



Universidad Tecnológica Nacional  
Rectorado  
Secretaría de Ciencia, Tecnología y Posgrado

SISTEMA DE INFORMACION DE CIENCIA Y  
TECNOLOGIA (SICyT)

FORMULARIO PARA PROYECTOS DE INVESTIGACIÓN Y DESARROLLO

Código del Proyecto: MSUTIME0005354TC

**1. Unidad Científico-Tecnológica**

FR Mendoza - LABORATORIO DHARMA

**2. Denominación del PID**

Calibración de sensores de bajo costo para monitoreo de calidad del aire mediante aprendizaje de máquinas

**3. Resumen Técnico del PID**

El presente proyecto propone investigar el uso de técnicas novedosas de aprendizaje de máquinas para calibrar sensores de calidad de aire de bajo costo. Por un lado, actualmente la tecnología para monitorear calidad de aire ha evolucionado notoriamente, brindando equipos compactos, fáciles de utilizar, y de precios muy accesibles. Esto permite mejorar las posibilidades de la sociedad para evaluar su exposición al aire contaminado. Sin embargo, aún es necesario mejorar la utilidad de los datos generados por dichos sensores, que es de muy baja confiabilidad. Hoy en día, la calidad del aire es uno de los principales causantes de problemas de salud, y por ello debería ser un tema de prioridad para la sociedad. La contaminación del aire está asociada con diversas enfermedades, síntomas y condiciones que afectan la salud y la calidad de vida. En muchas ciudades del mundo se está combatiendo contra este problema. Sin embargo, en la generalidad de los casos las ciudades no cumplen los requerimientos de los estándares de regulación de calidad del aire. Actualmente no existen legislaciones que regulen el derecho de los ciudadanos a acceder a datos de calidad de aire de las estaciones de monitoreo, las cuales son pertenencia de organismos o instituciones privadas. Por esta razón, hoy en día hay una tendencia mundial a incrementar la disponibilidad de dichos datos. Como parte de esta tendencia, se está investigando fuertemente la posibilidad de utilizar sensores de bajo costo para poder medir contaminación del aire. El principal desafío que demanda el uso de estas tecnologías es lograr que los datos generados alcancen la calidad de las directivas de los parlamentos de calidad ambiental (e.g., Air Quality Directive 2008/50/EC). La falta de calidad de los datos generados por estos sensores se debe a variaciones bruscas de las condiciones del ambiente, o a la baja calidad de los sensores electroquímicos que se utilizan. Sin embargo, existen estudios que plantean aplicar métodos de post-procesamiento de datos para calibrar dichos sensores (es decir, corregir las fallas de medición). Dichos métodos de post-procesamiento incluyen técnicas de aprendizaje de máquinas como regresiones múltiples, redes neuronales, entre otras (Spinelle et al., 2015). En esta investigación proponemos trabajar en una librería de software libre para poder llevar a cabo calibración de sensores de bajo costo utilizando aprendizaje de máquinas. Para ello, proponemos investigar los algoritmos que se han utilizado hasta la actualidad e implementarlos. Del mismo modo, añadir nuevos algoritmos y tecnologías de aprendizaje de máquinas, e investigar cuáles de ellas tiene capacidad de mejorar la calibración de sensores de bajo costo para medir la calidad de aire.

**4. Programa**

Medio Ambiente, Contingencias y Desarrollo Sustentable

**5. Proyecto**

**Tipo de Proyecto:** UTN (PID UTN) CON INCORPORACION EN PROGRAMA INCENTIVOS

**Tipo de Actividad:** Investigación Aplicada

**Campos de Aplicación:**

Rubro	Descrip. Actividad	Otra (especificada)
SALUD HUMANA (Desarrollo, protección y mejoramiento)	Preservación de la atmósfera	

**Disciplinas Científicas:**

Rubro	Disciplina Científica	Otras Disciplinas Científicas
INGENIERÍA EN COMUNICACIONES ELECTRÓNICA Y CONTROL	Electrónica	-
CIENCIAS DE LA COMPUTACIÓN E INFORMÁTICA	Inteligencia Artificial	-

**Palabras Clave**

sensores bajo costo monitoreo calidad aire aprendizaje máquinas

**6. Fechas de realización**

Inicio	Fin	Duración	Fecha de Homologación
01/01/2019	31/12/2020	24 meses	-

**7. Aprobación/ Acreditación / Homologación / Reconocimiento (para ser completado por la SCTyP - Rectorado)****7.1 Aprobación / Acreditación / Reconocimiento (para ser completado por la FR cuando se posea N° Resolución)**

N° de Resolución de aprobación de la FR:

**7.2 Homologación (para ser completado por la SCTyP - Rectorado)**

Código SCTyP : MSUTIME0005354TC Disposición SCTyP:

Código Ministerio:

**8. Estado (para ser completado por la SCTyP - Rectorado)**

EN TRÁMITE

**9. Aavales (presentación obligatoria de aavales)****10. Personal Científico Tecnológico que participa en el PID**

Apellido y Nombre	Cargo	Hs/Sem	Fecha Alta	Fecha Baja	Otros Cargos
BROMBERG, FACUNDO	INVESTIGADOR FORMADO	10	01/01/2019	31/12/2020	-
PEREZ, DIEGO SEBASTIÁN	INVESTIGADOR DE APOYO	10	01/01/2019	31/12/2020	-
SCHLUTER, FEDERICO	DIRECTOR	10	01/01/2019	31/12/2020	
CREMADES, PABLO	INVESTIGADOR FORMADO	6	01/01/2019	31/12/2020	-
CASTRO VIDELA, FERNANDO HORACIO	CO-DIRECTOR	10	01/01/2019	31/12/2020	
MONGE BOSDARI, DAVID ANTONIO	INVESTIGADOR FORMADO	10	01/01/2019	31/12/2020	-
BARI, SANTIAGO	BECARIO ALUMNO FAC.REG.	10	01/01/2019	31/12/2020	-
BARBEITO, MARÍA ESTHER	INVESTIGADOR FORMADO	6	01/01/2019	31/12/2020	-
GEL, ANA PAULA	BECARIO ALUMNO FAC.REG.	10	01/01/2019	31/12/2020	-

**11. Datos de la investigación****Estado actual de concimiento del tema***A nivel global*

Los humanos manufacturamos el aire que respiramos de manera creciente, especialmente en las ciudades (Fenger, 2009). Como ciudadanos, vivimos inmersos en una atmósfera compleja y nociva que es el producto de un frenético metabolismo urbano y representa en la actualidad una de las mayores amenazas hacia la salud a nivel mundial (WHO, 2016).

