



Electronic data publication in geochemistry

Hubert Staudigel

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA (hstaudigel@ucsd.edu)

John Helly

University of California, San Diego, San Diego Supercomputer Center, 9500 Gilman Drive, Mail Code 0527, La Jolla, California 92093, USA (hellyj@spsc.edu)

Anthony A. P. Koppers

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA (akoppers@ucsd.edu)

Henry F. Shaw

Division of Chemical Sciences, Geosciences, and Biosciences, SC-14, Office of Basic Energy Sciences, U.S. Department of Energy, 19901 Germantown Road, Germantown, Maryland 20874, USA (henry.shaw@science.doe.gov)

William F. McDonough

Department of Geology, University of Maryland, College Park, Maryland 20742, USA (mcdonough@geol.umd.edu)

Albrecht W. Hofmann

Max-Planck-Institut für Chemie, Postfach 3060 55020 Mainz, Germany (hofmann@mpch-mainz.mpg.de)

Charles H. Langmuir

Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA (langmuir@eps.harvard.edu)

Also at Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA.

Kerstin Lehnert

Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA (lehnert@ldeo.columbia.edu)

Baerbel Sarbas

Max-Planck-Institut für Chemie, Postfach 3060 55020 Mainz, Germany (sarbas@mpch-mainz.mpg.de)

Louis A. Derry

Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York 14853, USA (derry@geology.geo.cornell.edu)

Alan Zindler

National High Magnetic Field Laboratory and Department of Geological Sciences, Florida State University, 1800 E Paul Dirac Drive, Tallahassee, Florida 32306, USA (zindler@magnet.fsu.edu)

[1] Many disciplines of geochemistry have no data reporting standards, and their use of metadata is inadequately developed. This presents problems to the quality of the published science, and it limits the utility of computers in data analysis and the exploitation of Information Technology (IT). We discuss problems of data and metadata publication, in particular for geochemistry, and offer solutions to these problems in the form of consistent data publication formats and a proposal for publication of metadata in

geochemistry. Metadata are grouped according to types (location, sampling, characterization), and this grouping allows for the transfer of these formats to other Earth science disciplines. In a companion paper [Helly *et al.*, 2003], we illustrate how these metadata groupings can be used in an IT context. Formats presented here are comprehensive and allow for modification and expansion. It is the hope of the authors that this paper initiates a constructive discussion of data formats and metadata in geochemistry. The most recent contributions to this discussion may be found at <http://earthref.org/metadata/GERM/>.

Components: 6046 words, 3 figures, 13 tables, 1 dataset.

Keywords: Metadata; geochemistry; database; information technology; cyberinfrastructure; sample characterization.

Index Terms: 9810 General or Miscellaneous: New fields (not classifiable under other headings); 1094 Geochemistry: Instruments and techniques; 3094 Marine Geology and Geophysics: Instruments and techniques; 4594 Oceanography: Physical: Instruments and techniques.

Received 23 January 2002; **Revised** 24 June 2002; **Accepted** 25 June 2002; **Published** 29 March 2003.

Staudigel, H., *et al.*, Electronic data publication in geochemistry, *Geochem. Geophys. Geosyst.*, 4(3), 8004, doi:10.1029/2002GC000314, 2003.

Theme: Geochemical Earth Reference Model

Guest Editor: Hubert Staudigel

1. Introduction

[2] Geochemistry is still in the earliest stages of its exploitation of Information Technology (IT) as a tool for research, publication and data archiving. Nonetheless, it is clear that geochemistry, like most other Earth science disciplines, stands to reap substantial benefits from embracing IT. These benefits include wider dissemination of geochemical data, increased ease of use of data by different Earth science subdisciplines, and more efficient storage and retrieval of archived data. Current common practices in the publication of geochemical data, however, present unnecessary obstacles to effective use of IT in geochemistry. These obstacles typically involve the lack of standardized formats for data and/or the omission of essential metadata (supplemental data that “describe” the “real” data). Common problems include the following.

