Potential NIMO Host Institution Profile

Institution/Department

Name

The University of New Mexico - Department of Biology

Background

The University of New Mexico (UNM) is prepared to meet the needs of the next LTER Data Center. Currently, the lead PI candidate, Inigo San Gil, is negotiating with the UNM's Vice President for Research for a package aligned with the expectations of the Information Management Committee (IMC). Several formulas are being discussed, and to date, it is still soon to decide which would be the optimal solutions. The following are some details of the discussions.

UNM's Biology department currently hosts all LTER network infrastructures, demonstrating an excellent track record in providing the LTER Network with the services it requires.

UNM continues to recognize the uniqueness of LTER, which currently runs under the CREST center of excellence. UNM is willing to examine the options to provide the best service to the next Data Center, and has plans for a series of negotiating meetings with the current Chair of Biology, the current University Michael Dougher and his associate Kevin Malloy and the incoming VPR. In addition, we will explore the synergies that exist with DataONE (headquarters hosted at UNM) and the revamped University Libraries Digital and Data services.

Principal Investigator Candidate

Name

Inigo San Gil – PI

Current Position

Inigo San Gil - Research Associate Professor

Background

Inigo has a strong record of data driven research in his scientific career. Inigo is also an experienced manager and has a proven track of successes leading LTER network projects. Inigo's record, along with Dr. Mark Servilla support as Co-PI, offers a unique combination to lead the change over the LTER Network Office transition into the information management distributed operations and services. Data expertise, LTER track record and managerial experience required to lead the NIMO project are highlighted below:

Background in Data Sciences:

Inigo earned a Bachelor's degree in Physics specializing in Theoretical Physics and Fluid Dynamics. In 2001, Inigo received a Ph.D in Mechanical Engineering from Yale University. His thesis explored the order in chaos (Fractals in turbulence. His research on turbulence using direct numerical simulations made him an expert in super-computers, information technology as well as a seasoned big data analyst.

Inigo spent a year (98/99) in IBM's Watson Research Center, where he used IBM's Deep Blue clusters to run numerical simulations and data analysis. Later, Inigo conducted research for a year in Los Alamos National Lab (NM) where he ran numerical simulations in the fastest civilian supercomputer in the US around the year 2000. Inigo also used the supercomputing facilities at Cornell U., Maui, Minnesota (UMN) for his research, as well as several clusters for data visualization and big data analysis at Brown University and Yale.

Inigo regularly served as teaching assistant for Yale's popular "*Introduction to Computing for Engineers and Scientists"* Course and "*Applied Numerical Methods*". The teaching experience gave Inigo a chance to interact with the younger computing scientists and engineers willing to apply an array of programming languages to solve real world problems.

Managerial Experience :

Upon finishing his Ph.D, Inigo joined the Yale Core Facility for Bioinformatics, in the Yale School of Medicine. For two years, Inigo participated in the development of Bioinformatic data applications and studies, most noteworthy a time-series study on the gene expression for the fruit fly (Drosophila Melanogaster), whose lifecycle from the

point of view of its 30,000 gene's activities was exposed to the public through a web application Inigo designed using the best technologies available in 2002. (See Science cover, September 2002). Inigo won the recognition and continued support of the Yale School of Medicine as an effective manager for the inter-departmental Core Facility for Bioinformatics. Inigo has also served in the Executive Boards of the US National Phenology Network and the Genomics Standards Consortium, where he made critical contributions to the community approaches to standardization and collaboration.

LTER network wide project successes:

Inigo left his position as manager of the Yale Core Facility for Bioinformatics and moved to Albuquerque in 2004, following his wife's career choices with the Forest Service. Inigo started working at the LTER Network Office in a liaison position with the USGS National Biological Information Infrastructure program. His primary charge: Coordinate the LTER Network standardized metadata creation for inclusion in the NBII metadata clearinghouses in Oak Ridge National Lab. After a successful completion of the standardization project (and the first publication involving all LTER sites), Inigo addressed a recurring challenge for the LTER Network: The lack of satisfactory tools to create, maintain and re-use ecological data and its context. In conjunction with Marshall White, Vanderbilt and Melendez, Inigo launched the Drupal Ecological Information Management System (DEIMS) as a testament of how to work together in the network. This grass-roots project is growing and demonstrating that the LTER can maximize its potential, and work effectively in common tools to manage the core information services provided at the site and network level.