El monitoreo atmosférico ha sido clave en el descubrimiento, caracterización y actualización de los distintos fenómenos asociados a la contaminación del aire (Finlayson-Pitts, 2010) y es un aspecto fundamental de todos los estándares de calidad del aire establecidos para mitigar sus efectos (Jacobson, 2002). En general, se considera una práctica normal que sean solo los expertos al servicio de un organismo científico o ambiental público quienes tengan a su cargo la operación de costosas y complejas redes de monitoreo para generar y analizar datos acerca del estado del aire que respiramos. Además, sus informes han guiado, y convalidado, la planificación y el accionar de los tomadores de decisión encargados de intervenir en la regulación y la gestión de los distintos aspectos que afectan tanto la calidad del aire como la calidad de vida urbana (ie. transporte, salud, usos del suelo).

Sin embargo, esta práctica de monitoreo ambiental y asesoría experta parece estar en crisis y en transición hacia un nuevo paradigma más abierto e inclusivo, no solo por sus limitaciones técnicas (baja resolución espacial, altos costos, etc) (Mead et al., 2013; Kumar et al., 2015) y las objeciones al rol público de esta ciencia normal (Funtowicz & Ravetz, 2003; Guimaraes et al, 2009; Albagi, 2015), sino también por las posibilidades derivadas de la masificación de las tecnologías de la información en general y de la reciente proliferación de sensores de bajo costo (Snyder et al., 2013). Amateurs, makers, hackers, comunidades de afectadas y ciudadanos científicos están utilizando y compartiendo los diseños de equipos científicos abiertos DIY (do it yourself) cambiando no sólo cómo y por quien son obtenidos los datos, sino también por qué, para qué y cómo los mismos están siendo presentados, y accionados, en la esfera pública (Dye, n.d.). Algunos ejemplos de esta tendencia se pueden encontrar en el desarrollo de equipos como: AirQualityEgg, PublicLab, SmartCitizen, AirCasting.

Este cambio de paradigma presenta nuevos desafíos y oportunidades para el monitoreo de la calidad del aire. Por un lado, existe cierto optimismo en cuanto a las oportunidades técnicas, y políticas, abiertas por este tipo de monitores de bajo costo y la mayor participación ciudadana, pero por otro lado, existe una fuerte preocupación acerca de la calidad de los datos generados. En este sentido se pueden mencionar los esfuerzos de las agencias ambientales y la academia por anticipar, incorporar, acompañar y dirigir los cambios producidos por esta ciencia comunitaria y ciudadana en sus propios roles y actividades (Spinelle et al., 2013; EPA, (n.d.); CAP, 2015; WHO, 2017). En la actualidad, existe consenso en que este tipo de mediciones ciudadanas/comunitarias no pueden reemplazar las regulatorias sino que serían complementarias (Clements et al., 2017, Rai et al., 2017). Sin embargo, se hace énfasis en que para distintos tipos de datos pueden existir diferentes tipo de uso y que, a pesar de una menor calidad de los datos, estos pueden ser lo suficientemente buenos para un determinado fin, en especial para las comunidades afectadas (Gabrys et al., 2016). En la legislación europea (Air Quality Directive 2008/50/EC) se reserva la denominación de mediciones indicativas para aquellas que cumplan con unos determinados objetivos de calidad (DQO - data quality objectives) y la calibración y la evaluación de la performance de estos sensores low-cost se ha tornado en un tema central al cual se están dedicando numerosas discusiones, investigaciones, prácticas y contestaciones (EPA, 2018; AQ-SPEC, (n.d.); Gabrys et al., 2018; Castell et al., 2017).

La calibración se entiende como el proceso por el cual la respuesta de los sensores low-cost es comparada y ajustada a los datos de sensores de referencia. Este proceso no es trivial dadas las características de los sensores (no linealidad, interferencia de temperatura y humedad, drift línea base, etc) y es una tendencia creciente aplicar regresiones multivariable, técnicas de aprendizaje de máquinas y otros complejos algoritmos que ayuden a mejorar la calidad de los datos. Usualmente ambos sensores (low-cost y referencia) son co-localizados en un ambiente representativo de las condiciones de muestreo y se utiliza este periodo como entrenamiento para desarrollar el algoritmo de ajuste. Luego, se utiliza este algoritmo para corregir la respuesta del sensor en distintos ambientes. En la actualidad, el uso de aprendizaje de máquinas para calibración de sensores de bajo costo es un enfoque prometedor. La flexibilidad que ofrecen las herramientas de aprendizaje de máquinas hacen que este enfoque resulte atractivo, comparado con técnicas paramétricas que requieren la suposición de modelos paramétricos estáticos. Este tipo de técnicas poseen la habilidad de aprender relaciones entre las variables de interés de un problema, utilizando un conjunto de datos de entrenamiento. Respecto del uso de este tipo de técnicas para la calibración de sensores de calidad de aire de bajo costo, hasta la fecha se han publicado diversos estudios en los que se utilizan modelos aprendidos con diversas técnicas, como regresiones múltiples, redes neuronales, entre otras (Esposito et al., 2016; Spinelle et al., 2015, 2017, De Vito et al., 2008, 2009).

