN 0035-0451. Chile, 2009, pp. 153-169 (JCR of Social Sciences).

### **Revistas de padrón de CONACYT**

43. [Alexander Gelbukh](#), [Grigori Sidorov](#), Igor Bolshakov. *On Coherence Maintenance in Man-Machine Dialogue with Ellipses*. *Computación y Sistemas, Revista Iberoamericana de Computación*, ISSN 1405-5546 (padrón de CONACYT), 5(3), 2002, pp 204-214.
44. Alexander Gelbukh, Grigori Sidorov, *Hacia la verificación de diccionarios explicativos asistida por computadora*, *Revista "Estudios de lingüística aplicada" UNAM*, ISSN 0185-2647 (padrón de CONACYT), 21 (38), 2003, pp 89-108.
45. Alexander Gelbukh, Grigori Sidorov y Francisco Velásquez. *Análisis morfológico automático en español a través de generación*. *Revista "Escritos" (padrón de CONACYT)*, 0188-6126, 28, 2003, pp. 9-26.
46. [Alexander Gelbukh](#), [Grigori Sidorov](#), [Adolfo Guzman-Arenas](#). *Document Indexing with a Concept Hierarchy*. *J. Computación y Sistemas, Revista Iberoamericana de Computación*, ISSN 1405-5546 (padrón de CONACYT), Vol. VIII No. 4, April-June 2005, pp 281-292.
47. Y. Ledo, G. Sidorov, A. Gelbukh. Recuperación de información con resolución de ambigüedad de sentidos de palabras. *Iberoamericana "Computación y sistemas" (padrón de CONACYT)*, 11 (3), 2008, pp. 288-300.

### **Otras revistas de investigación con arbitraje**

#### ***Lecture Notes in Computer Science since 2008***

48. Alexander Gelbukh, Grigori Sidorov, Diego Lara-Reyes and Liliana Chanona-Hernandez. Division of Spanish Words into Morphemes with a Genetic Algorithm. *Lecture Notes in Computer Science*, N 5039, ISSN 0302-9743, Springer-Verlag, 2008, pp 19-26.
49. Grettel Barcelo-Alonso, Eduardo Cendejas-Castro, Grigori Sidorov, and Igor A. Bolshakov. Formal Grammar for Hispanic Named Entities Analysis. *Lecture Notes in Computer Science*, N 5449, ISSN 0302-9743, Springer-Verlag, 2009, pp 183-194.
50. Grigori Sidorov, Obdulia Pichardo-Lagunas, and Liliana Chanona-Hernandez. Search Interface to a Mayan Glyph Database based on Visual Characteristics. *Lecture Notes in Computer Science*, Vol. 5723, ISSN 0302-9743, Springer-Verlag, 2009, pp. 222-229.
51. Eduardo Cendejas, Grettel Barceló, Alexander Gelbukh and Grigori Sidorov. Hybrid Algorithm for Word-Level Alignment of Parallel Texts. *Lecture Notes in Computer Science*, Vol. 5723, ISSN 0302-9743, Springer-Verlag, 2009, pp. 293-294.
52. Eduardo Cendejas, Grettel Barceló, Alexander Gelbukh and Grigori Sidorov. Incorporating Linguistic Information to Statistical Word-Level Alignment, 2009, *Lecture Notes in Computer Science*, Vol. 5856, 2009, pp. 387-394.
53. Noé Alejandro Castro-Sánchez and Grigori Sidorov. Analysis of Definitions of Verbs in an Explanatory Dictionary for Automatic Extraction of Actants based on Detection of Patterns. *Lecture Notes in Computer Science* 6177, 2010, pp 233-239.

54. Alexander Gelbukh, Grigori Sidorov, Eduardo Lavin-Villa, and Liliana Chanona-Hernandez. Automatic Term Extraction using Log-likelihood based Comparison with General Reference Corpus. *Lecture Notes in Computer Science* 6177, 2010, pp. 248-255.

