Inter-Service Data Integration for Geodetic Operations

INDIGO –
User Assessment Report
June 2005
http://indigo.nasa.gov/

The Inter-Service Data Integration for Geodetic Operations (INDIGO) proposal to NASA's Cooperative Action Notice entitled "Earth Science REASoN Research, Education & Applications Solutions Network A Distributed Network of Data and Information Providers for Earth Science Enterprise Science, Applications, and Education" was selected by NASA in July 2003 and funded in July 2004. INDIGO aims to develop data systems to enable geodetic studies that integrate data from multiple space geodetic techniques. To ensure that INDIGO will be responsive to the needs of the user community, an initial stated goal was to assess the current services and user requirements, and prepare a report on INDIGO's plans for the Science Advisory Team. This document serves as that assessment report.

We interviewed 14 researchers from 12 universities and government agencies in 5 countries to learn about investigator needs. The experience level of the interviewees ranged from graduate students to well established principal investigators. Participation and informal conversations at professional meetings and workshops in Fall 2004 also contributed to our understanding of this topic. Conversation-starting questions are presented in appendix A. Although this sample size is small, we began to see commonalities in the responses and feel we grasp what the most pressing needs are. The major threads from our interviews are as follows.

- Investigators had surprisingly few complaints regarding the mechanics of acquiring and preparing data and metadata. Some did mention having had to get used to the procedures when entering the field, and others described brute-force methods which clearly could be simplified, even though they are not perceived as a major hindrance to research.
- We typically have thought of novice users preferring web-based tools for actions like data selection, and expert users preferring low-level protocols (often automated) such as FTP. Interestingly, one student interviewee with somewhat recent memories of learning how to get data and metadata specifically stated that "primitive" FTP access should not be discontinued, for the benefit of new users learning how the data are stored. This remark, which challenges our stereotype of preferences, confirms in any case that both low- and medium- to high- level methods of data discovery and access are valuable to our users and warrant our attention.
- Calibration issues remain troublesome for the space geodetic techniques. Simultaneous use of data from more than one technique can provide better determination of estimable parameters, but development is hampered by the

difficulty in comparing and combining results. Investigators have mentioned some key problem areas:

- Different atmospheric loading models between VLBI and GPS make comparison difficult.
- Insufficient documentation of the specifications and process of formation of the services' products.
- Groups work on different datasets, decreasing the potential gain from comparing the results.
- Not easy to discover spatiotemporally coincident multi-technique datasets. A tool for selecting a geographical area would be valuable. This capability should include regional densification sites and decommissioned sites.
- LEO satellites with multiple instruments offer multiple dynamic models (e.g., altitude or mass) and the investigator may not know how to select the appropriate model.
- Orbits tagged in GPS time can be a nuisance when working with other data in UTC and vice versa.
- Unclear what is the authoritative source for vectors ("local ties") between reference points of different techniques. It is tedious and error prone to assemble tables of vector components by hand.
- Available site information fails to capture "periods of concern" when anomalous data should be avoided or treated specially.
- Products such as EOPs are generated by the different services on varying schedules, necessitating interpolation or extrapolation, causing noise issues when trying to combine or compare.
- Need flexible way to examine time series solutions plot more than one on a graph and vary the local tie between them, and provide the data for statistical analysis.
- Conversions between reference frames are often needed.
- Some additional metadata types are needed: VLBI antenna dimensions and materials, time dependent masses of GPS satellites, weather models yielding the atmosphere along each path from sensor to signal source, shape and reflectivity of multi-technique LEOs, site geology.

Since study and comparison of results calculated by different methods lead to improvements in models and procedures, the field of inter-technique space geodesy will substantially and immediately benefit from systems and processes that enable easy comparison of single and multi-technique results with one another in sensible ways. This will be the specific goal of INDIGO services.

Some obvious actions and whether they can or should be taken on by INDIGO can now be specified:

1. Provide authoritative tables of inter-technique tie vectors. Offer vectors between each pair of reference points, whether measured or calculated. Time dependence must be considered.

Note: The IERS WG2 goals (per-site SINEX with covariance) are the optimal situation. The tabular information discussed here is an interim goal which would have great benefit to multi-technique investigators.
INDIGO? Can probably contribute. Study ITRF's site (contact: Z. Altamimi) and collaborate.