1. Data may be published in various units and normalizations. Elemental abundances can be given as molar or weight fractions, as elemental or oxide fractions, and in relation to volume or weight. Abundance data are sometimes referenced to unexplained normalizations or to laboratory-specific standards or procedures without providing a

cross-calibration to other laboratories. Isotopic measurements of the same isotopic system can be referenced to different standard materials, different instrumental mass fractionations or in inverted notations. Often there is no simple way of converting from one representation to the other.

2. Lack of conventions for the arrangement of data in tables. Depending on the numbers of samples or parameters, and on the layout of a journal, data for a given sample can be ordered by either rows or columns. Data categories can be found listed almost in any sequence.

3. Lack of numerical values for data appearing in figures. It is increasingly common for data to be presented only in graphical form, with no accompanying numerical data table. This practice is particularly common in high-profile letter journals that have restrictive page limitations.

4. Incomplete metadata. Descriptions of samples are rarely complete, and important sample characteristics may not be recoverable from a publication. Sample locations may be missing or given only as vague written descriptions that cannot be located on maps with sufficient accuracy and precision. Critical analytical details may not be described at all or are scattered through the body of a paper, in appendices, data table subscripts, or in cited papers.

[3] These problems may be more prevalent in some disciplines in geochemistry than others, but they have a profound influence on how geochemistry functions as a scientific discipline. Inconsistency in data formats, and the lack of simple conversions between different normalizations set up unnecessary obstacles that are often difficult to overcome. This limits the extent to which data are accessible to researchers in other disciplines of Earth system science and even to researchers in different specialties of geochemistry. These problems also make it unnecessarily arduous to meaningfully review the literature. As a result, geochemical data are much less efficiently used than they could be. Most importantly, many published data cannot be (re-)used because of the lack of critical metadata.

[4] Some of the most important problems could be easily remedied if the community would embrace the electronic publication of data using a consistent standard with a minimum of mandatory metadata that are enforced in the review process of journal submissions. Metadata flagged as “essential” in this paper are suggested as candidates for such mandatory metadata. The time is right for setting up such standards for data supplements. Most highly ranked journals have begun publication of electronic data supplements but their current use is minimal at best. *Geochemistry, Geophysics, Geosystems* is entirely electronic and has a substantial commitment to the efficient publication of data, as evidenced by the fact that one of its publication categories is entirely devoted to the publication of data. On the IT side, the development of data description languages, such as the Extensible Markup Language (XML) or metadata interchange formats (*.mif) are making it possible to “package” data into structures that are self-describing and can be automatically processed and used by any software that knows how to parse the description (see companion paper by *Helly et al.* [2003]). Given this state of affairs, the major remaining obstacle to more effective electronic data storage, retrieval, and use in geochemistry is the definition of an appropriate minimum set of metadata. This should not be particularly difficult. In most cases, there is little ambiguity about which metadata need to be supplied in order to document geochemical data in a scholarly manner. The pur-

pose of the metadata is to make the associated data maximally reproducible, searchable, easily usable, and comparable to other data in the same field. The main step remaining is putting these metadata in logical sequence and providing a format that can be easily read by humans and by computers. In this paper we propose such a data/metadata format for electronic data supplements. These supplements may be used in parallel with the typeset data tables in a paper or (ultimately) replace them. We focus here on solid (mostly geological) sample types, such as rocks and minerals, but we also apply this methodology to other types of geochemical samples, in particular fluids and gases.

[5] Full disclosure of details regarding the samples, the sampling process, and the analytical process is essential to a meaningful analysis of geochemical data. The current literature does not contain many examples with proper sample metadata, a habit that is often based on the size limitations of paper journals [see *GERM Steering Committee*, 2001]. In the age of electronic dissemination of data, it is now possible to publish all data and metadata without the restrictions imposed by journal layout. This shift toward electronic data dissemination should be used to begin disclosing a critical minimum amount of information on all samples for which data are published in the peer-reviewed literature. The metadata scheme offered here contains a set of “essential” metadata for a scholarly sample description that would resolve most of the problems listed above. Nevertheless, we put this scheme together not as a unique and ultimate solution but as a catalyst for a discussion on how the geochemical research community can most effectively take advantage of advances in Information Technology. A broadly accepted metadata standard may emerge after such a discussion.