***Otras revistas con arbitraje***

55. K. B. Saniev, O. G. Posypanov, G. O. Sidorov, y I. V. Tsverava. *Evaluación de Efectividad de Sistemas Manejados por Información*. (en ruso) - Industria Electrónica, 7, 1989. pág. 63-64.
56. I. V. Tsverava y G.O. Sidorov. *Métodos para Optimización de Interacción con las Bases de Conocimientos*. (en ruso) – Industria Electrónica, 7, 1989. pág. 62.
57. Alexander Gelbukh, Grigori Sidorov, and Adolfo Guzman-Arenas. *A system of Search, Classification, and Thematic Summarization of Documents with the use of a Hierarchical Thesaurus* (in Russian). In: Acta Academia 2000, International Informatization Academy (branch of Moldova) and United Nations, Evrica, Chisinau, Moldova, 2000, pp. 190-203.
58. A.F. Gelbukh, G.O. Sidorov. *Heaps law coefficients for Russian and English languages* (in Russian). In: Acta Academia 2001, International Informatization Academy (branch of Moldova) and United Nations, Evrica, Chisinau, Moldova, 2001, pp. 184-188.
59. Alexander Gelbukh, Grigori Sidorov *Coefficients of Zipf and Heaps laws for Russian and English*. J. Nauchno-Tehnicheskaya Informaciya, Seria 2 (NTI) (antes indexada por SCI), ISSN 0548-0027, vol 40, No 9, pp 32-37. Moscow, Russia, 2001.
60. A. Gelbukh, G. Sidorov, S. Galicia Haro, I. Bolshakov. *Environment for Development of a Natural Language Syntactic Analyzer*. In: Acta Academia 2002, Moldova, 2002, pp.206-213
61. Alexander Gelbukh, Grigori Sidorov. *Morphological Analysis of Inflective Languages Through Generation*. J. Procesamiento de Lenguaje Natural, No 29, September 2002. Sociedad Española para el Procesamiento de Lenguaje Natural (SEPLN), Depósito Legal B:3941-91 (indexed by INIST-CNRS, Institut de L'information Scientifique et Technique, Centre National de la Recherche Scientifique), ISSN 1135-5948, Spain, pp. 105-112
62. Alexander Gelbukh, Grigori Sidorov. *Selección automática del vocabulario definidor en un diccionario explicativo*. J. Procesamiento de Lenguaje Natural, No 29, September 2002. Sociedad Española para el Procesamiento de Lenguaje Natural (SEPLN), Depósito Legal B:3941-91 (indexed by INIST-CNRS, Institut de L'information Scientifique et Technique, Centre National de la Recherche Scientifique), ISSN 1135-5948, Spain, pp. 55-64.
63. Alexander Gelbukh, Grigori Sidorov, and Yoel Ledo Mezquita. *On Similarity of Word Senses in Explanatory Dictionaries*. International Journal of Translation, vol. 15, No 2, 2003, ISSN 0970-9819, pp. 51-60.
64. A. Gelbukh, G. Sidorov. *Automatic resolution of ambiguity of word senses in dictionary definitions*. Revista NTI (serie 2) (Información científica y técnica), ISSN 0548-0027, (antes indexada por ISI) Vol. 3, Rusia, 2004, pp 10-15.
65. Alexander Gelbukh, Grigori Sidorov, and Yoel Ledo-Mezquita. *Some Linguistic Methods of Improving of the Quality of Document Retrieval on the Internet*. IJEB (International Journal of Electronic Business), special issue on Internet and E-Business, ISSN 1470-6067, vol. 3, nom. 3/4, 2005, 264-275 pp.

66. Grigori Sidorov. *Linguist's Intelligent Workplace for Spanish Language*. In: Special issue "Advances in artificial intelligence and computer science", Journal "Research in computing science", ISSN 1665-9899, vol. 14, 2005, pp. 173-178.
67. David Pérez, José Tecuapacho, Héctor Jiménez-Salazar and Grigori Sidorov. *A term frequency range for text representation*. In: Special issue "Advances in artificial intelligence and computer science", Journal "Research in computing science", ISSN 1665-9899, vol. 20, 2006, pp. 113-118.
68. Grigori Sidorov and Noé Alejandro Castro-Sánchez. *Automatic emotional personality description using linguistic data*. In: Special issue "Advances in artificial intelligence and computer science", Journal "Research in computing science", ISSN 1665-9899, vol. 20, 2006, pp. 89-94.
69. Obdulia Pichardo-Lagunas and Grigori Sidorov. *Mayan Script Database*. In: Special issue "Advances in computer science and engineering", Journal "Research in computing science", ISSN 1665-9899, vol. 29, 2007, pp. 131-137.
70. J. Ernesto Gomez-Balderas, Grigori Sidorov, Hector Jimenez-Salazar. *Homonymy resolution and tagging of grammar classes without lexical information using neural networks*. In: Special issue "Advances in artificial intelligence and applications", Journal "Research in computing science", ISSN 1665-9899, vol. 32, 2007, pp. 305-314.
71. Guoxing Zhao, Miltos Petridis, Grigori Sidorov, Jixin Ma. A critical examination of node similarity graph matching algorithm. In: Special issue "Advances in artificial intelligence and applications", Journal "Research in computing science", ISSN 1665-9899, vol. 40, 2008, pp. 73-82.
72. Y. Ledeneva, G. Sidorov. Recent Advances in Computational Linguistics. "Informatica. International Journal of Computing and Informatics", 34 (2010) 3-18, ISSN 0350-5596, 2010.

