

# REPORT OF THE CHRONOS TOOLBOX WORKSHOP

October 27-28, 2003

San Diego

The workshop was charged to advise CHRONOS concerning the acquisition and implementation of software tools for its time-stratigraphic database services. The workshop was primarily concerned with user-end tools for data visualization, analysis, and interpretation. The participants (Appendix I) were drawn from the community of stratigraphic software developers. Management tools for data input, annotation, filtering, and retrieval were largely ignored.

This report attempts to summarize the issues raised in round table discussions, individual presentations, and informal conversations during the two days of the workshop, plus ideas exchanged after the meeting. After a statement of the general insights gained from the workshop, the report will develop the specific underpinnings of those insights in four parts: an extract of the CHRONOS mission and its charge to the workshop; a detailed catalog of the kinds of tools that CHRONOS should consider, with discussion of the options and issues pertaining to their implementation; issues concerning the interaction with developers; and specific recommendations for implementation and prioritization of the toolbox. Parts 1 and 4 together serve as an executive summary; parts 2 and 3 have more of the exhaustive character of a reference manual.

## 1. SUMMARY OF THE MAIN EMERGENT THEMES AND INSIGHTS

- **Most academic programmers want to share their software.**

Soliciting data and tools will be very different exercises, with hurdles in different places. Donors of data may need to be weaned from a sense of ownership and a fear of being exploited. But developers derive intellectual satisfaction from seeing their programs in use. Thus, for CHRONOS the main hurdle is not to encourage donation of tools but to manage the implementation and updating. Software tools, in a variety of languages and operating systems may be more difficult to unify and distribute than data. Databases improve by continual growth within the same structure; programs improve by continual modification in response to users' requests and ideas. A small CHRONOS staff might have sufficient expertise to manage all the component data structures. For the tools, however, the implementation will need to allow individual developers to update their tools after donation.

- **A question-driven toolbox may be a best initial investment.**

CHRONOS needs to start attracting favorable attention by providing valuable services in less time than it may take for the component data bases to grow to a critical mass. This can be achieved by focusing on answers to frequent and popular questions. Some of these can be answered by tools that use specific flat-file information or small custom data bases.

- **The data-driven toolbox may usefully be developed ahead of the databases.**

The offering of data-based tools need not wait for the growth of data bases. CHRONOS can become an attractive one-stop-shopping source for time-stratigraphic shareware. A "*CHRONOS ENABLED!*" seal would imply some level of approval by the developer community and the promise of future compatibility with CHRONOS data structures and converters. The shareware offerings should be supplemented with tutorial data sets that can be used in courses. Let the users apply the tools to their own data and look forward to sharing with the larger CHRONOS data services. After considerations of ease of implementation, CHRONOS should prioritize those data-driven tools that are least specific and offer the widest potential applications.

- **Data quality is a major concern for tool developers.**

Developers are keenly aware that good tools produce shameful results when applied to bad data or the wrong questions. To some extent, some tools can be offered that prevent common stratigraphic blunders and others may include safeguards against their application to inappropriate and inadequate data. Although data quality problems (garbage-in; garbage-out) lie largely outside the control of program developers, concerns about quality-control, particularly for taxonomic data, arose persistently, regardless of the topic of discussion. Participants encourage all CHRONOS efforts to promote taxonomic dictionaries and subjective annotations for taxonomic entries. The task is enormous and should take advantage of developments in related database initiatives.

- **Relationships to other database initiatives need to be fostered with care.**

CHRONOS and other data base initiatives that have some overlap in scope all gain considerable strength by being federated or connected in a GRID environment. All implementation choices should be made with a view to full GRID integration in future. Nevertheless, the inter-relationships must be fostered to avoid three pitfalls: 1) waste of limited resources in a duplication of effort; 2) counterproductive rivalry; and 3) over-dependence upon a weak link in the federation. Many databases specify age as a descriptor. CHRONOS, in contrast, is concerned with the evidence of age and sequence. CHRONOS database records will not become prisoners of the originator's interpretation of age or a pre-ordained level of resolution. Rather, CHRONOS will enable the community to keep its estimates of age current with state-of-the-art resolving power. CHRONOS should direct most of its resources toward time-related data and tools and take primary responsibility for all essential developments of this kind. For tools that are primarily spatial or biological, for example, CHRONOS should support the efforts of initiatives that focus on paleogeography or paleobiology.