2. Metadata Defined

[6] The term “metadata” is relatively new to the language of the Earth sciences, and for this reason we wish to define the term both in general and in the sense we are using it in this document. Metadata are “data about data” and they can have different functions and contents. By function, we distinguish

cataloguing metadata from application metadata. Cataloguing metadata includes any information that may be used in a card-catalogue to search for the existence of data. Dublin Core is one of the best established sets of cataloguing metadata used, in particular, by the library community (<http://dublincore.org/>). Application metadata summarizes all information relevant to any particular science application. Such metadata come from the research archives of scientists and typically explain how data in question were produced and processed, including the nature and location of analyzed samples. They may include information that may be useful for the cataloguing of data including, for example, sample location, sample type, or sample age. Here we focus on metadata with scientific contents.

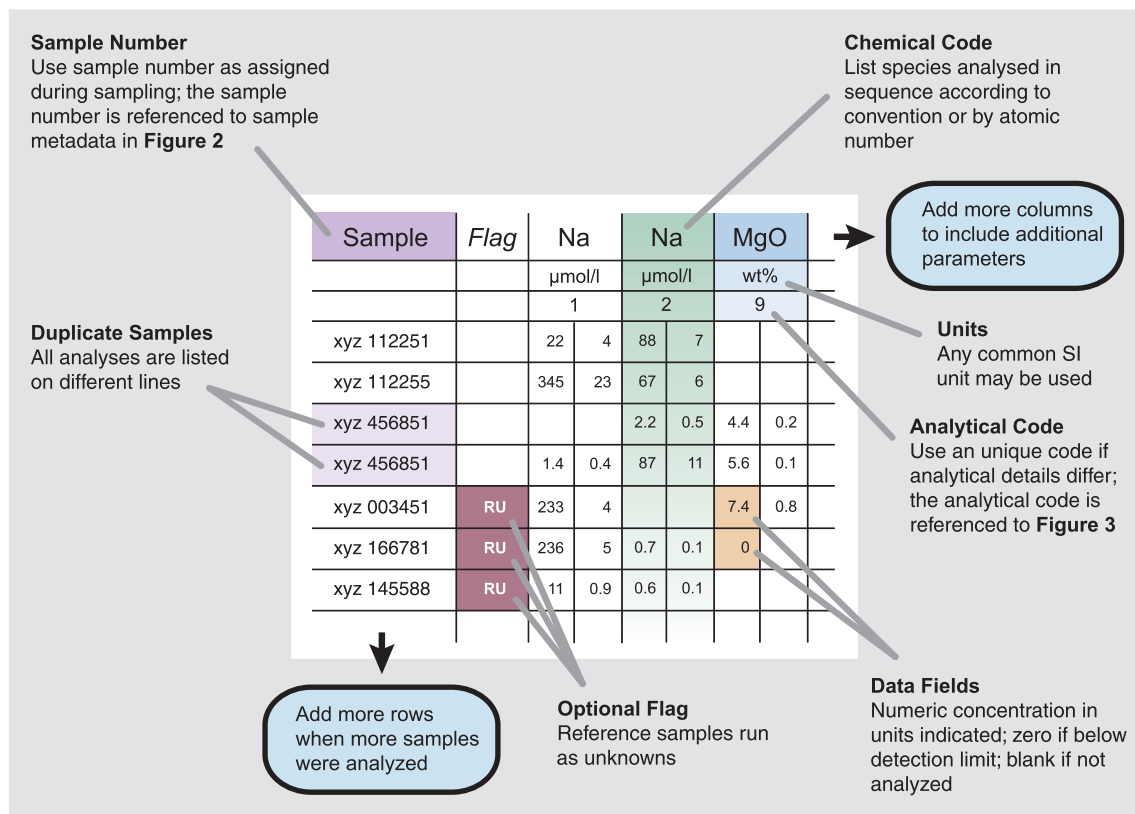
[7] Metadata may exist in the form of numeric or alphabetic entries such as keywords, abbreviations, or a “controlled vocabulary” specific to a particular science discipline. Metadata can be expressed in written descriptions, in data tables, or in formats that are optimized for automated computer processing. In this paper, we focus on a tabulation of metadata in a form that is both comprehensible to humans and can be easily processed by computers.

