ESTE PROGRAMA ES PÚBLICO, AJENO A CUALQUIER PARTIDO POLÍTICO. QUEDA PROHIBIDO EL USO PARA FINES DISTINTOS A LOS ESTABLECIDOS EN EL PROGRAMA.

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"

ipn.mx

CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN

RETRATOS es una edición periódica especial editada por el Centro de Investigación en Computación del Instituto Politécnico Nacional (CIC IPN), en homenaje a sus fundadores dentro del marco de los festejos del XXV Aniversario del CIC IPN.

Número 6 de ocho entregas de la 1ª edición.

Registro ante el Instituto Nacional del Derecho de Autor y Certificado de Licitud de Título y Contenidos por la Comisión calificadora de publicaciones y revistas ilustradas de la Secretaría de Gobernación en trámite.

Queda prohibida la reproducción total o parcial, sin el permiso expreso del Productor Editorial.

Agradecimientos

Al **Dr. Alexander Gelbukh** por su disponibilidad para la realización de la entrevista, así como las facilidades para obtener material gráfico.

Créditos

Director Interino:
Dr. Francisco Hiram Calvo Castro
Producción General:
Mtro. Mauricio Sebastián Martín Gascón
Producción Editorial:
Alejandra Berenice Landeros Barraza
Entrevistas y Redacción:
Daniel Chávez Fragoso
Diseño Gráfico:
María Paulina Alcántara Rodríguez
Logística y Apoyo Técnico:
Claudia Cortés Rivera
Logística General y Difusión:
Departamento de Relaciones Públicas e Imagen del Centro de Investigación en Computación.

Índice

5 Alexander Gelbukh, de Rusia para “México lindo y querido”

6 Un mundo cerrado

7 A veces, los ojos que no miran bien te llevan a buen puerto

8 Salir de la galaxia

10 Lograr que las máquinas entiendan nuestro lenguaje

12 Laboratorio de Procesamiento de Lenguaje Natural del CIC IPN: Una ventana al mundo, a un mundo más grande

14 Galería

DR. ALEXANDER GELBUKH

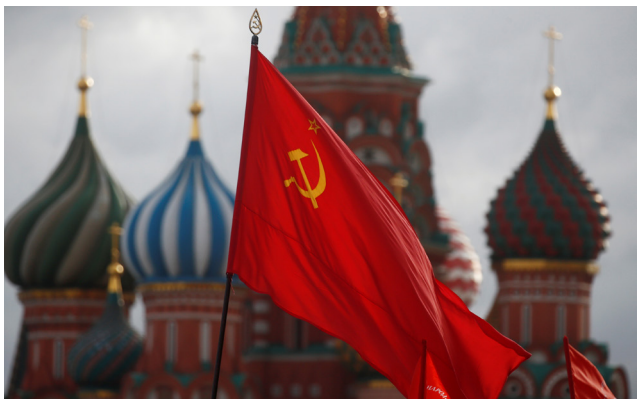
Alexander Gelbukh, de Rusia para “México lindo y querido”

“El ser humano es extremadamente insistente en hacer lo que le interesa y lo hace bien, un niño es incansable en jugar, no le pagan, quiere jugar todo el día y no se le puede mandar a la cama, quiere seguir jugando. Pienso que un científico es un niño que no creció, sus juguetes ahora son más grandes, a lo mejor son aceleradores de partículas, pero sigue jugando, es incansable no se le puede mandar a dormir y hace su trabajo verdaderamente bien.” Alexander Gelbukh.

El Dr. Gelbukh en Barcelona, España.

Un mundo cerrado

La Unión de Repúblicas Socialistas Soviéticas (URSS) dejó de existir oficialmente el 26 de diciembre de 1991, durante su presencia, de casi 70 años, las personas habitamos en un planeta dividido y enfrentado, al grado que la idea del fin del mundo por una guerra nuclear formaba parte de la vida cotidiana.



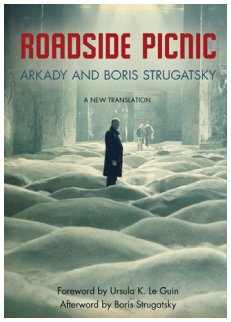
Bandera URSS.

Aunque la URSS era una potencia militar, gran parte de sus ciudadanos vivían aislados del resto del mundo, en Moscú, el corazón del bloque comunista, nació en 1962 Alexander Gelbukh, hijo único cuyo padre era doctor en filosofía y su madre bióloga, de ella aprendió el gusto por la naturaleza: “Mi mamá se graduó de la facultad de biología y me enseñó muchas cosas de los seres vivos, de las plantas, me gustaba mucho estudiar los insectos. En la ciudad no había muchas posibilidades de interactuar con la naturaleza, pero en el verano me reunía con amigos que tenían el mismo interés de estudiar las ranas o los insectos”.