Prior LTER Network Service and Leadership

Inigo's tenure as a liaison agent in the LTER Network Office exposed him to the details of the Information Management Systems for all the LTER sites (he wrote a synopsis of such systems). Inigo showed a unique ability for team building during his interactions with each member of the network, often changing the inertia on projects. Inigo's initiatives often transcended the official channels to expedite the projects. DEIMS and LTER in Github are clear examples when Inigo has taken the initiative when there is indecision or uncharted directions. During the coming transitions, the LTER Network needs informed, experienced and decisive leadership that does not shy of change and has clarity in its goals. Inigo and the IMC team will lead LTER and the broader community where needs to be.

Institutional Infrastructure

Technology

UNM will continue providing the excellent service that has provided for over a dozen years, hosting all the LTER infrastructure and related services.
UNM's Biology department currently hosts all LTER network infrastructures, with a dedicated server room with all the high technology safeguards (hydro sensors, temperature monitors, redundant AC units and dedicated stable power feed with backup generators.

While the IMC explores the transition to cloud services and distributed models available, the UNM will keep providing seamlessly the current service with some changes that will favor the cost of operations. Currently Inigo is conducting negotiations. On the table are UNM willingness to provide in-kind system administration services through is Center for Advanced Research Computing, DataONE co-location of services, reduction of standard F&A (fees, overheads), off-campus locations, and other options.

Non Technology Support Services

UNM provided meeting rooms and specialized training laboratories that will continue to host future data related trainings as it has done in the past.

San Gil Publication List:

Genomic standards consortium projects Field, Dawn; Sterk, Peter; Kottmann, Renzo; De Smet, J Wim; Amaral-Zettler, Linda; Cochrane, Guy; Cole, James R; Davies, Neil; Dawyndt, Peter; Garrity, George M; Standards in genomic sciences. 9,3,599, 2014

Framework for Building Collaborative Research Environment Devarakonda, Ranjeet; Palanisamy, Giriprakash; San Gil, Inigo; International Journal of Recent Contributions from Engineering, Science & IT (iJES) 2,4,pp. 11-15,2014

The Drupal Environmental Information Management System Provides Standardization, Flexibility and a Platform for Collaboration Gries, C; Vanderbilt, K; Reid, D; Melendez-Colom, E; San Gil, I; AGU Fall Meeting Abstracts,1 ,1646,2013

Clarifying concepts and terms in biodiversity informatics Field, D; Amaral-Zettler, L; Cochrane, G; Cole, JR; Dawyndt, P; Garrity, GM; Gilbert, J; Glöckner, FO; Hirschman, L; Karsch-Mzrachi, I; PLoS Biol9,e1001088, 2013

RCN4GSC Workshop Report: Modeling a Testbed for Managing Data at the Interface of Biodiversity and (Meta) Genomics, April 2011 Robbins, Robert J; Cochrane, Guy; Davies, Neil; Dawyndt, Peter; Kottmann, Renzo; Krishtalka, Leonard Kris; Morrison, Norman; Tuama, Éamonn Ó; San Gil, Inigo; Wooley, John; Standards in genomic sciences 7,1,153, 2012

RCN4GSC Meeting Report: Initiating a Testbed for Managing Data at the Interface of Biodiversity and Genomics/Metagenomics, May 2011 Robbins, Robert J; Beach, James; Blum, Stan; Dawyndt, Peter; Deck, John; Kottmann, Renzo; Morrison, Norman; Tuama, Éamonn Ó; San Gil, Inigo; Vieglas, David; Standards in genomic sciences 7,1,171,2012

RCN4GSC Workshop Report: managing data at the interface of biodiversity and (meta) genomics, March 2011 Robbins, Robert J; Amaral-Zettler, Linda; Bik, Holly; Blum, Stan; Edwards, James; Field, Dawn; Garrity, George; Gilbert, Jack A; Kottmann, Renzo; Krishtalka, Leonard; Standards in genomic sciences,7,1,159,2012

Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications Yilmaz, Pelin; Kottmann, Renzo; Field, Dawn; Knight, Rob; Cole, James R; Amaral-Zettler, Linda; Gilbert, Jack A; Karsch-Mizrachi, Ilene; Johnston, Anjanette; Cochrane, Guy; San Gil, Inigo; Wilson,Gareth; Wipat Anil et. Al. Nature biotechnology 29,5,415-420,2011

The genomic standards consortium Field, Dawn; Amaral-Zettler, Linda; Cochrane, Guy; Cole, James R; Dawyndt, Peter; Garrity, George M; Gilbert, Jack; Glöckner, Frank Oliver; Hirschman, Lynette; Karsch-Mizrachi, Ilene; PLoS Biol,9,6,e1001088,2011

Community Page-The Genomic Standards Consortium Field, Dawn; Amaral-Zettler, Linda; Cochrane, Guy; Cole, James R; Dawyndt, Peter; Garrity, George M; Gilbert, Jack; Oliver Glöckner, Frank; Hirschman, Lynette; Karsch-Mizrachi, Ilene; Meyer, Folker; San Gil, Inigo et. Al. PLoS-Biology 9,6,1189,2011

Examples of ecological data synthesis driven by rich metadata, and practical guidelines to use the Ecological Metadata Language specification to this end Inigo San Gil; Kristin Vanderbilt; Steve A. Harrington,; International Journal of Metadata, Semantics and Ontologies 6,1,46-55,2011

Identification of common and cell type specific LXXLL motif EcR cofactors using a bioinformatics refined candidate RNAi screen in Drosophila melanogaster cell lines Davis, Melissa B; SanGil, Inigo; Berry, Grace; Olayokun, Rashidat; Neves, Lori H; BMC developmental biology 11 1 66 2011

Gestión de la información ambiental en los espacios protegidos y en las redes de seguimiento del cambio global Bonet, FJ; San Gil, I; Revista Ecosistemas 19 2 2010

Metadata activities in biology San Gil, Inigo; Hutchison, Vivian; Frame, Mike; Palanisamy, Giri; Journal of Library Metadata 10 3-Feb 99-118 2010

A flexible online metadata editing and management system Aguilar, Raul; Pan, Jerry; Gries, Corinna; San Gil, Inigo; Palanisamy, Giri; Ecological Informatics 5 1 26-31 2010

Case studies of ecological integrative information systems: the Luquillo and Sevilleta information management systems San Gil, Inigo; White, Marshall; Melendez, Eda; Vanderbilt, Kristin; Metadata and Semantic Research 18-35 2010

The “Minimum Information about an ENvironmental Sequence”(MIENS) specification Yilmaz, Pelin; Kottmann, Renzo; Field, Dawn; Knight, Rob; Cole, James R; Amaral-Zettler, Linda; Gilbert, Jack A; Karsch-Mizrachi, Ilene; Johnston, Anjanette; Cochrane, Guy; San Gil, Inigo et al.; 2010

The Long-Term Ecological Research community metadata standardisation project: a progress report San Gil, Inigo; Baker, Karen; Campbell, John; Denny, Ellen G; Vanderbilt, Kristin; Riordan, Brian; Koskela, Rebecca; Downing, Jason; Grabner, Sabine; Melendez, Eda; International Journal of Metadata, Semantics and Ontologies 4 3 141-153 2009

OOS 49-6: Using data clearinghouses to evidence global warming: The value of metadata standards San Gil, Inigo; Servilla, Mark; The 94th ESA Annual Meeting 2009

Meeting report: the fourth Genomic Standards Consortium (GSC) workshop Field, Dawn; Glöckner, Frank Oliver; Garrity, George M; Gray, Tanya; Sterk, Peter; Cochrane, Guy; Vaughan, Robert; Kolker, Eugene; Kottmann, Renzo; Kyrpides, Nikos; OMICS A Journal of Integrative Biology 12 2 101-108 2008

Defining linkages between the GSC and NSF's LTER program: how the Ecological Metadata Language (EML) relates to GCDML and other outcomes San Gil, Inigo; Sheldon, Wade; Schmidt, Tom; Servilla, Mark; Aguilar, Raul; Gries, Corinna; Gray, Tanya; Field, Dawn; Cole, James; Pan, Jerry Yun; OMICS A Journal of Integrative Biology 12 2 151-156 2008