## 2. CHRONOS MISSION AND WORKSHOP AGENDA

For the purposes of this workshop, the essence of the CHRONOS mission is to add value to individual achievements in time stratigraphy by providing the larger community with simple and direct means to share raw data, analytical tools, and derived interpretive products. The community includes everyone with a need for a geologic time scale or an interest in the sequence of events in Earth history.

Members range from highly informed and skeptical academic researchers to geologically naïve lay-people. The success of CHRONOS will be measured by “community buy-in,” which means it must be a rewarding repository for time-stratigraphic data, an effective platform for developers of software tools to reach a wide user base, a source of rapid answers for those who simply ask what we know, and a satisfactory source of the underlying evidence for those who question how and why.

The CHRONOS steering committee has suggested the following guidelines for its interaction with software developers.

- CHRONOS will not dictate which compilers developers should use to make their tools.
- CHRONOS will provide an open and standards-based system that accommodates the needs of users and developers with various level of assistance.
- CHRONOS will be open and accessible during development so that the widest possible community will be able to test-drive its components as soon as feasible.

Initially, the CHRONOS hub is setting up the following services:

- **CHRONOS HUB SERVER:** standards-based and open; for XML, Web services etc. The current application server is TOMCAT (J2SE + JSP + Servlets). JBOSS or GERONIMO are likely in the future and, potentially, .NET via MONO ([www.go-mono.org](http://www.go-mono.org)).
- *MyCHRONOS* is for data storage and performance for typical clients. Users will send job requests to server and be notified when it is completed through a batch queuing system. Half capacity will be for users, half for auto-caching.
- **JDBC/ODBC PIPES** to other database nodes but likely not for all clients.

### 3. TOOLS FOR TIME-STRATIGRAPHY

The wide range of tools that CHRONOS should consider may be classified several ways. Classification by development history serves best to organize the implementation issues and matters of developer buy-in. Classification by purpose better displays the total wish-list of tools that was drawn up in the process of recommending priorities for CHRONOS' strategic plan.

#### 3.1 TOOLS CLASSIFIED BY DEVELOPMENT HISTORY:

In terms of the cost and development standards there are two extremes, commercial software and academic freeware.

- **Commercial Software** is characterized by a high cost to acquire. The terms of purchase will limit how widely the product can be distributed. For the high price, the advantages should be a wide user base, thorough testing, ease of support, upgrading, and maintenance, and generally good documentation. Integrated commercial products are available for generic data examination (e.g. SAS, S-Plus, Excel) and more specific stratigraphic tasks (e.g. Rockware). In addition to cost, there is the disadvantage of closed source code, with the corresponding limits to customization. Of course, commercial software is simply unavailable for many specifically stratigraphic tasks. Therefore, commercial software cannot offer a total solution and the problems of implementing academic software must be faced.

The data management component of CHRONOS seems to be avoiding closed source code and commercial prices where possible. The same principle would be sound for the tool boxes. Several developers note that the individual analytical components of many commercial packages for data analysis are rather trivial to program and are available in open libraries. If CHRONOS must provide custom data formatting for other purposes, then it would not be too onerous to write the basic analytical tools components, without purchasing whole commercial package. A related argument questions whether the integrated commercial packages offer more tools than the average user can safely employ. User's with greater analytical sophistication are likely to have access to such tools at their own workstations. Thus, the most sophisticated commercial packages may not be needed.

Some commercial software is so widely used that it has become a de-facto standard with which CHRONOS must be compatible. This is a matter of customer satisfaction. For example, CHRONOS users will surely expect compatibility with Excel spread sheets and ArcInfo GIS.

- **Academic Freeware** brings the advantage of zero purchase price and more specific applications, many of which have no commercial counterpart. This is particularly true for stratigraphic correlation tools (e.g. PAST; CONOP; BIOGRAPH; CONJUNCT, PCSLOT). These products vary from source-code modules to stand-alone, menu-driven applications with data managers. The typical disadvantages include the lack of professional support, very variable levels of documentation, a limited user base, and a wide variety of compilers and operating systems, some of them obsolete. In some cases the developers would have lost interest or cannot offer help; CHRONOS would need to resuscitate or renovate the products to a language or compiler that is familiar to the current CHRONOS community. In other cases, the product would still be actively and continually upgraded by the original developer; CHRONOS then would need to ensure that the tool is implemented in such a way that the new developments can easily be ported to the CHRONOS version. It would be counterproductive to migrate the tool to a language that is foreign to the developer; every enhancement would need to be written twice.