En las comunidades de práctica, tanto expertas como ciudadanas, existe el consenso de que es necesario poner en común capacidades y conocimientos para afrontar y resolver las limitaciones de los sensores de bajo costo y poder avanzar en acciones concretas, guiadas por los datos, para mejorar la salud pública en general. En este sentido, es necesario empoderar a los ciudadanos haciendo disponibles no solo las herramientas de sensado como ya ocurre (DIY/open source scientific hardware), sino también atender a sus limitaciones en cuanto al manejo de los datos, su análisis y su visualización (Clements et al., 2017). Herramientas abiertas de visualización como OpenAir (Carlslaw y Ropkins, 2012) y plataformas que permiten almacenar, descargar y visualizar datos como Luftdaten y Hackair son fundamentales para construir sentido (making sense) a partir de los datos y poder integrarlos en las historias particulares y los fines de los distintos usuarios (ie. data stories) (Gabrys et al., 2016). Poder evaluar, mejorar y expresar la calidad de los datos debería ser parte indispensable en la construcción de estos relatos con datos. Sin embargo, a pesar de la gran cantidad de publicaciones respecto a las calibraciones, los métodos y algoritmos de ajuste no son open source y por el contrario, parecen ser el nuevo foco de creación de valor de los sensores low-cost y en algunos casos ya están siendo patentados y comercializados (Zimmerman et al., 2018; Agilaire, n.d.).

### **A nivel local**

En Latinoamérica existe una falta crítica de datos de calidad del aire con algunas contadas excepciones en México, Brasil, Colombia y Chile (CAI, 2013; WB, 2017). Tanto en Argentina como otros países de Latam se han establecido estándares de calidad del aire, en su mayor parte copiados/adaptados de leyes y guías extranjeras (WHO, EPA), sin embargo, los protocolos, técnicas e instrumentos de monitoreo con los que dichas normas son evaluadas no han podido ser adaptados ni apropiados por los organismos de control locales debido a sus altos costos y a las dificultades técnicas que plantea su puesta en funcionamiento, resultando en un monitoreo de la calidad del aire históricamente insuficiente, y en muchos casos inexistente. En ciudades como Córdoba y Mendoza se han hecho esfuerzos para adquirir sofisticado instrumental de monitoreo automático extranjero, sin embargo, la falta de previsión respecto a los requerimientos técnicos y económicos de operación, mantenimiento y, especialmente, de calibración ha hecho que estos equipos no se hayan usado de manera sistemática. Aun cuando los repuestos y los gases patrón puedan ser conseguidos en el extranjero, la experiencia técnica local en la reparación y calibración es, no solo costosa, sino deficiente.

Un contexto donde la cuestión ambiental es relegada sumado a las dificultades para acceder a la información pública y las características inherentes de la contaminación del aire urbano (ie. invisible, interterritorial, etc) refuerzan y cierran el círculo vicioso por el que la calidad del aire es mantenida ausente en la esfera pública. Solo en casos puntuales donde ha habido una clara definición de contaminadores y afectados la contaminación del aire ha tomado estado público a partir de un conflicto y ha tenido efectos territoriales, jurídicos e institucionales (Murgida, 2014; Merlinsky, 2014). Algunos ejemplos de estos conflictos son los de las poblaciones cercanas a los polos petroquímicos de Dock Sud (dentro de la cuenca Matanza-Riachuelo), Bahía Blanca y Luján de Cuyo, y los de la ciudad de Puerto Madryn con la planta de aluminio de Aluar (Puliafito et al., 2013).

La ciencia local, a pesar de su amplia producción sobre contaminación atmosférica (Puliafito et al., 2013), tampoco ha contribuido a revertir la ausencia de esta problemática en foros públicos donde se evalúan y definen las distintas políticas públicas (transporte, salud, usos del suelo) que afectan tanto la calidad del aire como la calidad de vida urbana. A esto debe sumarse la poca llegada de la

producción científica a los ciudadanos que pudieran estar interesados. Estas situaciones a su vez han redundado en un limitado interés para el financiamiento y la ampliación de los estudios en esta temática.

En general, mucha de las acciones públicas en materia de contaminación del aire han estado más orientadas a cumplir con compromisos internacionales que a resolver inquietudes internas (ej. comunicaciones cambio climático) (Murgida, 2014). Este también parece ser el caso de la Red Federal de Monitoreo Ambiental (<https://redfema.ambiente.gob.ar/>) donde el Ministerio de Medio Ambiente ha centralizado los datos de monitoreo de agua, suelo y aire según lo acordado en XIX Reunión del Foro de Ministros de Medio Ambiente de América Latina y el Caribe (<http://www.pnuma.org/forodeministros/19-mexico/documentos.htm>). Sin embargo, el Ministerio solo ha solicitado los datos pero no ha ofrecido a cambio ningún beneficio a quienes los generan más allá de su publicación, ni tampoco ha establecido ningún mecanismo de control de calidad o aseguramiento de la calidad de dichos datos. Solo recientemente, e impulsado por organismos internacionales, se está trabajando con los institutos de normalización (INTI) y los operadores de las estaciones de monitoreo en Argentina, y otros países de Latinoamérica, para evaluar la calidad de los datos y generar capacidades de monitoreo. Esto se da en el marco del Plan de Acción Regional sobre la Contaminación Atmosférica para América Latina y el Caribe (UNEP, 2014).

Ante este panorama, son varios los grupos de ciudadanos y científicos en Latinoamérica los que están buscando en los sensores de bajo costo una respuesta ante la falta de datos de calidad del aire. Es interesante destacar que varias de estas iniciativas son de código abierto y han surgido fuera de las instituciones (ie. hackerspaces) con el objetivo de poder accionar los datos junto a afectados en la esfera pública (Cremades, 2013, ACA, ETER).

Por otro lado, pero en un mismo sentido, las prácticas de ciencia abierta están ganando espacio y consideración entre grupos e instituciones científicas también en Argentina. Diversos grupos científicos locales han comenzado a tomar nota del atractivo y los posibles beneficios de la ciencia abierta. Este enfoque promueve una mayor eficiencia y calidad en la producción de conocimiento científico y aumenta la democratización de la información, ofreciendo además una mejor respuesta a las demandas sociales de conocimiento. Al mismo tiempo, estas primeras experiencias de ciencia abierta están enfrentando ciertos problemas y barreras que se relacionan con la relativa novedad de sus prácticas, el desconocimiento de los pares y/o algunos organismos de financiamiento o evaluación y la necesidad de generar nuevas métricas de evaluación (Arza et al. 2015).