### **Revistas de divulgación**

73. Grigori Sidorov. *Problemas actuales de lingüística computacional*. Revista digital universitaria, UNAM, México. Vol. 2 No. 1, ISSN 1607-6079, marzo 2001, 11 p.
74. Grigori Sidorov. *La capacidad lingüística de las computadoras*. Revista "Conversus", Vol. 36, 2005, ISSN 1665-2665, pp. 28-37.

### **Patentes**

75. K. B. Saniev, G. O. Sidorov, I. V. Tsverava y O. G. Posypanov. *Patente URSS N 1697533 "Método para Reconocimiento de Objetos y Sistema para Implementación del mismo"*. - Rusia, 1991.

### **Software con derechos de autor registrado**

76. F. Castillo Velásquez, L. Chanona Hernández, A. Gelbukh, G. Sidorov. AGME: programa de análisis y generación morfológica para el español, 24/10/2006, Num de reg. 03-2006-100911021300-01.

77. O. Pichardo Lagunas, G. Sidorov. Interfaz de búsqueda en una base de datos de los glifos maya basada en la descripción visual estructural, 2009, 03-2009-01291563500-01.

### **Trabajos inéditos**

78. G. O. Sidorov. *Diseño e implementación de los modelos, algoritmos y datos lingüísticos para los sistemas con el análisis y la generación morfológicos, para el ruso.* (en ruso) – Resumen de la tesis doctoral (publicado como un folleto) – Moscú, Rusia, 1995. - 29 pág.

### **Informes técnicos**

79. Grigori Sidorov, Alexander Gelbukh. *Demonstrative pronouns as markers of indirect anaphora.* Technical report. Series Azul, N 31, CIC, IPN, ISBN 970-18-3327-9, 1999. Reprint of the article published in Proc. 2nd International Conference on Cognitive Science and 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99), July 27-30, 1999, Tokyo, Japan, pp. 418-423
80. Manuel Montes-y-Gómez, Aurelio López-López, Alexander F. Gelbukh, Grigori Sidorov, Adolfo Guzmán-Arenas. *Text mining: new techniques and applications.* Technical report. No 34, CIC, IPN, 1999. ISBN 970-18-3512-3. 14 pp.
81. [Alexander F. Gelbukh](#), [Grigori Sidorov](#), Igor Bolshakov. *Coherence maintenance in human-machine dialogue: indirect antecedents and ellipses.* Technical report, Serie Azul, N 76, August 2000. CIC, IPN, ISBN 970-18-5428-4.
82. [Alexander F. Gelbukh](#), [Grigori Sidorov](#), Igor Bolshakov. *Dictionary-based method for coherence maintenance in human-machine dialogue with indirect antecedents and ellipses.* Technical report, Serie Azul, N 77, August 2000. CIC, IPN, ISBN 970-18-5492-2.
83. [Alexander F. Gelbukh](#), [Grigori Sidorov](#). *Word sense disambiguation in an explanatory dictionary and its usage in semantic network construction.* Technical report, Serie Azul, N 90, May 2001. CIC, IPN, ISBN 970-18-6617-7.
84. [Alexander F. Gelbukh](#), [Grigori Sidorov](#). *Algorithm of word sense disambiguation in an explanatory dictionary.* Technical report, Serie Azul, N 89, April 2001. CIC, IPN, ISBN 970-18-6617-7.
85. [Alexander Gelbukh](#), [Grigori Sidorov](#), and [Adolfo Guzman-Arenas](#). *Document indexing with a concept hierarchy.* Technical report, Serie Azul, N 101, June 2001. CIC, IPN, ISBN 970-18-6975-3.
86. [Alexander Gelbukh](#) and [Grigori Sidorov](#). *Selección automática del vocabulario definidor en un diccionario explicativo.* Technical report, Serie Azul, N 104, June 2001. CIC, IPN, ISBN 970-18-6978-8.
87. [Alexander Gelbukh](#), [Grigori Sidorov](#), and [Adolfo Guzman-Arenas](#). *Simple method for hierarchical conceptual indexing of documents using relational data model.* Technical report, Serie Roja, N 101, May 2002. CIC, IPN, ISBN 970-18-8082-X, 9 pp.
88. [Grigori Sidorov](#) and [Alexander Gelbukh](#). *Automatic detection of semantically primitive words using their reachability in an explanatory dictionary.* Technical report, Serie Azul, N 130, March 2002. CIC, IPN, ISBN 970-18-7902-3, 13 pp.