[8] While it would be desirable to have extensive metadata available for every data point, it is clear that there must be some balance between the utility of the metadata and the effort required to collect and archive them. For this reason, it is important to find a minimum set of metadata that provides a reasonably complete description of the data, consistent with the goals of maximizing the utility and reuse of data. Furthermore, metadata formats must be flexible in order to accommodate the metadata needs of particular studies. In the metadata format discussed below, we have defined a minimum set of “essential” metadata that should be reported in any scholarly publication; all other metadata are considered important and useful but optional. Each investigator is free to include additional new types of metadata, if a particular type of scientific contribution will benefit from their inclusion. This should accommodate the needs of specialist communities and provide the potential for modifying the format so it can serve as a data input format for various databases.

[9] The distinction of data and metadata can be confusing, whereby the same value or entry may be used in either way. For example, the latitude/longitude of a sample location may be used as metadata in a catalogue or as an archiving principle, or it may be used as data in a scientific analysis of sample properties as a function of regional distribution. We have grouped geochemical metadata in a modular framework that is transparent and applicable to different types of data and samples. Some of the metadata modules can be used for almost any geochemical sample type (e.g., sample identification, geographic location, sampling procedure, analytical procedure), while others are more specific to particular sample types. In the following text we explain the overall presentation of data and sample-related metadata with the use of illustrated examples. Specific formats for the description of geological, water, and gas samples will be presented next to a common format for sample identification, geographic location and the sampling process. In all cases, these formats should be considered proposals intended to stimulate further discussion.

3. Data Files

[10] Electronic data supplements should be published as downloadable comma-delimited ASCII files that are not displayed as typeset data tables in the text version of a paper. Comma-delimited ASCII data files are universal and usable across any computer platform and software. They may be produced from spreadsheets (see the .csv format option in Excel) and read by all standard editors and easily read back into spreadsheet software relatively. Numeric data entries should not use commas but commas are okay in text strings (in a comma-delimited format text strings are marked by double quotes). In Figure 1 we have given a schematic example for such an electronic data file in tabular form (see appendix 1 for examples). Data for particular samples are arranged in rows with data categories in columns. Each row is a separate data record with a unique sample identifier that is tied to the sample metadata table (Figure 2; appendix 2). Each data column gives the chemical symbol for the parameter analyzed, the SI unit used, and an ana-



Appendix 1

Figure 1. The main features of a geochemistry data table. Data categories are listed as columns and sample categories as rows. The first column gives the sample number which also provides the link to the sample metadata in Figure 2. The first three rows give the parameters analyzed, (SI) units and the analytical code for each parameter column. The analytical code links this particular analysis to the relevant metadata explaining the analytical techniques (see Figure 3). The sequence of elements should be listed either following a well established convention, or atomic number.

lytical code that is tied to a metadata description of analytical techniques as provided in a separate table (Figure 3; appendix 3). As a result, data are linked to their metadata through the sample label in the first column and an analytical code that is given in the third field of each column.

[11] Data files may have any number of columns, for all elements or isotope ratios analyzed. Data columns without data are not listed, and fields without data remain blank. Values that fall below the detection limit or below the calibrated range of an instrument should be given as “bdl.” It is preferable to have all the data in one table, rather than several separate tables. Multiple types of analyses of the same element in the same sample should be displayed in separate columns with separate analytical codes.

[12] Multiple analyses of a given sample are displayed as separate lines (data records) but with identical sample labels. Analytical data on known reference samples are listed in the last rows of the data table. We encourage a consistent sequence of data columns. Such sequences should follow a broadly used “conventional” form. However, if there is no widely shared convention, we recommend listing elements by atomic number. This sequence is universally recognizable, without knowledge of particular geochemical element characteristics, such as the exact sequence of relative compatibility of particular trace elements with a particular magmatic fractionation process. However, as long as data columns are clearly flagged, their sequence is clearly a second order problem in data publication. We illustrated some examples for