## Referencias

1. ACA, Agentes Calidad del Aire. [http://wiki.unloquer.org/personas/brolin/proyectos/agentes\\_calidad\\_aire](http://wiki.unloquer.org/personas/brolin/proyectos/agentes_calidad_aire), (Visitada en Junio 2018)
2. Agilaire, <http://agilaire.com/pdfs/ANT.pdf>, (Visitada en Junio 2018)
3. Air quality egg, <http://airqualityegg.com/>, (Visitada en Junio de 2018)
4. Aircasting, <http://aircasting.org/>, (Visitada en Junio de 2018)
5. Albagli, S., 2015, Open science in question. En Open Science, open issues. Albagli, S., Macie, M.L., Hannud Abdo A. (Eds). IBICT y Unirio. Brasília y Rio de Janeiro. Brasil 292. ISBN 978-85-7013- 111-9. <http://livroaberto.ibict.br/handle/1/1061>
6. Arza, V., Fressoli, M. 2016. Proyecto: Ciencia abierta en Argentina: experiencias actuales y propuestas para impulsar procesos de apertura. Informe Final CIECTI. <http://stepsamericalatina.org/wp-content/uploads/sites/21/2016/07/Informe-Final-CIECTI.pdf>
7. Clean Air Partnership. 2015. Collaborative Air Quality Monitoring Strategy: Background and Opportunities. Report prepared for: Healthy Public Policy Team, Toronto Public Health. February, 2015.
8. Clements, A.L.; Griswold, W.G.; RS, A.; Johnston, J.E.; Herting, M.M.; Thorson, J.; Collier-Oxandale, A.; Hannigan, M. Low-Cost Air Quality Monitoring Tools: From Research to Practice (A Workshop Summary). *Sensors* 2017, 17, 2478.
9. Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., and Bartonova, A., 2017, Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?, *Environ. Int.*, 99, 293–302, <https://doi.org/10.1016/j.envint.2016.12.007>.
10. Carslaw, D.C., Ropkins, K., 2012, Openair d An R package for air quality data analysis, *Environmental Modelling & Software* 27-28.
11. Clean Air Institute. 2013. La Calidad Del Aire En América Latina: Una Visión Panorámica. [www.cleanairinstitute.org/calidaddelaireamericalatina/cai-report-spanish.pdf](http://www.cleanairinstitute.org/calidaddelaireamericalatina/cai-report-spanish.pdf)
12. Cremades, P., Castro, F., Fernandez, R., Clausen, R., Puliafito, E., (2013), Desarrollo de un monitor abierto de calidad del aire (MACA), VII Encuentro de Investigadores y Docentes de Ingeniería EnIDI, Setiembre 2013, San Rafael, Argentina. EnIDI 2013, Los Reyunos, San Rafael. Mendoza, Argentina
13. Dye, T., <http://tdenviro.com/brief-history-air-quality-sensing/>, (Visitada en Junio de 2018)
14. Esposito, E., De Vito, S., Salvato, M., Bright, V., Jones, R. L., & Popoola, O. (2016). Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. *Sensors and Actuators B: Chemical*, 231, 701-713.
15. ETER, Monitor Abierto de Calidad del Aire, <https://etermonitor.net/>, (Visitada en Junio 2018)
16. EPA. 2018. Workshop <https://www.epa.gov/air-research/workshop-webinar-deliberating-performance-targets-air-quality-sensors>
17. Fenger, J., 2009. Air pollution in the last 50 years – From local to global. *Atmospheric Environment*, 43(1), pp.13–22. <http://linkinghub.elsevier.com/retrieve/pii/S1352231008008960>
18. Finlayson-Pitts, B. J., 2010, Atmospheric Chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), 6566–6567. <http://doi.org/10.1073/pnas.1003038107>