89. Alexander Gelbukh, Grigori Sidorov y Liliana Chanona Hernández *Compilación automática del corpus representativo para el español a través de Internet*, Informe técnico No 146, Serie Azul, CIC-IPN, México D.F., 2002, 78 p, ISBN 970-18-8242-3.
90. Alexander Gelbukh, Grigori Sidorov y Liliana Chanona Hernández. *Automatic calculation of similarity of word senses of a word in explanatory dictionaries and its applications*, Informe técnico No 184, Serie Azul, CIC-IPN, México D.F., 2003, 10 p.
91. Alexander Gelbukh, Grigori Sidorov y Erika Hernández Rubio. *Enriquecimiento automático de diccionario de colocaciones basándose en el parser sintáctico*. Informe técnico No 194, Serie Azul, CIC-IPN, México D. F., 2004, 20 p.
92. E. Gomez, A. Vera, O. Olivas, G. Sidorov. *Métodos estadísticos en procesamiento de lenguaje natural y su uso en alineación de los corpus paralelos*. CIC-IPN, México D. F., Informe técnico No 217, Serie Azul, 2006, ISBN 970-36-0331-9, 36 p.

### **Congresos nacionales**

93. G. O. Sidorov y I. V. Tsverava. *Acerca del estatus de los compuestos verbales con modificadores, en el ruso y el dungano*. (en ruso) – Memorias del 5º Congreso de los orientistas jóvenes de toda la Unión Soviética, Vol. 2. - Moscú, Rusia, “Nauka”, 1989. pág. 170-172.
94. K. B. Saniev, G. O. Sidorov, I. V. Tsverava, y O. G. Posypanov. *Algoritmo macroestructural para reconocimiento óptico de caracteres*. (en ruso) - Memorias del Congreso: Sistemas y Dispositivos Optoelectronicos. - Moscú, Rusia, 1990. pág. 22-25.
95. M. A. Ananyan, K. B. Saniev, G. O. Sidorov, y I. I. Suskov. *Kariotipización computarizada de cromosomas humanas con aberraciones por radiación*. (en ruso) - III congreso: Investigaciones en Radiación. Vol 2. - Puschino, Rusia, Academia de Ciencias de Rusia, 1997. pág. 64.
96. Grigori Sidorov, Alexander Gelbukh. *A Hierarchy of Linguistic Programming Objects*. Proc. ENC'99, Segundo Encuentro de Computación, Pachuca, Hidalgo, September 12 – 15, 1999, 6 p.
97. Alexander Gelbukh and Grigori Sidorov. *Recuperación de documentos con comparación semántica suave*. Proc. TAINA-2002, Workshop on Soft Computing at MICAI'2002: 2nd Mexican International Conference on Artificial Intelligence, Merida, Mexico, April 2002, pp 253-261.
98. Paredez Viscarra, C. A., J. Vallejo Scott, y G. Sidorov. *Notificador de páginas Web nuevas y modificadas*. 7 verano de investigación científica del pacífico, 2002, México, p. 247.
99. Gabriela Medina Garay y Grigori Sidorov. *Desarrollo de las herramientas de análisis automático de los diccionarios*. 9 verano de investigación científica del pacífico, 2004, México, p. 441.
100. Juan Manuel Núñez Maldonado y Grigori Sidorov. *Herramienta de transformación automática de los diccionarios bilingües*. 10 verano de investigación científica del pacífico, 2005, México. 1 p.