Rock Samples	Water Samples	Gas Samples
Sample ID <ul style="list-style-type: none"> ● Sample Number ● Sample Description ● Laboratory Identifier 	Sample ID <ul style="list-style-type: none"> ● Sample Number ● Sample Description ● Laboratory Identifier 	Sample ID <ul style="list-style-type: none"> ● Sample Number ● Sample Description ● Laboratory Identifier
Geographical Data <ul style="list-style-type: none"> ● Sample Latitude ● Sample Longitude ● Lat-Lon Accuracy ● Elevation ● Elevation Accuracy ● Depth in Core / Section ● Location Description ● Reference ● Remarks / Keywords 	Geographical Data <ul style="list-style-type: none"> ● Sample Latitude ● Sample Longitude ● Lat-Lon Accuracy ● Elevation ● Elevation Accuracy ● Depth in Core / Section ● Location Description ● Reference ● Remarks / Keywords 	Geographical Data <ul style="list-style-type: none"> ● Sample Latitude ● Sample Longitude ● Lat-Lon Accuracy ● Elevation ● Elevation Accuracy ● Depth in Core / Section ● Location Description ● Reference ● Remarks / Keywords
Sampling Data <ul style="list-style-type: none"> ● Sampling Time ● Time Zone ● Collector / Scientist ● Archiver / Curator ● Expedition Name ● Sampling Platform ● Sampling Method ● Reference ● Remarks / Keywords 	Sampling Data <ul style="list-style-type: none"> ● Sampling Time ● Time Zone ● Collector / Scientist ● Archiver / Curator ● Expedition Name ● Sampling Platform ● Sampling Method ● Sampling Container ● Sample Treatment ● Water Temperature ● Reference ● Remarks / Keywords 	Sampling Data <ul style="list-style-type: none"> ● Sampling Time ● Time Zone ● Collector / Scientist ● Archiver / Curator ● Expedition Name ● Sampling Platform ● Sampling Method ● Sampling Container ● Sample Treatment ● Gas Temperature ● Gas Pressure ● Humidity ● Reference ● Remarks / Keywords
Characterization <ul style="list-style-type: none"> ● Chemical Type ● Metamorphic Type ● Sediment Grain Size ● Texture ● Alteration ● Tectonic Setting ● Reservoir ● Reference ● Remarks / Keywords 	Characterization <ul style="list-style-type: none"> ● Water Type ● Water Classification ● Density ● Salinity ● Potential Temperature ● Alkalinity ● Turbidity ● Reference ● Remarks / Keywords 	Characterization <ul style="list-style-type: none"> ● Gas Type ● Gas Classification ● Reference ● Remarks / Keywords
Mineral <ul style="list-style-type: none"> ● Mineral Name ● Occurrence Type ● Mineral Shape ● Mineral Size ● Reference ● Remarks / Keywords 		
Fossil <ul style="list-style-type: none"> ● Fossil Group ● Fossil Species ● Fossil Preservation ● Reference ● Remarks / Keywords 		
Age Data <ul style="list-style-type: none"> ● Absolute Age in Ma ● Age Accuracy ● Stratigraphic Age ● Stratigraphic Unit ● Reference ● Remarks / Keywords 		

Appendix 2a

Appendix 2b

Appendix 2c

Figure 2. Metadata for sample descriptions for rocks, waters (fluids) and gases. Essential metadata are flagged with red dots and examples for these metadata descriptions can be found in appendix B1–B3. Note the similarity and consistency in metadata categories for Sample ID, Geographical Data and Sampling data for all sample types. See text and examples for detailed description of parameters.

data tables templates in appendix A. Common abbreviations used in these data tables can be found in appendix D.

[13] Petrologists commonly report major element analyses as oxides using the assumption that cations

in silicate rocks bond exclusively with oxygen and do not form any metallic bonds between them. This practice works well for most cations, except for Fe, which can display variable oxidation states. In this case, a choice has to be made for which oxidation state to use in its representation in a data column.