Between the two extremes of commercial programs and freeware there are hybrid development histories that require different treatment. CHRONOS can expect at least three of them:

- **Academic Costware** is offered at a modest price, largely to cover the cost of development. (e.g. RASC; GRAPHCOR; TIMETREK; PALEOMAP)
- **Proprietary Academic/Industrial Software**, not developed to the point of commercial release, might be obtainable by permission. (e.g. Rot Plotnick's STATISTICS, originally developed under the auspices of Argonne National Laboratories)
- **Commercial Partnerships** may allow research groups access to commercial software at below-market cost. For example, GEON has a working relationship with GIS developers.

#### THE CASE FOR A CHRONOS PANEL OF DEVELOPERS:

Evidently, for CHRONOS, academic freeware brings the potential burden of customizing a working relationship with each individual developer. Nevertheless, this is exactly the type of product development that CHRONOS should foster, enlarging the user base, the flow of ideas, and the rate of innovation. A panel of stratigraphic software developers ought to be able to encourage standards and assist with new developments and upgrades. Ideally, CHRONOS would foster a better development environment by bringing together programmers who currently operate in rather isolated settings.

Although academic freeware may be available with open source code, most developers prefer not to allow unrestricted access to their source code. Some of this reluctance stems from the same kind of protective bonding that scientists develop with their data. There is also a more general objection to freely distributed source code: there may evolve a bewildering proliferation of poorly documented versions and hybrids that eventually become incompatible and discourage users. The preference is that the source code be open to CHRONOS and, through the developers' panel, to any legitimate enquirer. By controlling access to the source code, CHRONOS might be able to channel all beneficial modifications into a single version.

The purview of the panel could include:

- recommending minimum standards to meet "*CHRONOS ENABLED!*" approval
- advising on the use of CHRONOS resources for testing and translating source code, generation of better documentation, development of uniform pop-up help messages, providing tutorials and sample data
- advising CHRONOS on priorities for tool implementation
- oversight of the distribution of source code
- development of explicit terms of use and disclaimers
- development of language for citation of tools distributed by CHRONOS and used for published research. Tool developers generally lament the lack of academic credit for programs as compared with topical papers.
- facilitating proposals for funding of tutorial workshops

#### CHRONOS CAN REDUCE COMMON HINDRANCES TO TOOL USE

After CHRONOS makes tools readily available to the community, there will remain at least two common reasons not to use the tools. The first hurdle is to learn what the tools can do and how to make them do it. CHRONOS should encourage good documentation and sample data sets. Ultimately, tools attract community attention most effectively when they are used in publications that solve significant problems. CHRONOS can be cited as a source in publications. The second hurdle is the chore of formatting the input files; CHRONOS could eliminate this chore entirely. These hurdles are higher for academic freeware than for commercial software. Commercial software tends to a

common look and feel and often includes format converters. Academic programs usually begin for the developers own use; each tends to have a unique look-and-feel and its own input idiosyncrasies.

#### CHRONOS CAN BUILD ON AN EXISTING MODEL FOR DEVELOPER COLLABORATION

Some of the academic freeware programs (e.g. Oyvind Hammer's PAST) bundle several tools under a common menu and include data management modules. Chronos should give such tools a high priority for three reasons. First, they present the user community with lower hurdles. Second, some of their modules are scaled-down versions of more complex specialized tools and serve as stepping stones to them (e.g. PAST to RASC and PAST to CONOP). Third, when an academic program is scaled down for bundling into an integrated environment, the developer of the specialized module provides source code with the understanding that the programmer of the umbrella program will modify it as necessary. In the process, both programs benefit from exposure to more users and there are now two programmers familiar with the internal operation of the specialized module. Thus, bundled tools, like PAST, indicate one successful pathway that CHRONOS can follow to developer collaboration.

### 3.2 TOOLS CLASSIFIED BY APPLICATION

The following catalog of tools was drawn up by workshop participants as "dream list" from which priorities could be drawn. The catalog is split into three categories according to the degree of interaction with CHRONOS data bases: user-driven, program driven, or none.