## **Resúmenes en congresos internacionales**

101. Alexander Gelbukh, Grigori Sidorov, and Igor A. Bolshakov. *Role of lexical meaning in resolution of hidden anaphora for demonstrative pronouns*. Proc. of 7<sup>st</sup> International Pragmatics Conference, Budapest, Hungary, 9-14 July 2000, p. 266.
102. Yoel Ledo Mezquita, Grigori Sidorov, y Alexander Gelbukh. *Búsqueda inteligente en bibliotecas digitales*. Resumen publicado. In: Proc. of Int. Conf. TelEduc-2002. Habana, Cuba, 2002.
103. Alexander Gelbukh, Grigori Sidorov, and L. Chanona-Hernández. *Is word sense disambiguation useful in information retrieval?* SSGRR 2003s, International Conference on Advances in Infrastructure for e-Business, e-Education, e-Science, e-Medicine, and Mobile Technologies on the Internet, track ISY, Scuola Superiore G. Reiss Romoli, L'Aquila, Italy, July 28 – August 3, 2003, p. 17.
104. Alexander Gelbukh, Grigori Sidorov. *Representividad morfológica en los corpus*. En: Memorias del VIII encuentro internacional de lingüística en el noroeste, Hermosillo, México, noviembre 2004, p 208.
105. N. Castro Sanchez, G. Sidorov, *Análisis Cuantitativo del discurso político electoral*. In: Rodríguez Alfano, Lidia (ed.). VIII Congreso de la ALED, Programa y resúmenes. ISBN: 978-607-433-268-1. Monterrey, México: Facultad de Filosofía y Letras, Universidad Autónoma de Nuevo León, 2009, p. 137.

## **Congresos internacionales publicados por IEEE**

106. Alexander Gelbukh, Grigori Sidorov, A. Guzmán-Arenas. *Document classification with a weighted topic hierarchy*. Proc. 1st International Workshop on Document Analysis and Understanding for Document Databases (DAUDD'99), 10th International Conference and Workshop on Database and Expert Systems Applications (DEXA), Florence, Italy, September 1, 1999. IEEE Computer Society Press ([http://mcculloch.ing.unifi.it/~docproc/DAUDD99/daudd\\_program.html](http://mcculloch.ing.unifi.it/~docproc/DAUDD99/daudd_program.html)), pp. 566 - 570.
107. A. Gelbukh, G. Sidorov, and A. Guzmán-Arenas. *A Method of Describing Document Contents through Topic Selection*. Proc. SPIRE'99, International Symposium on String Processing and Information Retrieval, Cancun, Mexico, September 22 – 24. IEEE Computer Society Press, 1999 (<http://garota.fismat.umich.mx/spire99>), pp. 73-80.
108. Alexander Gelbukh, Grigori Sidorov, and Igor A. Bolshakov. *Coherence Maintenance in Human-Machine Dialogue: Indirect Antecedents and Ellipses*. Proc. DEXA-2000, 11<sup>th</sup> International Conference and Workshop on Database and Expert Systems Applications, NLIS-2000, 2<sup>nd</sup> International Workshop on Natural Language and Information Systems, Greenwich, England, September 4-8, 2000. **IEEE Computer Society Press**. pp. 96–100.
109. [Grigori Sidorov](#), [Alexander Gelbukh](#). *Automatic extraction of semantic primitive words*. Proc. International IEEE SMC-2001 Conference: Systems, Man, and Cybernetics. Tucson, USA, October 7–10, 2001, **IEEE** (Abstract in printed version, *in extenso* in CD-ROM version), ISBN 0-7803-7087-2, 0-7803-7088-0 (microfiche), 0-7803-7089-9 (CD-ROM), ISSN 1062-922X, IEEE Catalog N 01CH37236C. pp. 1683–1687 <http://crater.sie.arizona.edu/nlpke.html>, <http://crater.sie.arizona.edu/index2.html>.

110. Alexander Gelbukh, Sang-Yong Han, Grigori Sidorov. *Compression of boolean inverted files by document ordering*. NLP-KE 2003 (International conference on Natural Language processing and knowledge engineering), Beijing, China, 2003, pp 244-249.