Otro de los intereses de Alexander era la ciencia ficción, sus autores favoritos eran los hermanos Arkadi y Boris Strugatski, en cuya famosa novela, Roadside Picnic, traducida al español como Picnic junto al camino, se basó Andréi Tarkovski para rodar en 1979 la célebre película Stalker. Sobre los hermanos Strugatski el Dr. Gelbukh señala: “Mi historia favorita es El lunes empieza el sábado, es sobre los investigadores y la vida en un centro de investigación, porque eso dicen que los investigadores no buscan tomar vacaciones y el fin de semana ya empiezan a trabajar en su investigación, como un niño empieza a jugar desde la mañana, porque le interesa”.

Hermanos Strugatsky.

Quizá la mezcla del gusto por la naturaleza y la ciencia ficción derivaron en un gran interés por el universo, así lo recuerda el Dr. Gelbukh: “Me gustaban las cosas de la física, tenía mucho interés en la mecánica cuántica, eso era lo que quería estudiar, leía mucha ciencia ficción, astronomía y astrofísica, no me interesaba la astronomía tipo cómo se llamaban las constelaciones, sino más bien cómo funcionaban las estrellas y las constelaciones”.



Novela “Roadside Picnic”, autores Arkadi y Boris Strugatski .

A veces, los ojos que no miran bien te llevan a buen puerto

Alexander Gelbukh quería estudiar en la facultad de física, pero lo rechazaron por un examen... de la vista, así lo recuerda: "Parece ser que tenía problemas para distinguir los matices sutiles de los colores, me sugirieron la facultad de matemáticas, allí sí me admitieron y me gustó mucho, allí me quedé. Algunos estudiantes se dedicaban sólo a salir con las chicas y a tomar cerveza, pero mis amigos y yo, además de eso, estudiábamos mucho, cuando trataba de presumir algo a las chicas o verme bien, les explicaba los teoremas para el examen, y en esta facultad las chicas valoraban a los chicos inteligentes".

El desempeño académico de Alexander fue siempre excelente y en la facultad de matemáticas llevaban una materia de programación que, posteriormente, sería muy importante en su vida: "Teníamos una clase de programación, se conectaban los monitores a una supercomputadora, no había computadoras personales. Me gustaba mucho jugar con la computadora, no hablo de videojuegos, sino de programar, allí había algo de matemáticas, trataba de hacer un programa que imprimiera su propio código, lo cual no es trivial".

Al concluir los estudios de matemáticas de nivel de licenciatura continuó con la Maestría en Ciencias en el mismo Departamento de Matemáticas de la Universidad Estatal de Moscú y empezó a trabajar como programador en el Centro de la Información Científica y Técnica de toda Rusia de la Academia de Ciencias, donde desarrollaban un sistema de inteligencia artificial, con base en su trabajo sobre este sistema realizó el doctorado en la misma institución. Su tesis ya trataba de lingüística computacional, sobre su inicio en esta rama de la informática recuerda: "El jefe me dio la tarea de hacer un programa que entienda el lenguaje natural, le dije «ahorita lo hago», pasó un mes de trabajo intenso y luego otro y resultó no tan fácil como el jefe creía, actualmente sigo en lo mismo, ahora entiendo que no me va a salir muy pronto, pero la ciencia que conlleva el trabajo de hacerlo es muy interesante".

Monitores de supercomputadoras.

Universidad Estatal de Moscú

Academia de Ciencias de Rusia.

Salir de la galaxia

La vida de Alexander Gelbukh estaba bastante encaminada, se había casado con una brillante compañera de la facultad de matemáticas (actualmente Doctora en Matemáticas) y trabajaba como programador, lo cual le permitía sobrellevar la difícil situación que se vivía en la URSS.

Desde 1985 Mijaíl Gorbachov había impulsado una serie de reformas bajo los términos de glasnost (“transparencia” o “libertad de expresión”) y perestroika (“reconstrucción”, o “transformación”) encaminadas a sacar del estancamiento a la economía soviética, aunque hubo un poco más de libertad de expresión, la crisis económica se recrudeció. Así recuerda el Dr. Gelbukh esta etapa: “La dictadura comunista era muy dura, cruel, en algún momento se nos permitió opinar, era un cambio muy brusco e inesperado, hasta que un día en el trabajo alguien llegó con la noticia de que había un golpe de estado, una «Junta», (desde el golpe de Estado de la junta militar chilena a Allende, en Rusia se usa la palabra Junta, en español, como alusión al fascismo, represalias y opresión de la libertad) había quitado a Gorbachov del poder, querían acabar con los cambios liberales, porque contradecían la doctrina comunista. Boris Yeltsin, el gobernador del estado (República) de Rusia se opuso a la Junta, él estaba en la Casa Blanca de Rusia (el edificio del gobierno estatal, lo llamamos así por su color y por analogía con la Casa Blanca de los Estados Unidos) y miles fuimos a rodear el edificio para protegerlo”.