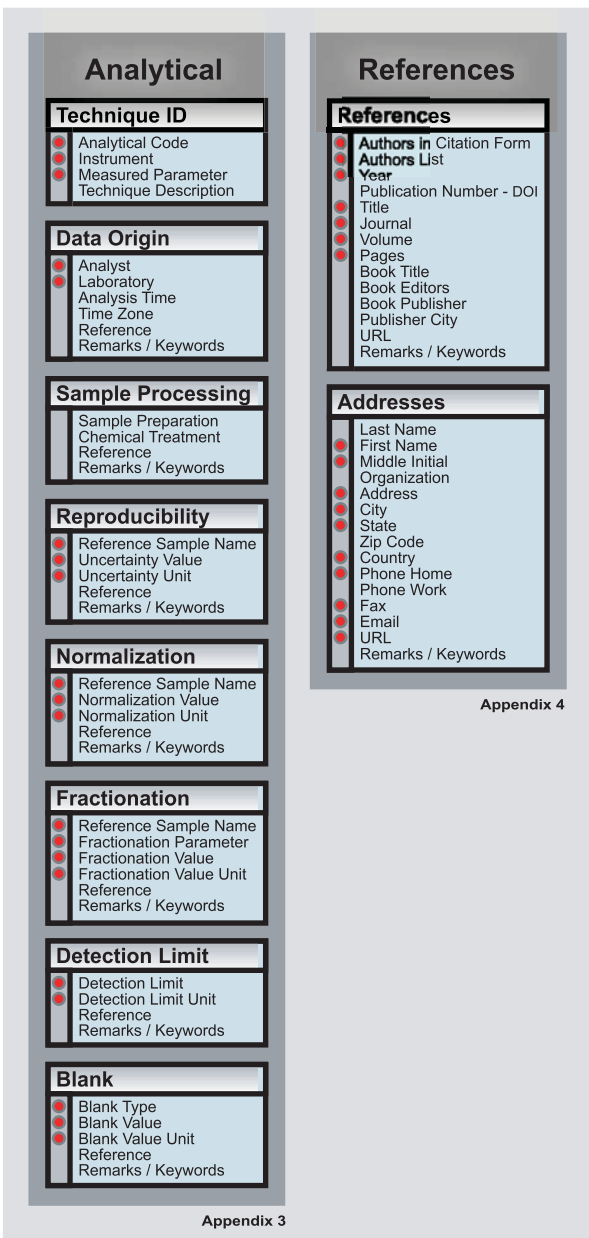


Figure 3. Metadata for analytical techniques, references and author/sampler/curator addresses. Essential metadata are flagged with red dots.

We suggest here to use the one oxidation state that is likely to be dominant in a particular chemical system but not to list FeO or Fe₂O₃ abundances based on the assumption of a fixed Fe²⁺/Fe³⁺ ratio. FeO and Fe₂O₃ abundances should be given only when the abundances for each valence states are determined analytically. If data are given as a bulk analysis, a “(t)” should be added to the chemical

symbol to indicate that the total Fe inventory is represented in this particular form.

4. Sample-Related Metadata

[14] We consider sample-related metadata for a range of commonly studied sample types from solids (including unconsolidated sediments) to fluids and gases. The sample-related metadata are illustrated in a summary diagram that lists all metadata types included in the rock, water and gas categories (Figure 2) and we present some examples in a spreadsheet environment (appendix B1, B2, and B3). These metadata describe the sample location, the sampling method, and the sample itself while the analytical metadata (see section 4.1; Figure 3) include information on the origin of the data, how the data were normalized, and the uncertainties of the data.

[15] The metadata required for different kinds of geochemical samples vary even though the categories of metadata remain remarkably similar (Figure 2). Solids may include igneous, metamorphic or sedimentary rock types, minerals and fossils. However, they can also be man-made such as experimental charges. Metadata fall into two logical groups, the first of which consists of information that is common to almost all sample types: sample identification, geographic data and sampling data. The second group of metadata consists of information that is unique to specific types of samples with respect to their sample description, classification or age. Many metadata categories provide for the names of scientists and references, in short citation form. Full addresses and references are given in separate address and reference files (Figure 3).

[16] The first metadata category is devoted to sample identification. The first entry in this category carries an alphanumeric sample identification (Figure 2). If there are several choices of sample numbers, the preferred choice should be the sample number that was assigned during the sampling process, because this relates to most of the other metadata provided here. The sample number is essential. A unique sample identifier can be derived from this number, combined with the location, sampling time and the scientist name. The second

and third entry may be used for a short sample description and laboratory number, respectively.