19. Funtowicz, S., Ravetz, J. 2003. Post-normal science. International Society for Ecological Economics (ed.), Online Encyclopedia of Ecological Economics.
20. Gabrys, J, Pritchard, H, Barratt, B (2016) Just good enough data: figuring data citizenships through air pollution sensing and data stories. *Big Data & Society*. 3(2): 1-14.
21. Guimarães Pereira, Â., Raes, F., De Sousa Pedrosa, T., Rosa, P., Brodersen, S., Jørgensen, M.S., Ferreira, F., Querol, X., Rea, J., 2009. Atmospheric composition change research: Time to go postnormal? *Atmospheric Environment* 43, 5423–5432.
22. Hackair, <http://www.hackair.eu/>, (Visitada en Junio 2018)
23. Jacobson, M. 2002. *Atmospheric Pollution: History, Science, and Regulation*- Cambridge University Press. 412 pp. ISBN-10: 0521010446
24. Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R., 2015. The rise of low-cost sensing for managing air pollution in cities. *Environment international* 75, 199–205.
25. Luftdaten, <https://luftdaten.info/>, (Visitada en Junio 2018)
26. Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J. J., McLeod, M. W., Hodgson, T. F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J. R., and Jones, R. L.: The use of electrochemical sensors for monitoring urban air quality in lowcost, high-density networks, *Atmos. Environ.*, 70, 186–203, <https://doi.org/10.1016/j.atmosenv.2012.11.060>, 2013.
27. Merlinsky, G. comp, et al., 2013, *Cartografías del conflicto ambiental en Argentina*. 1a ed. Buenos Aires: Ciccus, 326 p. ISBN 9789876930338
28. Murgida, A., Guebel, C., Natenzon, C., Frasco, L., (2013), *El aire en la agenda pública: el caso de la ciudad autónoma de Buenos Aires*. En *Respuestas urbanas al cambio climático en América Latina*, CEPAL, 160p, Santiago de Chile, Chile.
29. Pritchard, H., Gabrys, J., Houston, L., 2018, *Re-calibrating DIY: Testing digital participation across dust sensors, fry pans and environmental pollution*, *New Media & Society*, <https://doi.org/10.1177/1461444818777473>
30. PTB, 2018, [https://www.ptb.de/tc/fileadmin/ResumenTematico\\_Calidad\\_del\\_Aire.pdf](https://www.ptb.de/tc/fileadmin/ResumenTematico_Calidad_del_Aire.pdf) (Visitada en Junio 2018)
31. Puliafito S.E., Allende, D., Panigatti, C. 2013. *Contaminación atmosférica e hídrica en Argentina : contribuciones de la IV Reunión Anual PROIMCA y II Reunión Anual PRODECA* Puliafito S.E., Allende, D., Panigatti, C (Eds). Buenos Aires : Universidad Tecnológica Nacional. Facultad Regional Mendoza, 2013
32. PublicLab, <https://publiclab.org/>, (Visitada en Junio de 2018)
33. Rai, A.C., Kumar, P., Pilla, F., Skouloudis, A.N., Di Sabatino, S., Ratti, C., Yasar, A., Rickerby, D., 2017. End-user perspective of low-cost sensors for outdoor air pollution monitoring. *Science of The Total Environment* 607-608, 691-705.
34. Spinelle, L., Aleixandre, M., Gerboles, M., European Commission, Joint Research Centre, Institute for Environment and Sustainability, 2013. *Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution*. Publications Office, Luxembourg.
35. Spinelle L., Gerboles M., Villani M.G., Aleixandre M., Bonavitacola F., *Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: ozone and nitrogen dioxide*, *Sens. And Act. B: Chem.*, 215, (2015), pp. 249-257.
36. Smartcitizen, <http://www.smartcitizen.me/es/>, (Visitada en Junio 2018)
37. Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. *The Changing Paradigm of Air Pollution Monitoring*. *Environmental Science & Technology* 47, 11369–11377. doi:10.1021/es4022602
38. UNEP,2014,<http://www.pnuma.org/forodeministros/19-mexico/documentos/Borrador%20Plan%20de%20Accion%20Regional%20Contaminacion%20Atmosferica%20Version%20Final%20301013.pdf> (Visitada en Junio 2018)
39. WHO, 2017, Lewis, A.C., Zellweger, C., Schultz, M.G., Tarasova, O.A., (GAW), R.G.S.A.G., *Technical advice note on lower cost air pollution sensors*. World Meteorological Organization and Global Atmosphere Watch.
40. World Health Organization, 2016, *Ambient air pollution: a global assessment of exposure and burden of disease*. World Health Organization, <http://www.who.int/iris/handle/10665/250141>
41. World Bank, 2017, *Filling the Gaps: Improving Measurement of Air Quality in Developing Countries Workshop*, <http://www.worldbank.org/en/events/2017/07/25/filling-the-gaps-improving-measurement-of-air-quality-in-developing-countries-workshop#2>
42. Zimmerman, N., Presto, A.A., Kumar, S.P.N., Gu, J., Haurlyuk, A., Robinson, E.S., Robinson, A.L., Subramanian, R., 2018. A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.* 11, 291-313.

## Grado de Avance

MACA – Monitores abiertos de calidad del aire

Los MACA son prototipos para el monitoreo de la calidad del aire urbana desarrollados a partir de módulos de hardware abierto. Tienen como misión comenzar a resolver la ausencia de datos públicos y abrir la participación de otros ciudadanos y científicos para intentar colocar la contaminación atmosférica en otra esfera, la pública. Si bien el avance en conseguir estos objetivos no ha sido el esperado, los prototipos han sido utilizados como vehículos de aprendizaje para recorrer la red socio-técnica local y para reflexionar acerca de cómo se mantienen/producen los datos científicos y las infraestructuras que los sustentan, cómo y quienes construyen conocimiento a partir de esos datos y cómo es transmitido, y percibido, ese conocimiento a/en la esfera pública.

La descripción del proyecto y de los prototipos, como así también algunas publicaciones, pueden encontrarse en [www.monitorabierto.wikidot.com](http://www.monitorabierto.wikidot.com). La documentación técnica y los datos obtenidos con el segundo prototipo pueden encontrarse en [https://gitlab.com/nanocastro/Repo\\_maca](https://gitlab.com/nanocastro/Repo_maca). Durante el inicio de 2018 se ha trabajado en la evaluación de campo (field testing) de los prototipos a partir de los datos obtenidos en una campaña de medición donde se co-localizaron junto a equipo de referencia. Al comenzar el análisis de datos, y revisando la bibliografía al respecto, surgió la inquietud y la necesidad de profundizar en la calibración de los sensores low-cost mediante técnicas de aprendizaje de máquinas que permitan minimizar/eliminar interferencias de temperatura, humedad e interferencias cruzadas de otros gases. Los resultados parciales de la investigación en curso pueden consultarse en [http://rpubs.com/dmonge/calibracion\\_maca](http://rpubs.com/dmonge/calibracion_maca).

Actualmente estamos pasando a una segunda etapa de evaluación con el prototipo dividido en dos, uno para el monitoreo de gases y por otro el monitoreo de material particulado (PM). El primero estaría equipado con dos pares de sensores de óxidos metálicos MICS (e2v) para monóxido de carbono (CO), dióxido de nitrógeno (NO<sub>2</sub>) y ozono (O<sub>3</sub>). El segundo contaría con tres sensores distintos para PM, el Plantower 7003, el NOVA SDS011 y también el Shinyei PPD42.

La idea es aprovechar un Proficiency Test (PT) a nivel Latinoamérica del que participa el Laboratorio de Análisis Instrumental de la Facultad de Ingeniería de la UN Cuyo para poner a medir estos dos prototipos junto a los equipos de referencias. Además, se sumarían réplicas de monitores ya existentes que utilicen los mismos sensores de PM (Luftdaten y Dustduino) y se sumaría también un prototipo de ETER. Para el análisis de los datos en esta nueva etapa se han incluido en el proyecto los laboratorios Dharma (UTN - FRM) e ITIC (FCEN - UN Cuyo).