A wide variety of tools will be able to interact with information drawn by the user from the main dynamically growing data bases that are hosted by CHRONOS or within the federation of databases to which CHRONOS has pipes. These are here termed *data-based* tools; the user selects both the data and the tool; the data bases are not constructed exclusively for any one data-based tool; the tools are likely to be put to a variety of uses depending on the data retrieved by the user. Data-based tools are contrasted with *question-based* tools for which the user enters a question in a standard format and does no data selection. The question-based tool interrogates specific data bases, typically small ones constructed explicitly to serve a tool that answers one type of enquiry; e.g. "What is the duration of Period X in the time scale published by A?" Researchers and students in the geosciences are likely to use tools of both kinds. Lay persons are more likely to use the question-based tools. The third category, *demonstration based*, allows the user to learn by interactive play, without asking questions or requesting data searches. Interactive animations fall into this category; they can be fully self-contained and independent of CHRONOS data bases.

#### 3.2.1 DATA-BASED TOOLS

An asterisk indicates generic database management tools that are outside the workshop mandate.

- **Data Input\***

In order to encourage the community to donate data, CHRONOS databases strive for a minimum of required information in each record. The minimum information necessary to run any given tool, however, is likely to exceed the minimum for acceptance into the database. CHRONOS should consider an advisory sheet that indicates which non-mandatory fields are necessary for particular tools. The need for such advice will be clearer when a pattern of tool use emerges.

- **Annotation and Quality Control Tools\***

- **Taxonomic Dictionaries**

These tools assist in the standardization of taxon names and may automate the handling of synonyms. They are to be the subject of separate workshops and can benefit by collaborating

with other database initiatives, such as the Paleobiology Database. They are mentioned here to stress their importance to the meaningful application several analytical tools.

- **Retrieval/Query Tools \***

These data management functions have a crucial interface to the toolbox. They should include modules that are capable of rendering the data into formats suitable to the contents of the toolbox. Data formatting is one of the biggest disincentives to using academic freeware.

- **Data Display (visualization) Tools**

The Geosciences are visually oriented and accustomed to graphical displays of data. Favorable first encounters with CHRONOS will depend upon effective graphical displays and mapping of the results of database queries.

*Generic* (such as provided with Excel, SAS, S-Plus, Systat, etc)

- one-variable frequency distributions

- X-Y graphics with point IDs and error analysis

- GIS Mapping (paleogeographic distribution of data)

  - several tools involved

  - map client browser

  - map renderer (ARC IMS)

  - map server (knows spatial geometry; e.g. ARC SDE)

  - map database (DB2, ORACLE)

- 2-D and 3-D contouring, trend surfaces

*Stratigraphic* (some provided in ROCKWARE, PAST, and STATISTICS)

- Stereonets and rose diagrams

- Columnar sections, and well logs, and fence diagrams

- pollen diagrams

- Range Charts

- Shaw Plots (graphic correlation type X-Y plots of 2 range charts)

- depth-age plots

- Fischer Plots (thickness series data for cyclostratigraphy)

- **Data Analysis Tools**

Users will vary considerably in their sophistication in the use of analytical methods.

CHRONOS should take care not to purchase expensive commercial software that is beyond the needs of users who do not have their own computational facilities. Beyond the potential waste of financial resources, there are more substantial risks: users may be intimidated by advanced options or, worse, they may be lured into mistakes. First, do no harm.

*Generic* (such as provided in Excel, SAS, Systat, etc)

- 1 – 2 variable descriptive statistics

- Splining and regression with confidence intervals and analysis of residuals

- Multivariate analysis, PCA, clustering analyses,

- Geostatistics – spatial statistics, autocorrelation, semivariograms

- Fractal analysis of spatial patterns - lacunarity

- Time Series tools (much Fortran and Matlab freeware and shareware available)

  - correlation,

  - interpolation, resampling,

  - spectral analysis, tuning, filtering

*Stratigraphic* (many provided in PAST and STATISTICS)

For these tools, especially those that attempt difficult and partly subjective tasks, like correlation, CHRONOS needs to support many different tools as possible. Different tools are adapted to different data types and answer significantly different questions; it is intellectually sound for the user to be forced to consider the choices on a case-by-case basis. There will rarely be one “right” or “best” tool for all situations.