[17] Relevant geospatial information should be given for all samples in decimal degrees latitude and longitude and meters elevation above mean sea level. Negative values indicate southern latitudes, western longitudes and depths. Uncertainties are given in the same units. An effort should be made to provide the most accurate latitude and longitude information possible since this numerical location estimate is most essential to any location description. Conservative uncertainties should be assigned that give a realistic estimate to which extent a location estimate can be trusted. In some earth science communities UTM coordinates may represent acceptable alternatives to latitude and longitude notation. We discourage their use because computer handling of global location information is greatly simplified in units of latitude and longitude. If UTM coordinates are given, the zone must be provided. Specific national reference grids in topographic sheets cannot be used instead of latitude and longitude because they cannot be globally recast into latitude and longitude. They can be listed in an additional set of entries, if wanted. Location estimates may include a depth in a core or a land-section, but a reference has to be given to a description of the drill core or the profile taken. This reference may be an illustration in the same paper or in another publication. In addition, locations may be described by an appropriate geographic name (e.g., a mountain range, island, ocean basin, river, lake). Geographic names should avoid political terms if possible and be specific, but commonly known and identifiable in major geographic indexes. If there are commonly used English names for a mountain range or a country, the local name can be listed as well. A remark entry is added for any location related information, such as a description of the location, the sampling of a time-variant feature like an ash dusting collected after an eruption, an aerosol in an eruption cloud, or floodwater.

[18] Sampling data (Figure 2) include information related to the sampling processes, including time, collector, archiver, methods, sampling platform, keywords on sampling and citations related to the

sampling process. Such information is included in the metadata for two reasons: (1) to help with the interpretation of the data and to correlate samples from the same expedition or sampling method and (2) to assign uniform identifiers to samples that are analyzed by different laboratories and appear in different publications. The preferred format for the sampling time is decimal universal time (GMT) in the yyyy:mm:dd:hh.hh format. If for some reason local time is more appropriate (very rarely) the time zone must be specified remembering to consider daylight savings time. Sample time is considered essential for all samples because it is an important parameter for time series sampling and because it helps assign a unique sample identifier for a particular sample. The precision of sampling time may be chosen as appropriate for the likely uses of the data; for a rock sample, simply the year might be sufficient to indicate a particular sampling season, while for a stream water sample time may need to be precise to the hour. The three following entries include information about the sampling platform (e.g., ship/vessel/airplane name) and expedition name, the scientist responsible for the sampling effort, and the scientist in charge of sample curation. These data allow the tracing of the data to other related data, an archival facility, or an investigator to whom sample requests could be directed. For water samples, there are additional entries on the sample treatment during sampling, including filtration, biocides, acidification, the temperature of the water sample, and the sample container. Obviously, not all of these categories will apply to all samples and inapplicable categories may be omitted. The sampling methods may provide important information with respect to potential sources of sampling related blank problems (e.g., diamond drilling versus hammer sampling; Niskin bottle versus Ti-syringe; squeeze extraction of pore waters; filtration of particles). A citation entry can be used for published or web-available publications on the sampling method used. Remarks may give more details on the sampling.

[19] Some of the above data may not be recoverable for data publications on samples that were taken in the past or with less than ideal field characterizations. In those cases, metadata should be estimated if possible (e.g., latitude and longitude from “pre-

GPS” field surveys), but they should not be listed if they cannot be reconstructed with confidence.

4.1. Sample Characterizations for Rocks

[20] Rocks are defined here as any geological specimen. These include igneous, metamorphic and sedimentary rocks, unconsolidated sediments, meteorites, and even experimental charges. In this scheme, rocks may be characterized based on chemistry, metamorphic grade, grain size, texture, and in terms of their “setting”. The emphasis of descriptors in this metadata category should be the utility in a search for key rock groups, and not the need for classification in a specialist database. The main descriptor for rocks should be based on its chemistry/mineralogy, even if it is based only on a rough visual inspection. Even very rough terms such “basalt” (*sensu lato*, for any mafic extrusive) are acceptable, but it is better to use a more specific general term (hawaiite, picrite). Metamorphic and sedimentary rocks should also be primarily characterized using a term that includes the chemical composition but also by metamorphic grade and grain size characteristics, respectively. Any term that may be found in a general textbook may be used. Specialist classifications should be reserved for the “remarks” entry. Additional entries provide room for the description of igneous, metamorphic or sedimentary textures, and alteration. Alteration descriptions may include a general term for a low temperature alteration overprint, a percent estimate of the fraction of minerals replaced, or high temperature hydrothermal overprint. Tectonic setting or “reservoir” are keywords for whether a particular sample is relevant to the study of particular tectonic settings (mid-ocean ridges, arcs, continents) or particular geochemical reservoirs (core, mantle, continental crust). Multiple terms may be used if necessary. The purpose of these keywords is to help relating samples to a particular geological context, rather than assigning a particular reservoir or tectonic setting. Citations may be used for further descriptions of the rock type, and remarks may include specialist rock type descriptors, text descriptions of the sample or the relationship to a particular project.