La Junta estaba conformada por miembros de la “línea dura” del partido comunista y de la KGB que buscaban, de acuerdo a su comunicado oficial, “restaurar el honor y la dignidad del hombre soviético”, cerraron los periódicos y los medios que no les eran afines, mientras que Boris Yeltsin y otros funcionarios emitieron una declaración señalando que la junta encabezaba un golpe reaccionario. Sobre esto recuerda el Dr. Gelbukh: “Tomamos un carro de uno de nosotros y fuimos a hacer propaganda y a hablar con los soldados sobre el porqué no debían apoyar a la Junta, que ya había llamado a cuatro divisiones de tanques a Moscú, en la carretera se veían tanques de horizonte a horizonte con los cañones hacia la ciudad, construimos barricadas con todo lo que encontrábamos en la calle, pero no eran nada contra los tanques, también teníamos cocteles Molotov y pasamos toda la noche alrededor de la Casa Blanca esperando, pero aunque la Junta ordenó atacar, los soldados no obedecieron”.

“Golpe de agosto”

Gorbachov y Yeltsin.

Tras el fallido “Golpe de agosto”, la economía se siguió deteriorando, el salario oficial del Dr. Gelbukh era muy poco, su trabajo vía remota como programador para una empresa de Estados Unidos le permitía tener una situación más cómoda, pero no la percibía como algo estable y tenía que hacer lo que le ordenaban, no lo que le interesaba hacer.

Para la creación del Centro de Investigación en Computación del IPN (1996), el Dr. Adolfo Guzmán Arenas, director fundador, se dio a la tarea de sumar talento internacional a la planta de profesores, invitó a venir al Dr. Igor Bolshakov, recomendado por un ministro del Gobierno de la Ciudad de Moscú, a quien había conocido en una visita de funcionarios de la URSS a su laboratorio, el Dr. Bolshakov quería que le acompañara el Dr. Gelbukh, que se encargaba de la programación en sus proyectos y de quien había sido asesor en la tesis de doctorado.

Además de la inestable situación económica, una razón importante para que el Dr. Gelbukh saliera de la URSS fue la de conocer otros países: “Mi generación percibía a la Unión Soviética como algo permanente, para siempre, yo imaginaba mi vida como dentro de un vaso, sus paredes eran las fronteras de la Unión Soviética y nunca podría salir del vaso. Uno sabe que no puede ir a otra galaxia, es un hecho indudable, pero imagínese que un día despierta en la galaxia de Andrómeda, así me he sentido los últimos 25 años”.

En 1997 los Drs. Bolshakov y Gelbukh llegaron al CIC y fundaron el Laboratorio de Procesamiento de Lenguaje Natural, así lo recuerda el Dr. Gelbukh: “Llegué a vivir a Tlatelolco, al mismo edificio del Dr. Bolshakov, la adaptación fue sencilla, la idiosincrasia del ruso y el mexicano es parecida y el lenguaje es fácil porque se pronuncia como se escribe. Al llegar fundamos el laboratorio de procesamiento de lenguaje, era de interés del Dr. Guzmán, el entonces Director del CIC, él había trabajado en algo parecido, al inicio éramos los dos, luego llegó el Dr. Mikhail Alexandrov quien ahora está en Barcelona”.

Dr. Igor Bolshakov y Dr. Gelbukh en el Simposium Internacional de Computación, 1998.

Dr. Gelbukh en convivio realizado en el CIC, 2003.

Lograr que las máquinas entiendan nuestro lenguaje

El año de su llegada, 1997, el Dr. Gelbukh escribió su primer artículo en inglés, en 1998 participó en su primer congreso internacional, para él, publicar en Springer era un sueño, mismo que realizó en 1999 y en 2001 editó su primer libro para esa casa editorial global.