## Objetivos de la investigación

El objetivo general de esta investigación consiste en mejorar las posibilidades actuales de utilización de sensores de bajo costo para poder medir contaminación del aire. Concretamente, se buscará contribuir desde dos objetivos específicos:

### Objetivo específico 1.

Desarrollo de una herramienta de software libre para visualización, almacenamiento y análisis de datos obtenidos por sensores de bajo costo. Dicha herramienta debe incluir funcionalidades y algoritmos de aprendizaje de máquinas para llevar a cabo la calibración de los datos obtenidos por los sensores. Se incluirá código abierto de algoritmos que aparecen actualmente en el estado del arte, pero que no están accesibles por licencias privativas. Asimismo, también se incluirá el código de los métodos que se desarrollen durante la ejecución del proyecto.

### Objetivo específico 2.

Mejorar la calibración de los sensores de bajo costo. Lo que se intenta lograr es que los datos generados alcancen la calidad de las directivas de los parlamentos de calidad ambiental. Para esto, se investigará el desempeño de los modelos aprendidos por diversos métodos de aprendizaje de máquinas, como por ejemplo regresiones múltiples, redes neuronales, random forest, modelos probabilísticos gráficos (redes Bayesianas, o redes de Markov), entre otros.

## Descripción de la metodología

En general, la metodología a aplicar en el desarrollo de la investigación será iterativa. Partiremos de revisiones bibliográficas y exploración teórica, para luego realizar campañas de medición de datos. Posteriormente, se implementará y se correrán experimentos con algoritmos de aprendizaje de máquinas sobre los datos de calidad de aire obtenidos durante las campañas de medición de datos. Se analizarán los resultados obtenidos por diversos algoritmos del estado del arte con fines comparativos, evaluando distintas métricas relacionadas con la calidad y eficiencia de los mismos. Por último, se trabajará en la preparación de publicaciones de las contribuciones alcanzadas.

Más específicamente, se utilizará la metodología de desarrollo de modelos de aprendizaje automático. Para ellos, proponemos las siguientes etapas de trabajo:

### - Construcción de equipos de monitoreo

Se construirán equipos de monitoreo con sensores de bajo costo basados en los prototipos actuales de MACA. Esta tarea estará a cargo de los integrantes del equipo expertos en este tipo de equipamientos (Fernando Castro Videla, Pablo Cremades).

### - Campañas de medición para obtención de datos

Se programarán campañas de medición de datos en tiempos y lugares representativos, a determinar durante la investigación. Durante dichas campañas se medirá por un tiempo acotado la calidad de aire utilizando en paralelo los prototipos construidos, junto a equipos de referencia (datos confiables). Contamos con la disponibilidad de equipos de monitoreo de referencia mediante el Laboratorio de Análisis Instrumental de la Facultad de Ingeniería de la Universidad Nacional de Cuyo (<http://ingenieria.uncuyo.edu.ar/laboratorio-de-analisis-instrumental>). La directora de dicho laboratorio, la Lic. María Esther Barbeito, es parte del equipo de trabajo de este proyecto.

### - Armado de datasets para entrenamiento y aprendizaje de modelos

Se crearán los datasets con los cuales se aprenderán modelos para calibrar los sensores de bajo costo. Los datasets se generarán en un formato de series de tiempo para monitoreo y análisis (ver InfluxDB, <https://www.influxdata.com/>). Dichos datasets serán realizados respetando estándares conocidos, como por ejemplo el AQDC (Air Quality Data Commons) del grupo ASW (Air Sensor Workgroup) del Fondo de Defensa del Ambiente (<https://www.edf.org/health/air-sensor-workgroup>). Dichos datasets también se publicarán para libre acceso de quienes estén trabajando en calibración de sensores de bajo costo.

**-Diseño de experimentos.**

Se seleccionarán las variables de interés a tener en cuenta para incluir en los modelos que se aprenderán desde los datos. Posteriormente, se analizará qué algoritmos de aprendizaje de máquinas incluir y comparar en los experimentos para mejorar la calibración de los sensores. Algunas herramientas a tener en cuenta serán: regresiones múltiples, random forests, redes neuronales y redes de Bayes.

**-Programación y ejecución de experimentos.**

Se utilizará el lenguaje R, por su versatilidad y sencillez junto con el entorno de programación R-Studio. Se utilizarán también algunas librerías de aprendizaje de máquinas disponibles en R, como por ejemplo *corelearn* para regresiones múltiples (<http://cran.r-project.org/web/packages/CORElearn/>), *nnet* para aprendizaje de redes neuronales (<https://cran.r-project.org/web/packages/nnet/>), *randomForest* para clasificación y regresión con random forests (<https://cran.r-project.org/web/packages/randomForest/>), y *bnlearn* para aprendizaje de redes de Bayes (<http://www.bnlearn.com/>). Para la corrida de experimentos sistemáticos con dichos algoritmos y los datos obtenidos en las campañas de medición, se requerirá de acceso a un cluster de cómputo, que obtendremos y pondremos a punto también durante la ejecución del proyecto.

**-Análisis de resultados**

El desempeño de los modelos aprendidos utilizando los algoritmos de aprendizaje de máquinas seleccionados será medido con diversas medidas comúnmente utilizadas, como por ejemplo 'accuracy', 'precision', y 'recall'. Se compararán los resultados obtenidos con los de las herramientas disponibles que utilizan software privativo y licencias pagas para su uso, como por ejemplo ANT de Agilaire (<http://agilaire.com/advanced-normalization-tool-ant-for-airvision/>).

**-Publicación de software y datos**

Se publicará con licencia GPL el código de las herramientas desarrolladas, junto con los datos obtenidos en las campañas de medición. Se utilizará la plataforma GitLab como repositorio de licencia abierta para publicar el trabajo realizado.