### **Otros congresos internacionales in extenso**

111. M. N. Mikhailov y G. O. Sidorov. *Sistema tutorial de gramática del ruso para los hablantes no nativos*. (en ruso) – Memorias del Congreso Internacional: Multimedia en enseñanza de idiomas. - Universidad Estatal "Lomonosov" de Moscú, Moscú, Rusia, 1995. pág. 50-52.
112. G. O. Sidorov. *Sistema para acentuación automática de palabras en el ruso*. – (en ruso) Memorias del II Simposium Internacional de AIMLIR (Asociación internacional de maestros de literatura e idioma rusos), Fonética en el sistema de lenguaje. (en ruso) - Universidad de enseñanza de idiomas "Patrice Lumumba", Moscú, Rusia, 1996. pág. 173-174.
113. G. O. Sidorov. *Sistema con lematización para el análisis de texto*. (en ruso) – Memorias del Congreso internacional: Lingüística funcional: Principios de formación de texto. - Symferopol, Ucrania, 1996. pág. 186-187.
114. K. B. Saniev y G. O. Sidorov. *El sistema inteligente para kariotipización automatizada*. (en ruso, resumen en inglés) - Memorias del VI Congreso internacional: Knowledge-Dialogue-Solution (KDS-97). Yalta, Ucrania, 1997. pág. 567-572.
115. G. O. Sidorov and A. F. Gelbukh. *A method of detection and resolution of hidden anaphora* (in Russian, abstract in English). Proc. Annual International Conf. on Applied Linguistics Dialogue-99, May 30 – June 5, 1999, Moscow, Russia. pp 288-297.
116. Alexander Gelbukh, Grigori Sidorov, and Adolfo Guzmán-Arenas. *Text categorization using a hierarchical topic dictionary*. Proc. Text Mining workshop at 16th International Joint Conference on Artificial Intelligence (IJCAI'99), Stockholm, Sweden, July 31 – August 6, 1999 ([http:// www.dsv.su.se/ ijcai-99](http://www.dsv.su.se/ijcai-99)), pp 34-35.
117. Alexander F. Gelbukh and Grigori Sidorov. *On Indirect Anaphora Resolution*. Proc. PACLING-99, Pacific Association for Computational Linguistics, ISBN 0-9685753-0-7, University of Waterloo, Waterloo, Ontario, Canada, August 25-28, 1999 ([http:// www.lpaig.uwaterloo.ca/ ~b2hui/ pacling/ presenters.html](http://www.lpaig.uwaterloo.ca/~b2hui/pacling/presenters.html)), pp. 181-190
118. G. Sidorov, I. Bolshakov, P. Cassidy, A. Gelbukh, and S. Galicia-Haro. *'Non-adult' semantic field*. Proc. 3rd Tbilisi Symposium on Language, Logic, and Computation. Batumi, Georgia, September 12–16, 1999 ([http:// www.illc.uva.nl/ Batumi](http://www.illc.uva.nl/Batumi)).
119. Alexander Gelbukh, Grigori Sidorov, and Adolfo Guzman-Arenas. *A system for search and classification of the documents with the use of a hierarchic thematic dictionary* (in Russian, abstract in English). Accepted to Proc. 8th International Conference Knowledge-Dialogue-Solution (KDS-99), Yalta, Ukraine, September 13-18, 1999.
120. Grigori Sidorov, Alexander Gelbukh. *Demonstrative pronouns as markers of indirect anaphora*. Proc. 2nd International Conference on Cognitive Science and 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (ICCS/JCSS99), July 27-30, 1999, Tokyo, Japan, pp. 418-423