De su inicio en el estudio del procesamiento del lenguaje natural, hasta hoy, esta rama de la informática ha evolucionado de ser puramente teórica, a ser una de las más buscadas por las empresas de TIC para generar aplicaciones prácticas, así lo percibe el Dr. Gelbukh: "Esta ciencia floreció con Internet 2 y con los teléfonos inteligentes, significó el acceso público a internet y la gente empezó a dejar allí un montón de textos y de datos, lo que formó los famosos datos grandes con los cuales se pudieron utilizar los métodos del aprendizaje automático".

Dentro del procesamiento del lenguaje natural, el Dr. Gelbukh ha desarrollado líneas de investigación como la detección de plagio, la tecnología de respuesta a preguntas (tipo Google), extracción de información de un discurso, análisis de sentimientos, análisis de personalidad, generación de resúmenes, registro de las lenguas indígenas y muchas otras, con la mayoría ha obtenido destacados resultados, algunos de ellos son:

La detección de plagio, entre sus alumnos el Dr. Miguel Sánchez Pérez, que trabaja actualmente en Oracle Guadalajara, obtuvo en 2014 el primer lugar en la categoría Alineación de Textos, del *11th Evaluation Lab on Uncovering Plagiarism, Authorship, and Social Software Misuse* que se celebra cada año en distintos países, en esa ocasión se realizó en la Universidad de Sheffield, Reino Unido. Sobre el plagio el Dr. Gelbukh comenta: "Es un problema cultural, desde la primaria los alumnos dan copiar y pegar a la tarea, este comportamiento afecta la evaluación y por tanto a la educación en todos los niveles". Actualmente, sólo ocho años después, cualquier maestro de primaria o secundaria puede descargar alguna aplicación para verificar que sus estudiantes no hayan plagiado la tarea.

Dr. Miguel Ángel Sánchez Pérez.

Detección de discurso de odio, el Dr. Segun Taofeek Aroyehun trabajó asesorado por el Dr. Gelbukh en este tema y obtuvieron el **Premio de Investigación para América Latina 2020 que otorga cada año Google** y el **Premio de Doctorado 2020 para América Latina** que concede anualmente **Microsoft**.

Dr. Segun Taofeek Aroyeun.

Ese mismo año, otro de sus alumnos, Navonil Majumder obtuvo el **Premio a la Mejor Tesis de Posgrado del IPN**, su tesis aborda el ***análisis de sentimientos en redes sociales***.

Actualmente su alumno de doctorado, Maaz Amjad, aborda el tema de *l**a detección automática de noticias falsas***, este trabajo es codirigido por el Dr. Grigori Sidorov y la Dra. Alisa Zhila, otra exalumna del Dr. Gelbukh y del CIC IPN que ha trabajado en áreas de inteligencia artificial de empresas como Yahoo, Microsoft e IBM y en 2012 obtuvo el reconocimiento **Miscrosoft Research Latin America Fellowship Program 2012–2013**.

Sobre los innumerables datos que ofrecen las redes sociales el Dr. Gelbukh señala: “De los datos y las opiniones que deja la gente en internet se puede obtener información muy importante, para que las personas puedan influir en las decisiones, en cómo una empresa desarrolla un producto o el gobierno se entere de qué necesitan, pero para eso tenemos que convertir un montón de textos en datos estructurados, entender, de forma automática, de qué se trata cada comentario, de qué producto se habla, de qué marca (esto se llama análisis de aspectos), de qué lugar fue enviado el mensaje, quién lo hizo, es hombre o mujer, de qué edad y sobre todo, entender automáticamente si el comentario es positivo o negativo, lo cual no es trivial y además, determinar automáticamente si el comentario es verdadero o es comprado, todo sirve para hacer una correlación y obtener información muy útil para la toma de decisiones bien informadas, tanto en la industria, como en la política y el gobierno, lo cual es la democracia en el sentido original de la palabra «influencia de la gente en la toma de decisiones»”.

Dr. Navonil Majumder, Premio a la Mejor Tesis de Posgrado del IPN, 2020.

Estudiante de doctorado, Maaz Amjad.

Dra. Alisa Zhila.

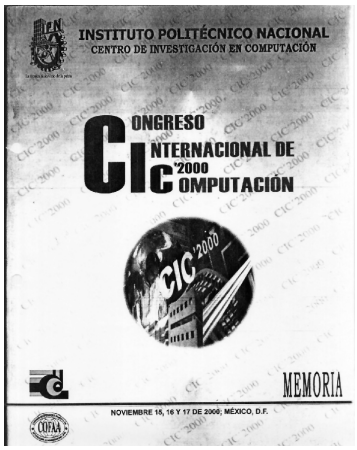
Laboratorio de Procesamiento de Lenguaje Natural del CIC IPN: Una ventana al mundo, a un mundo más grande