**-Publicación de resultados**

Luego de analizar los resultados obtenidos se llevará a cabo la escritura de un reporte técnico a publicar en revista internacional y/o conferencia.

**12. Contribuciones del Proyecto****Contribuciones al avance científico, tecnológico, transferencia al medio**

El presente proyecto promete generar una transferencia tecnológica de gran interés para la sociedad en general, estudiando temas de fundamental impacto ambiental, como la contaminación del aire. La investigación aplicada propuesta en este proyecto será publicada de modo tal que facilite y promueva el uso de sensores de bajo costo para ayudar a incrementar la disponibilidad de datos sobre la calidad de aire en la región.

**Contribuciones a la formación de Recursos Humanos**

El grupo conformado en este proyecto tiene integrantes en su mayoría de los laboratorios DHARMA (Dpto de sistemas), y LabFD (Dpto de Electronica). Este proyecto incluye investigadores formados y también incluye becarios alumnos. En el desarrollo del proyecto, considerando las tecnologías y herramientas que se utilizarán para el mismo, se realizarán talleres y mesas redondas, cuyo objetivo es la formación, intercambio y consolidación de las experiencias profesionales científicas y tecnológicas, de todos los recursos humanos involucrados.

Los becarios alumnos podrán capacitarse en las bases teóricas y las herramientas para la aplicación de algoritmos de aprendizaje de máquinas a problemas como el presentado en esta propuesta. Además, durante el proyecto tendrán la posibilidad familiarizarse con el concepto, y las herramientas, de IoT mediante la fabricación de, y la interacción con, los prototipos de monitoreo.

**13. Cronograma de Actividades**

Año	Actividad	Inicio	Duración	Fin
1	Construcción de equipos de medición de calidad de aire con sensores de bajo costo	01/01/2019	5 meses	31/05/2019
1	Realizar campañas de medición de calidad de aire. Obtención de datos y estandarización.	01/01/2019	6 meses	30/06/2019
1	Revisión bibliográfica del estado del arte	01/01/2019	4 meses	30/04/2019
1	Análisis y discusión de técnicas de aprendizaje de máquinas, y adaptabilidad a la calibración de sensores de bajo costo	01/03/2019	2 meses	30/04/2019
1	Diseño de herramienta de software libre para calibración de sensores de bajo costo	01/03/2019	2 meses	30/04/2019
1	Implementación de algoritmos de aprendizaje de máquinas.	01/05/2019	5 meses	30/09/2019
1	Publicación de herramienta de acceso libre para calibración de sensores de bajo costo	01/06/2019	2 meses	31/07/2019
1	Diseño de experimentos de calibración con aprendizaje de máquinas	01/11/2019	2 meses	31/12/2019
2	Implementación de experimentos de calibración con aprendizaje de máquinas	01/01/2020	5 meses	31/05/2020
2	Corrida de experimentos de calibración de sensores con aprendizaje de máquinas	01/04/2020	4 meses	31/07/2020
2	Análisis de resultados. Escribir publicaciones académicas y participación en congresos con los resultados encontrados	01/08/2020	5 meses	31/12/2020

**14. Conexión del grupo de Trabajo con otros grupos de investigación en los últimos cinco años**

Grupo Vinc.	Apellido	Nombre	Cargo	Institución	Ciudad	Objetivos	Descripción
sinc(i)	Milone	Diego	DIRECTOR	Universidad	Santa Fe	Colaborar en la	Co-autoría en

		Humbergo		Nacional del Litoral, Santa Fe		investigación post-doctoral del Dr. Federico Schlüter	trabajo publicado recientemente: Schlüter, F., Strappa, Y., Milone, D. H. , Bromberg, F. "Blankets Joint Posterior score for learning Markov network structures." International Journal of Approximate Reasoning 92 (2018): 295-320.
Intelligent Systems Group (ISG)	Santana Hermida	Roberto	INVESTIGADOR FORMADO	Universidad del País Vasco	SAN SEBASTIAN, ESPAÑA	Colaborar en la investigación post-doctoral del Dr. Federico Schlüter (co-director)	Durante 2016 se realizó una visita al ISG a través de la beca para estadias en el exterior de becarios postdoctorales, realizando una estadia de 3 meses en el país Vasco, de la cual Roberto Santana fue director durante la estadia.
Intelligent Systems Group (ISG)	LOZANO	JOSE ANTONIO	INVESTIGADOR FORMADO	UNIVERSIDAD DEL PAIS VASCO	SAN SEBASTIAN, ESPAÑA	Realizar jornadas de intercambio academico entre laboratorios de investigaciones afines	Se realizaron jornadas de visita por universidades de Europa para generar intercambios academicos, a través de fondos de la SPU
Computational Intelligence Group	LARRAÑAGA	PEDRO	INVESTIGADOR FORMADO	UNIVERSIDAD POLITECNICA DE MADRID	MADRID, ESPAÑA	Realizar jornadas de intercambio academico entre laboratorios de investigaciones afines	Realizar jornadas de intercambio academico entre laboratorios de investigaciones afines
MLDL and Artificial Intelligence Lab (DTAI).	Davis	Jesse	INVESTIGADOR FORMADO	Dept. of Computer Science-Katholieke Universiteit	Leuven, Belgica	Realizar jornadas de intercambio academico entre laboratorios de investigaciones afines	Se realizaron jornadas de visita por universidades de Europa para generar intercambios academicos, a través de fondos de la SPU.