121. Grigori Sidorov, Alexander Gelbukh. *La resolución de correferencia con un diccionario de escenarios*. Proceedings of CIC-99, Congreso Internacional en Computación, November 15 - 19, 1999, Mexico D.F., pp 382-391.
122. Alexander Gelbukh and Grigori Sidorov. *A dictionary-based algorithm for indirect anaphora resolution*. In Proc. Venecia per il Trattamento Automatico delle Lingue (VEXTAL), Venice, Italy, November 22 to 24 1999 ([http:// byron.cgm.unive.it/ eventi/ VEXTAL/ EVEXTAL.html](http://byron.cgm.unive.it/eventi/VEXTAL/EVEXTAL.html)), pp 169-173.
123. Sidorov, G. *Methods of analysis of combinability of words in Russian*. In: Taal en cultur (Language and culture), materials of the conference "Belgium - Holland - Russia", (in Russian), ISBN 5-02-022599-1, UNIVESIDAD DE MAASTRICHT Y "NAUKA", Maastricht, Holland and Moscow, Russia, 1999, pp. 294-302.
124. Igor A. Bolshakov, Alexander F. Gelbukh, Grigori Sidorov. *On Subject-Predicate Agreement in Spanish*. Proc. of International conference CICLing2000, Mexico, February 13-19, 2000. pp. 59-67
125. G. O. Sidorov, I. A. Bolshakov, P. Cassidy, S. Galicia-Haro, A. F. Gelbukh. *A Comparative analysis of the semantic field "non-adult" in Russian, English, and Spanish* (in Russian). Proc. Annual International Conf. on Applied Linguistics Dialogue-2000, June 2000, Moscow, Russia.
126. Alexander Gelbukh, Grigori Sidorov, and Igor A. Bolshakov. *Coherence Maintenance in Man-Machine Dialogue with Ellipses*. Proc. MICAI-2000, 1st Mexican International Conference on Artificial Intelligence, Acapulco, Mexico, April 2000, pp 386-392.
127. Alexander Gelbukh, Grigori Sidorov. *Sistema inteligente de recuperación de textos políticos de la base de datos documental del Senado de la República Mexicana*. Proc. CIC-2000, Congreso Internacional en Computación, November 15 - 17, 2000, CIC, IPN, Mexico D.F., ISBN 970-18-5540-X, pp. 315-321.
128. Alexander Gelbukh, Grigori Sidorov. *Algorithm of word sense disambiguation in an explanatory dictionary*. Proc. of COMPLEX-2001, Workshop on Computational Lexicography, Birmingham, Great Britain, June 28-30, 2001, pp. 35-40.
129. Grigori Sidorov, Alexander Gelbukh. *Word sense disambiguation in a Spanish explanatory dictionary*. Proc. of TALN-2001 (Tratamiento automático de lenguaje natural), Tours, France, July 2-5, 2001, legal deposit 06017105-200, pp 398-402.
130. Alexander Gelbukh, Grigori Sidorov, and Adolfo Guzman-Arenas. *Document Indexing With a Concept Hierarchy*. Proc. of NDDL-2001, 1<sup>st</sup> International Workshop on New Developments in Digital Libraries at ICEIS-2001, International Conference on Enterprise Information Systems, Setúbal, Portugal, July 6-7, 2001, ISBN 972-98050-4-0, pp 47-54.
131. Alexander Gelbukh, Grigori Sidorov, and Liliana Chanona-Hernández. *System for automatic building of a representative corpus using Internet*. Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities (ALLC / ACH 2002), Tübingen, July 24-28, 2002, pp 118-120.
132. Grigori Sidorov, Alexander Gelbukh, y Liliana Chanona Hernández. *Corpus virtual virtual: un diccionario grande de contextos de palabras españolas compilado a través de*

- Internet*. Proc. of Workshop “Multilingual information access and natural language processing” of IBERAMIA 2002 (8th Iberoamerican conference on Artificial Intelligence), Sevilla, Spain, November, 12, 2002, pp 7-14.
133. Francisco Velásquez, Alexander Gelbukh, y Grigori Sidorov. *AGME: un sistema de análisis y generación de la morfología del español*. Proc. of Workshop “Multilingual information access and natural language processing” of IBERAMIA 2002 (8th Iberoamerican conference on Artificial Intelligence), Sevilla, Spain, November, 12, 2002, pp 1-6.
  134. Alexander Gelbukh, Grigori Sidorov, and Sang-Yong Han. *Comparison of Some Global Coherence Optimization Heuristics for Word Sense Disambiguation*. In: J.-L. Diaz de León Santiago, G. Gonzalez-Santos, J. Figueroa (Eds.) “Avances en: Ciencias en la computación”, Proc. of Int. Conf. CIC, vol. 3, IPN, 2003, pp 181-185.
  135. Alexander Gelbukh, Grigori Sidorov, Marina Chubukova, and Erika Hernández-Rubio. *Dictionaries of word combinations: Which method of compilation is better?* In: Proc. of International Conference Dialogue-2004 on computational linguistics and its applications, Moscow, Russia, June 2004, pp. 133-138.
  136. Alexander Gelbukh, Grigori Sidorov. *Analizador morfológico disponible: un recurso importante para PLN en español*. En: Memorias de talleres del congreso internacional Iberamia-2004, Puebla, México, noviembre 2004, pp. 209-216.
  137. J. Ángel Vera Félix, Grigori Sidorov. *Proyecto de preparación del corpus paralelo alineado español-inglés*. En: Memorias del 5o encuentro internacional de computación ENC-2004, Colima, México, septiembre 2004, Sociedad Mexicana de Ciencias en Computación, pp. 235-242.
  138. Noé Alejandro Castro Sánchez, J. Ángel Vera Felix, Igor Bolshakov, Grigori Sidorov. *Formalización del sistema de nombres hispanos*. En: Memorias del 5o encuentro internacional de computación ENC-2004, Colima, México, septiembre 2004, Sociedad Mexicana de Ciencias en Computación, pp. 289-295.
  139. Tania Lugo, Alexander Gelbukh, Grigori Sidorov. *Unsupervised learning of word combinations for syntactic disambiguation*. En: Memorias del 5o encuentro internacional de computación ENC-2004, Colima, México, septiembre 2004, Sociedad Mexicana de Ciencias en Computación, pp. 311-319.
  140. Alexander Gelbukh and Grigori Sidorov. *Exploring large document collections with a dynamic hierarchy*. In: Proc of the Third International Conference on Information, Nov. 29-Dic. 2, 2004, Tokyo, Japan, pp. 573-576.
  141. Alexander Gelbukh, Grigori Sidorov. *On Automatic Morphological Analysis of Inflective Languages*. In: Proc. of International Conference on computational linguistics and its applications Dialogue-2005, 2005, Russia, pp 92-96.
  142. Grigori Sidorov. *Etiquetador Morfológico y Desambiguador Manual: Dos Aplicaciones del Analizador Morfológico Automático para el Español*. En: Memorias del VI encuentro internacional de computación ENC-2005, México, Puebla, septiembre, 2005, pp. 147-149.