En el laboratorio de Procesamiento de Lenguaje Natural del CIC IPN han cursado el posgrado alumnos de Bangladesh, Colombia, Cuba, India, Irán, Nigeria, Paraguay, Paquistán, Estados Unidos, Rusia, Ucrania, Vietnam y México, sobre este interés internacional y los resultados que ellos obtienen el Dr. Gelbukh comenta: “El interés de estudiantes de diferentes países es porque somos un muy buen grupo de investigación, he colaborado con varios equipos, de las mejores universidades de los países del Primer Mundo y mi impresión es que no somos menos que ellos, somos muy buenos a nivel internacional, publicamos bien, hacemos cosas interesantes, ganamos premios importantes, organizamos congresos internacionales y nos conocen. Tenemos en el Laboratorio un ambiente enriquecedor con algunos de los mejores alumnos de todas partes del mundo, ellos hablan entre sí en inglés, así cuando van a un congreso o a una estancia de investigación ya no tienen problema con el idioma”.

Procesamiento de Lenguaje Natural

Un aspecto que ha contribuido a la internacionalización del Laboratorio de Procesamiento de Lenguaje Natural es que además de asistir a congresos, en el año 2000 organizaron por primera ocasión el CIC-Ling (Conference on Intelligent Text Processing and Computational Linguistics) el cual se realizó en ocho ocasiones en México, pero se ha consolidado en el plano internacional y se ha llevado a cabo en Israel, Rumania, Japón, India, Grecia, Nepal, Egipto y Turquía, se espera que la próxima edición se realice nuevamente en India.

Entre congresos, estancias e invitaciones a dirigir o evaluar estudiantes de otras naciones, el Dr. Gelbukh ha viajado a 60 países, actualmente, por ejemplo, dirige las tesis de varias alumnas en Irán. Sobre la relación con los alumnos comenta: “He graduado a 36 doctores en ciencias, es una gran satisfacción ver que mis alumnos son exitosos, que les hemos abierto la puerta a un mejor futuro y que ellos, a su vez, están abriendo la puerta a un mejor futuro a otra gente”.



Memoria-Congreso Internacional de Computación, 2000.

Para el Dr. Gelbukh la inversión del gobierno mexicano en las becas a estudiantes que realizan una carrera científica es clave: “Es muy bueno porque eso es el futuro del país, la ciencia es conocimiento, es el camino hacia la educación, hacia la cultura y hacia una mejor vida para todos. Me gusta mucho la ciencia por la libertad, yo hago lo que me gusta hacer, no lo que el jefe me diga y por eso lo hago bien, al límite de mis capacidades para la satisfacción mía, no la del jefe”.

El Dr. Alexander Gelbukh recibe constantemente invitaciones para eventos y asesorías en otros países y aunque, acude a los que más le es posible, hizo de México y del CIC IPN el lugar para desarrollar su pasión por la ciencia: “En México formamos un buen grupo, era difícil pensar ya en ir a otro lado y volver a empezar, además, con el programa de becas he podido tener buenos estudiantes de todo el mundo. El CIC IPN es mi oportunidad de mejorar el mundo y también es una ventana al mundo, yo me siento como en otra galaxia, mientras vivía en la URSS, no me imaginaba que algún día iba a poder hablar con alguien de Bangladesh o de Irán, o de Colombia o de Paraguay o de Nigeria”.

Finalmente, a pregunta expresa «¿Cómo es México para usted?» su respuesta fue breve pero cabal: “Lindo y querido”.

Dr. Gelbukh con sus alumnos de laboratorio, 2022.

> “
>
> **He graduado a 36 doctores en ciencias, es una gran satisfacción ver que mis alumnos son exitosos.**

Dr. Gelbukh ofrece palabras en el CIC-Ling, 2002.

El Dr. Luis Villa y el Dr. Gelbukh durante una conferencia en la Semana del Emprendedor, 2016.

Galería

Dr. Gelbukh durante el Simposium Internacional de Computación, 1998.

El Dr. Gelbukh durante la expo LEI, realizada en el CIC, 2016.

El Dr. Gelbukh con su alumno, el Dr. Segun Taofeek Aroyehun en el CIC.

El Dr. Gelbukh recibiendo la presea Lázaro Cárdenas por desempeño académico en maestría, en representación de Navonil Majumder, 2018.

Dra. Helena Monserrat Gómez Adorno, alumna del Dr. Gelbukh y presea Lázaro Cárdenas por mejor desempeño académico, 2019.

El Dr. Gelbukh y Dr. Francisco Hiram Calvo Castro, director interino del CIC, durante una entrevista, 2022.

XXV
Aniversario
CIC • IPN