(Bio)Stratigraphic correlation:

Graphic Correlation GRAPHCOR (Ken Hood)

Unitary Associations BIOGRAPH (Jean Guex) and in PAST (Oyvind Hammer)

Appearance Event Ordination CONJUNCT (John Alroy)

Ranking and Scaling RASC (Fritz Agterberg and Felix Gradstein) and in PAST

Constrained optimization CONOP (Pete Sadler) and PAST (Oyvind Hammer)

(Litho)Stratigraphic correlation

slotting (e.g. PCSlot for palynology)

wireline log zonation and facies discrimination (single wells)

depth constrained cluster analysis (DCCA, Gill et al.)

wireline log correlation (multiple wells)

CORRELATOR, R. A. Olea

Cross section generation Color Log

synthetic sequence and log generation

Markovian simulation (OZ MACHINE)

Fourier simulation (SYNTHLOG)

Pollen diagrams and time series (e.g. POLSTA; David Green, ANU)

randomizing, bootstrapping,

back-stripping and burial history

confidence intervals and stratigraphic error analysis

### 3.2.2 QUESTION-BASED TOOLS

**Time scale converters and calculators** “How do I convert DNAG ages to Harland et al. ages?”

This is the most common kind of time-stratigraphic question addressed to experts; it requires look-up tables and a conversion algorithm that makes an age to age conversion, given two time scales.

– age-in : age-out (given import and export timescales)

–

The tables are **time stratigraphic dictionaries** that store definitions and usages. There simpler questions that can use these tables, without any calculation: “What is the age of the Namurian in the 2004 IGCP time scale?” “Into what stage do 123 million year old rocks fall according to Harland et al?” These date ashes were said Ludlovian using the DNAG time scale, does that change using the 2004 IGCP scale?” “What does Viséan mean?” Some of these questions concern matters of terminological translation for which GEON already has expertise.

– age-in : name-out (given time scale)

– name-in : age-out (given time scale)

– name-in : name-out (given import and export timescales) GEON?

– name-in : history-out, definition-out, usage-out (given time scale)

**Time Machine** “What was the latitude and longitude of this site in the Permian?”

i.e. give age and modern lat-long; get paleo lat-long for that age

Chris Scotese's PRIM box is an array designed to support this question.

Boxels in a PRIM box have 3 coordinates modern-lat and long + and ancient time; they contain the paleo lat and long for each ancient time.

PRIM boxes can be used to store multiple models of plate reconstruction and paleogeography. As with time scales and correlation tools, CHRONOS should emphasize the options and differences. This is an important aspect of education about science as a way of thinking; it should have priority over the temptation to simplify by arbitrary authoritative decision.

Chris Scotese's STIM box is an array designed to support other questions about a site they store any data by modern coordinates, as in a PRIM box the data at each boxel site may be and site-time specific information

CHRONOS could populate STIM boxels with time-stratigraphic information to answer a variety of site and time specific questions

e.g. "What index fossils could I use here for Permian rocks?"

"What is the biostratigraphic resolving power?"

"What was the climate?"

i.e. STIM cubes summarize of interpretations and derived information for which an age interpretation has been made

NOTE: the main CHRONOS database are explicitly NOT forced into time categories, but contain the data suitable for making age interpretations; some STIM boxes might be suited to automatic updating from the main data bases. STIM boxes can store spatio-temporal data beyond the primary CHRONOS needs; they will be a vehicle for collaboration between federated databases. CHRONOS should take care not to spread its resources too far beyond the temporal components.

**Accumulation Rate Estimators** "What is the average thickness of foraminiferal ooze accumulated in 10 million years?" What is the likely variability in accumulation rates on carbonate platforms?

Such questions arise frequently because the stratigraphic record presents us with thickness scales where a time scale is needed. The questions must be answered with care because the mapping functions are complex and probably fractal in nature – accumulation is unsteady and expected rate falls with increasing time span,. Order-of-magnitude answers should be standard stratigraphic expertise; more precise answers may be needed by modelers. Sadler has a relational database of accumulation rates with over 300,000 entries, differentiated by time span and environment. It can answer these questions by providing the scaling laws for sediment accumulation. The database might be left static and independent of the main databases; alternatively, it might be augmented automatically from geochronometric records in the main databases.

**Frequently Answered Questions** – For other specific questions that CHRONOS frequently needs to answer, the derived maps and data analyses might be stored with the questions in an FAQ format.