The EcoTrends web portal: an architecture for data discovery and exploration Servilla, Mark; Costa, Duane; Laney, Christine; San Gil, Inigo; Brunt, James; Proceedings of the Environmental Information Management Conference 42288 2008

The minimum information about a genome sequence (MIGS) specification Field, Dawn; Garrity, George; Gray, Tanya; Morrison, Norman; Selengut, Jeremy; Sterk, Peter; Tatusova, Tatiana; Thomson, Nicholas; Allen, Michael J; Angiuoli, Samuel V; Nature biotechnology 26 5 541-547 2008

Meeting report: the fifth Genomic Standards Consortium (GSC) workshop Field, Dawn; Garrity, George M; Sansone, Susanna-Assunta; Sterk, Peter; Gray, Tanya; Kyrpides, Nikos; Hirschman, Lynette; Glöckner, Frank Oliver; Kottmann, Renzo; Angiuoli, Sam; OMICS A Journal of Integrative Biology 12 2 109-113 2008

A standard MIGS/MIMS compliant XML Schema: toward the development of the Genomic Contextual Data Markup Language (GCDML) Kottmann, Renzo; Gray, Tanya; Murphy, Sean; Kagan, Leonid; Kravitz, Saul; Lombardot, Thierry; Field, Dawn; Glöckner, Frank Oliver; OMICS A Journal of Integrative Biology 12 2 115-121 2008

COS 19-8: Defining and assessing data quality in online ecological information systems Brunt, James W; Servilla, Mark S; San Gil, Inigo; Costa, Duane; 2007

LTER IMC community of practice: a learning environment Millerand, Florence; Kaplan, Nicole; Gil, Inigo San; O'Brien, Margaret; Baker, Karen; 2007

PASTA: A Network-level Architecture Design for Automating the Creation of Synthetic Products in the LTER Network Servilla, Mark; Brunt, James; San Gil, Inigo; Costa, Duane; Ecological Informatics 2006

Enabling synthesis of spatially and temporally diverse data collected from environmental sensors within the Long Term Ecological Research Network Vande Castle, JR; Servilla, MS; San Gil, I; Michener, BK; Waide, RB; Scuderi, LA; AGU Fall Meeting Abstracts 1 822 2006

PASTA: A network-level architecture design for generating synthetic data products in the LTER Network Servilla, Mark; Brunt, James; San Gil, Inigo; Costa, D; LTER Databits 2006

Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans Reinke, Valerie; San Gil, Inigo; Ward, Samuel; Kazmer, Keith; Development 131 2 311-323 2004

Fractal character of isoscalar surfaces in shear free turbulence and some effects of shear on the turbulence structure San Gil, Inigo; 2000

Effects of shear on small-scale properties in the Kolmogorov flow San Gil, Inigo; Chen, Shiyi; Sreenivasan, Katepalli; APS Division of Fluid Dynamics Meeting Abstracts 1 1999

The effects of large scales on the inertial range in high-Reynolds-number turbulence Sreenivasan, Katepalli R; Dhruva, Brindesh; San Gil, Inigo; arXiv preprint chao-dyn/9906041 1999

The fractal behavior of isosurfaces of a passive scalar Sreenivasan, Katepalli R; San Gil, Inigo; Chen, Shiyi; APS Division of Fluid Dynamics Meeting Abstracts 1 1997

Asymmetry of velocity increments in fully developed turbulence and the scaling of low-order moments Sreenivasan, KR; Vainshtein, SI; Bhiladvala, R; San Gil, I; Chen, S; Cao, N; Physical review letters 77 8 1488 1996

Instantaneous Wall Shear Stress and Velocity Fluctuations in Turbulent Pipe Flow Bhiladvala, Rustom; San Gil, Inigo; Sreenivasan, KR; APS Division of Fluid Dynamics Meeting Abstracts 1 1996

LTER IMC Community of Practice: A Learning Environment Baker, Karen; Kaplan, Nicole; San Gil, Inigo; O'Brien, Margaret; Millerand, Florence; Yar mey, Lynn;