2. Arrange "campaigns" where investigators can study very well-defined multi-technique datasets available from a single location and compare results. Collect copious metadata (software, models, etc) about how the results were formed and whom to contact for discussion.
INDIGO? Can certainly contribute indirectly by improving site metadata.
CDDIS DC is a natural place to collect such information. Coordinate with IERS CPP.

3. Provide upgraded site information, including anomalous periods for each data type.
INDIGO? Yes. Coordinate with IGS Reference Frame Coordinator, Remi Ferland, who has worked on anomalous data period specification for the IGS. The IVS is also working with such information for ITRF2004 submissions. This site information would incorporate the site metadata mentioned in item 2 above.

4. Provide ability to discover suitable spatiotemporally coincident multi-technique data sets.
INDIGO? Yes. Coordinate with GPS Seamless Archive Centers (GSAC) project regarding extending GSAC to other data and product types. GSAC is a proven system for searching and obtaining GPS data and orbits from distributed archives without requiring the user to know access details or even which archives are involved. The GSAC is built upon open source technologies which enable use on a variety of computer platforms and offers INDIGO the possibility to leverage significant previous planning, code development, and testing. The GSAC code would require extension to support SLR and VLBI data and more types of products. Note: JPL policy prohibits public dissemination of JPL-developed software from JPL servers and websites. It is possible under the policy to release JPL enhancements to open source software back to the maintainer, but a New Technology Report must be filed and the Intellectual Property office and Caltech Office of Technology Transfer must approve license terms. INDIGO would be wise to confirm that GSFC, NVI, and JPL contributions to GSAC would be able to be disseminated in a sensible way.

5. Provide upgraded documentation on the formation of products. INDIGO? Yes. Coordinate with each service's analysis coordinators.

Following reconfirmation of the suitability of these goals, the INDIGO team will begin specific design of systems to achieve them. Several interviewees expressed concern that an INDIGO implementation could "break" existing data provision routines. In general, the design principles will include minimizing impact to operational processes and respecting the varying needs of experienced and new users.

Appendix 1. Assessment questions

What geodetic techniques are you using?

What kinds of data, metadata, and products do you need to gather and where do you typically get them?

What kind of preparation do you have to do to use them? (Examples: format translations, reference frame conversions, unit conversions, etc.). If you have had to tediously put things together by hand, please describe.

Is there some data type which you would like to have access to, but you have not found, or have not found in a usable state?

If you have used GPS RINEX data, have you used the GSAC (GPS Seamless Archive Center)? If yes, what has been your experience? How could it be improved? If no, why not?

Do you have a wishlist of capabilities regarding the provision of data, metadata, products and/or information that would make your investigations easier? If so, please note which are top priority.

Can you think of other investigators who we should talk to?

What other questions should we ask?

Please tell us about any experiences or requests you have in this area which have not been mentioned yet.

Appendix 2. Persons interviewed

This list is not exhaustive of everyone whose comments contributed to the conclusions of this report, but demonstrates the breadth of investigators who kindly contributed to the assessment.

Kyohong "Kevin" Choi, University of Colorado, USA
John Ries, Center for Space Research, The University of Texas at Austin, USA
Arthur Niell, Haystack Observatory, Massachusetts Institute of Technology, USA
Seth Gutman, Forecast Systems Laboratory, National Oceanographic and Atmospheric Administration, USA
Jim Ray, National Geodetic Survey, National Oceanographic and Atmospheric Administration, USA
Pascal Willis, Institute Géographique National, France; and Jet Propulsion Laboratory, California Institute of Technology, USA
Erricos Pavlis, Joint Center for Earth Systems Technology, University of Maryland Baltimore County, USA
Tom Herring, Massachusetts Institute of Technology, USA
Richard Gross, Jet Propulsion Laboratory, California Institute of Technology, USA
John Dawson, Geoscience Australia, Australia
Christopher Cox, Raytheon Information Solutions, NASA Goddard Space Flight Center, USA
Per Helge Andersen, Forsvarets Forskningsinstitutt (FFI) and Institute of Theoretical Astrophysics, University of Oslo, Norway
Zuheir Altamimi, Institute Géographique National, France
Zinovy Malkin, Institute of Applied Astronomy, RAS, Russia
SungPil Yoon, Center for Space Research, The University of Texas at Austin, USA