### 3.2.3 DEMONSTRATION-BASED TOOLS

Any tool might be used in the classroom to answer questions or solve problems, particularly if sample data sets are provided. This third category of tool is primarily for demonstration or free-

form “play.” These tools are not prompted by a particular questions or data sets. The classic example is user-interactive paleogeographic animations. Another possibility would be one-dimensional sequence-stratigraphic models of cyclic sediment accumulation as a function of steady subsidence and fluctuating sea level.

**Time trek** (courtesy of Alan Smith, Cambridge) is a point-and-click application and demonstration of plate tectonic reconstructions  
an application of time-constrained data

### 3.3 TOOL IMPLEMENTATION

Academic freeware tools will a significant part of CHRONOS. The fashion in which such tools are made available will have considerable implications of the user (ease of access, necessary hardware), the developer (need to rewrite), and CHRONOS personnel (need to rewrite, maintain).

#### 3.3.1 USER ACCESS

Tools may be provided to users in a variety of ways. The following options are listed in order of increasing

- Locked into specific Operating System and to be downloaded by the user to hardware running that OS. This is commonly the format for which academic software is initially written.
- (Re)compiled for cross-platform use, but still to be downloaded. This might be achieved by compiling under Microsoft .NET. This is not an open source implementation but it can be fast and straightforward for tools developed with languages and compilers that are compatible with the .NET suite; e.g. Microsoft Visual Basic and C; Digital Visual Fortran, Intel Fortran.
- Hosted applications (via WWW, web services, GRID)
- Full GRID environment with certificates to control access levels (to understand certification, imagine current access to online journals, but improved)

The possibilities for the PAST program are illustrative

- if downloaded as a tool for a specific operating system, e.g. WIN 32, PAST is available only for Windows machines or Windows emulators.
- if migrated to .NET, PAST has potential for cross-platform access via MONO
- if re-written into XML or JAVA, cross-platform application is achieved; rewriting in JAVA may be the most labor intensive option except for tools whose original compilers are not .NET eligible.
- full GRID integration

CHRONOS can help with the second and third options. Rewriting poses a problem if the original developer is not fluent in the new language. It is not feasible to insist that the developer learns the new language. Failing this, however, either the translation chore is repeated every time the developer modifies the tool, or the tool splits into two species. GEON should help with the fourth option and CHRONOS should be always attempt to choose options that will simplify the eventual grid integration.

#### 3.3.2 TOOL EFFICIENCY

The efficiency of some tools, particularly those that sequence stratigraphic events, falls appreciably as the volume of data increases. CHRONOS databases will allow users to download larger volumes of data than are currently compiled by most individual researchers. Thus, some tools may be overwhelmed in the distributed environment. CHRONOS expertise

and hardware resources might be used for migrating such tools to parallel computing, in anticipation of future data overload.

### 3.3.3 DEVELOPER TRAINING

Provide workshops or other forums for developers and potential developers to learn programming practices that will make their products more easily “*CHRONOS ENABLED!*”

### 3.3.3 USER TRAINING

Provision of tools is not enough. CHRONOS should make a good faith effort to ensure that the tools steer users away from silly errors. First, provide tools that perform correctly those chores that are popularly mishandled, e.g. time scale conversion. Second, provide tools with a good online instruction and examples. Third, schedule workshops at national meetings to demonstrate the more advanced tools.

Facilitate the distribution of tools and data sets to graduate course that can provide beta-testing while training potential users. Defray the expenses of graduate students who wish to travel to learn tools and their operation by working in the developer’s laboratories for short intervals. (Similar exchanges are encouraged to train graduate students at geochronometric laboratories – Earthtime workshop, Washington DC, October 2003.)