143. Alexander Gelbukh, Grigori Sidorov, José Ángel Vera-Felix. *Dictionaries in tasks of automatic processing of pairs of translated texts*. In: Proc. of International Conference Dialogue-2006 on computational linguistics and its applications, Moscow, Russia, June 2006, pp. 110-115.
144. Grigori Sidorov, Noé Alejandro Castro-Sánchez. *System for linguistically-based evaluation of psychological profiles*. In: International Conference Dialogue-2006 on computational linguistics and its applications, Moscow, Russia, 2006, pp. 464-467.
145. Grigori Sidorov, Omar Olivas Zazueta. Resolución de anáfora pronominal para el español usando el método de conocimiento limitado. In: “Avances en la Ciencia de la computación”, Proc. of 7 Int .Conf. ENC-2006, San Luís Potosi, México, 2006, pp.276-281.
146. Gaspár Ramírez, James L. Fidelholtz, Héctor Jiménez, Grigori Sidorov. Elaboración de un diccionario de verbos del español a partir de una lexicografía sistemática. In: “Avances en la Ciencia de la computación”, Proc. of 7 Int .Conf. ENC-2006, San Luís Potosi, México, 2006, pp.270-275.
147. Grigori Sidorov. *Herramienta interactiva para procesamiento semiautomático de textos a nivel morfológico para el español*. IX encuentro internacional de lingüística en el noroeste, 2006, Hermosillo, Sonora, pp 263-265.
148. Ledo-Mezquita, Y., G. Sidorov, V. Cubells. Combined Lesk-based Method for word sense disambiguation. In: Proc. of Int. Conf. on Computing CIC-2006, México, 2006, pp. 105-108.
149. Alexander Gelbukh, Grigori Sidorov and Liliana Chanona-Hernandez. *Dynamic programming with lexical similarity calculation in alignment of parallel texts and its evaluation*. In: Proc. of International Conference Dialogue-2007 on computational linguistics and its applications, Moscow, Russia, June 2007, pp 121-129.
150. J. Ernesto Gómez Balderas, Grigori Sidorov, Héctor Jiménez Salazar. *Redes neuronales basados en contextos para la resolución de homonimia morfológica en el español*. 14th International Congress on Computer Science Research CIICC'07, Orizaba, 2007, pp. 94-104.
151. Alexander Gelbukh, Grigori Sidorov, Diego Lara-Reyes, Liliana Chanona-Hernandez, Marina Chubukova. *Genetic Algorithm for Automatic division of words into morphemes*. In: Proc. of International Conference Dialogue-2008 on computational linguistics and its applications, Moscow, Russia, June 2008, pp. 97-102.
152. Obdulia Pichardo Lagunas, Grigori Sidorov. *Diccionario de los glifos maya con descripción visual estructural*. In: Proc. of International Conference EURALEX-2008, Barcelona, Spain, July 2008, pp 747-751.
153. Grigori Sidorov, Alberto Barrón-Cedeño and Paolo Rosso. English-Spanish Large Statistical Dictionary of Inflectional Forms. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta. European Language Resources Association (ELRA), 2010, pp. 277-281.
154. Alexander Gelbukh, Grigori Sidorov, Eduardo Lavin-Villa, Liliana Chanona-Hernandez. Automatic detection and classification of single-word terms in a specific

domain corpus using log-likelihood similarity with general purpose corpus. In: Proc. of International Conference Dialogue-2010 on computational linguistics and its applications, Moscow, Russia, May 2010, pp. 82-89.