## 15. Presupuesto

**Total Estimado del Proyecto: \$ 0,00**

### 15.1. Recursos Humanos - Inciso 1 e Inciso 5

Primer Año			
Becarios Inciso 5	Cantidad	Pesos	Origen del financiamiento
1. Becario Alumno Fac.Reg.	2	\$ 10800,00	Facultad Regional
2. Becario Alumno UTN-SAE	0	\$ 0,00	-
3. Becario Alumno UTN-SCTyP	0	\$ 0,00	-
4. Becario BINID	0	\$ 0,00	-
5. Becario Posgrado-Doctoral en el país	0	\$ 0,00	-

6. Becario Posgrado Doctoral en el extranjero	0	\$ 0,00	-	-
7. Becario Posgrado - Especialización	0	\$ 0,00	-	-
8. Becario Posgrado - Maestría en el país	0	\$ 0,00	-	-
9. Becario Posgrado - Maestría en el extranjero	0	\$ 0,00	-	-

Docentes Investigadores y Otros - Inciso 1	Cantidad	Pesos
1.Administrativo	0	\$ 0,00
2.CoDirector	1	\$ 47143,56
3.Director	1	\$ 77937,09
4.Investigador de apoyo	1	\$ 71219,00
5.Investigador Formado	4	\$ 136034,04
6.Investigador Tesista	0	\$ 0,00
7.Otras	0	\$ 0,00
8.Técnico de Apoyo	0	\$ 0,00

Totales	Inciso 5	Inciso 1	Total
Primer Año	\$ 10800,00	\$ 332333,69	\$ 343133,69

**Segundo Año**

Becarios Inciso 5	Cantidad	Pesos	Origen del financiamiento	
1. Becario Alumno Fac.Reg.	2	\$ 10800,00	-	-
2. Becario Alumno UTN-SAE	0	\$ 0,00	-	-
3. Becario Alumno UTN-SCTyP	0	\$ 0,00	-	-
4. Becario BINID	0	\$ 0,00	-	-
5. Becario Posgrado-Doctoral en el país	0	\$ 0,00	-	-
6. Becario Posgrado Doctoral en el extranjero	0	\$ 0,00	-	-
7. Becario Posgrado - Especialización	0	\$ 0,00	-	-
8. Becario Posgrado - Maestría en el país	0	\$ 0,00	-	-
9. Becario Posgrado - Maestría en el extranjero	0	\$ 0,00	-	-

Docentes Investigadores y Otros - Inciso 1	Cantidad	Pesos
1.Administrativo	0	\$ 0,00
2.CoDirector	1	\$ 47143,56
3.Director	1	\$ 77937,09
4.Investigador de apoyo	1	\$ 71219,00
5.Investigador Formado	4	\$ 136034,04
6.Investigador Tesista	0	\$ 0,00
7.Otras	0	\$ 0,00
8.Técnico de Apoyo	0	\$ 0,00

Totales	Inciso 5	Inciso 1	Total
Segundo Año	\$ 10800,00	\$ 332333,69	\$ 343133,69

TOTAL GENERAL	Inciso 5	Inciso 1	Total General
Todo el Proyecto	\$ 21600,00	\$ 664667,38	\$ 686267,38

**15.2 Bienes de consumo - Inciso 2**

Año del Proyecto	Financiación Anual	Solicitado a
1	\$ 9.900,00	UTN - SCTyP
2	\$ 9.900,00	UTN - SCTyP
<b>Total en Bienes de Consumo</b>		<b>\$ 19.800,00</b>

**15.3 Servicios no personales - Inciso 3**

Año	Descripción	Monto	Solicitado a
1	Viáticos	\$ 6.600,00	UTN - SCTyP
2	Inscripciones a congresos	\$ 6.600,00	UTN - SCTyP
<b>Total en Servicios no personales</b>			<b>\$ 13.200,00</b>

**15.4 Equipos - Inciso 4.3 - Disponible y/o necesario**

Año	Disp/Nec	Origen	Descripción	Modelo	Otras Espec.	Cantidad	Monto Unitario	Solicitado a
1	Necesario	Compra	Equipos para estacion de monitoreo atmosferico	Sensores para material particulado y CO NO2 y O3	sensores y plataforma Arduino	1,00	\$ 18.000,00	UTN - SCTyP
2	Necesario	Compra	Equipo de cómputo	CPU para computo con procesador i7 5820k	-	1,00	\$ 18.000,00	UTN - SCTyP
<b>Total en Equipos</b>							<b>\$ 36.000,00</b>	

**15.5 Bibliografía de colección - Inciso 4.5 - Disponible y/o necesario**

Año	Disp/Nec	Origen	Descripción	Modelo	Otras Espec.	Cantidad	Monto Unitario	Solicitado a
<b>Total en Bibliografía</b>							<b>\$ 0,00</b>	

**15.6 Software - Disponible y/o necesario**

Año	Disp/Nec	Origen	Descripción	Modelo	Otras Espec.	Cantidad	Monto Unitario	Solicitado a
-	-	-	-	-	-	-	-	-
<b>Total en Software</b>							<b>\$ 0,00</b>	

**16. Co-Financiamiento**

Año	RR.HH.	Bienes de Consumo	Equipamiento	Servicios no personales	Bibliografía	Software	Total
1	\$343.133,69	\$9.900,00	\$18.000,00	\$6.600,00	\$0,00	\$0,00	\$377.633,69
2	\$343.133,69	\$9.900,00	\$18.000,00	\$6.600,00	\$0,00	\$0,00	\$377.633,69
<b>Total del Proyecto</b>	<b>\$686.267,38</b>	<b>\$19.800,00</b>	<b>\$36.000,00</b>	<b>\$13.200,00</b>	<b>\$0,00</b>	<b>\$0,00</b>	<b>\$755.267,38</b>

**Financiamiento de la Universidad**

Universidad Tecnológica Nacional - SCyT	\$ 69.000,00
Facultad Regional	\$ 21.600,00

**Financiamiento de Terceros**

Organismos públicos nacionales (CONICET, Agencia, INTI, CONEA, etc.)	\$ 664.667,38
Organismos / Empresas Internacionales / Extranjeros	\$ 0,00
Entidades privadas nacionales (Empresas, Fundaciones, etc.)	\$ 0,00
Otros	\$ 0,00
<b>Total</b>	<b>\$ 755.267,38</b>

**Avales de aprobación, Financiamiento y Otros**

	Orden	Nombre de archivo	Tamaño
<a href="#">Descargar</a>	1	IMG_20180622_112226778_HDR.jpg	1513396

**Currículums (Currículums de los integrantes cargados en el sistema)**

Imprimir

Exportar a PDF