## 4. SPECIFIC RECOMMENDATIONS and PRIORITIES

- Always worry about the users’ first encounter, with CHRONOS in general and tools in particular. Try to fill some immediate needs, painlessly (i.e. address frequently asked questions). Provide “something for nothing,” easily and immediately (tools to download!). Try to ease the learning curves and the ugly chores (like input formatting). Do not let the first-time visitor focus too heavily on promised developments or advanced data services with a steep learning curve. Do not overwhelm the first-time user with the impression that comprehensive and sophisticated tools suites are essential for CHRONOS data.
- Prioritize tools with regard to difficulty for CHRONOS and for users. Start with those that are easiest to implement and to use. Easiness will be a balance of two issues. Ease of implementation is a matter of cost and amount of modification. Ease of use is a matter of the interface mechanics and the required level of mathematical insight.
- Start with question-based tools that do not require that the main data bases catch up to their needs. Set up answers to the most common and irritating tasks that can be independent of the main data bases, by using small custom databases or flat files.
  - time scale converters
  - time machine (paleo lat/long values)
  - accumulation rate statistics (for facies and time span)
- For data-based tools that require the main data bases, start with the integrated stratigraphic tool boxes PAST and STATISTICS; they provide the most power for the least user effort; they also serve as attractive introductions to some of the more sophisticated individual tools.
- For the more advanced and specialized tools, start by offering pointers to shareware and a suite of free downloads. The number of hits and the feedback may help guide the prioritization for hosting and integrating the different tools. Otherwise prioritize by ease of implementation – largely a matter of original language and operating system.
- Where several tools attempt the same task with different assumptions or methods, do not attempt to find the “one size that fits all.” Offer as full a suite as possible and explain the differences; i.e. encourage intelligent user selection.

- Provide format conversions for the popular tools, hosted and downloaded. Consider providing Excel macros for format conversion.
- Get the Neptune and Paleostrat data formats, with sample data sets to tool developers as soon as possible. Post the input formats for the tools on the wish list - let the developers solve some format compatibility questions themselves.
- Plan for parallelization of those tools that are computationally intense and can “crunch” large data sets. If the main data bases function and grow as planned, user will extract and run dangerously large data sets through the tools.
- Consider establishing a tool developers’ panel to handle source code access, standardization, citation, etc. (keepers of the real or imaginary “*CHRONOS ENABLED!*” seal of approval).
- Foster the emphasis on taxonomic quality control. Without this, the correlation tools can produce ridiculous answers. Ask the taxonomists to consider whether this trait might be used to search automatically for *potentially* questionable names – i.e. “correlate” new data sets to standard biostratigraphic sequences and treat anomalies (exceptionally long ranges, numerous unusual coexistences) as warning signs.
- Seek opportunities to collaborate with overseas colleagues. In addition to presence at international meetings, think about a theme session suitable for sponsorship by European colleagues through the Geological Society of London, for example.

## **APPENDIX I: PARTICIPANTS**

(loosely grouped by expertise; several individuals represent multiple categories)

### NATIONAL SCIENCE FOUNDATION

Rich Lane  
Walt Snyder

### CHRONOS STEERING COMMITTEE

Cinzia Cervato, Iowa State  
Bruce Wardlaw, USGS

### INFORMATION TECHNOLOGY

Chaitan Baru, UC San Diego Supercomputer Center, GEON distributed resources  
Geoff Bohling, Kansas Geological Survey, Statistics  
Doug Greer, UC San Diego Supercomputer Center, programmer/analyst  
Dogan Seber, UC San Diego Supercomputer Center

### GEOCHRONOMETRIC/PALEOMAGNETIC DATA/TOOLS

Anthony Koppers, UC San Diego IGPP

### PALEOGEOGRAPHIC TOOLS:

Allister Rees, Arizona, Paleogeographic Atlas Project, GEON  
Chris Scotese, Texas Arlington  
Alan Smith, Cambridge

### PALEOBIOLOGICAL TOOLS

John Alroy, NCEAS Santa Barbara Paleobiology Database

### GENERAL STRATIGRAPHIC TOOLKITS

Oyvind Hammer, Oslo (PAST)  
Roy Plotnick, Illinois Chicago (STRATISTICS)

### CORRELATION/SEQUENCING TOOLS

Fritz Agterberg, Geological Survey of Canada (RASC)  
Vladimir Davydov, Boise State (Graphic Correlation)  
Lucy Edwards, USGS (Graphic Correlation, No-space Graphs)  
Pete Sadler, UC Riverside (CONOP)

### TIME-SERIES TOOLS

Linda Hinnov, Johns Hopkins (digital presentation)

(appendices under construction)

**APPENDIX II STIM and PRIM cubes** (courtesy of Chris Scotese)

**APPENDIX III PAST MENU TREE**

**APPENDIX IV STATISTICS MENU TREE** (courtesy of Roy Plotnick)

(FORTRAN 95 AND WINTERACTER)

- Univariate and bivariate statistics
- Fourier Analysis
- Fisher Plots
- Autocorrelation and semivariograms
- Spatial statistics and pattern analysis

**APPENDIX V CONOP OBJECTIVE FUNCTIONS AND OUTPUT MENU TREE**