[21] For mineral or fossil samples an additional category is added to the metadata. The mineral

characterization include essential descriptors such as mineral name and the type of occurrence (e.g., vein filling, groundmass phase, phenocryst, or xenocryst). Additional entries may be used to describe mineral shape (euhedral/anhydral), the size, and a citation and remarks.

[22] For fossils, essential descriptors include genus, species and preservation. Any common type of preservation index is acceptable, based on color, or based on extent of mineral (aragonite) replacement. In addition, there are entries for a citation and remarks or text descriptions.

[23] Rocks must also be characterized with respect to their age. This information is essential. The age may be given as an absolute age (with uncertainty), as a bio/magneto stratigraphic age, or both. An optional entry offers the opportunity to give a local stratigraphic unit. Every age has to be supported either with a reference (or some explanatory text) that allows judgment of the quality of the age data. In many cases, the age is well known and can be assigned based on high-precision geochronological data. However, even if the age is known only very roughly (such as 22 ± 20 Ma), such an age is still useful in searches over broad geological time periods.

4.2. Sample Characterization for Water Samples

[24] Water samples have similar metadata types for sample identification, geographic data and sampling data as geological samples (see descriptions above; Figure 2). In addition to geographic coordinates, sampling date, and physical properties such as temperature, river metadata should include discharge and suspended sediment information when possible. These can be added in a separate entry to the geographic data. Sampling data for waters should include filtration (type and size), acidification, biocides and temperature of sampling. Water samples are characterized with respect to water type (hydrothermal water, groundwater, pore water, seawater, precipitation, lake, river water, etc.) and a water classification that gives a more specific (or additional) description (e.g., black smoker, bottom water, North Atlantic Deep Water). Water samples are also characterized with a variety of physical and

chemical properties (hydrographic data), including pressure, potential temperature, density, and particulate contents (Nephelometer readings). Hydrographic data standards for seawater are well established, such as for the World Ocean Circulation Experiment (“WOCE”; *Swift and Diggs* [2001]). Such previous recommendations should be followed to a maximum extent possible, even though some reorganization in groupings may be beneficial in many cases. However, most other geochemical disciplines in water chemistry do not have the same level of organization as in oceanography and standards vary widely in the types of data and metadata reported. In absence of well described specific standards, we recommend establishment of standards along the lines of the description presented here.

4.3. Sample Characterization for Gas Samples

[25] Location estimates and much of the sampling metadata are similar to waters and rocks. Specific gas related sampling metadata include the sampling method, type of container, filtration, sampling pressure, temperature and humidity. Sample characterization includes a gas type for the general classification (natural gas, air, hydrothermal) and a gas characterization for the more specialized classification (plume, solfatara).

4.4. Exceptions

[26] There are data generation/collection efforts in geochemistry and cosmochemistry that do not fit into the above scheme. For example, meteorites are not usefully referenceable in a geospatial reference frame, and averaged data may integrate over a very large, even global, scale. In these cases, only some of the sample metadata categories will apply and other metadata may be omitted from this scheme.

5. Analytical Details

[27] Each geochemical parameter in a data column of Figure 1 is associated with an analytical code that links to analytical metadata in Figure 3. Examples for analytical metadata files are given in appendix C. These metadata provide information on the type of analytical techniques used, the origin of the data,

how samples are processed in the laboratory, how the parameters are represented, how they were determined, their uncertainties, and blanks. These data serve two main purposes: (1) to allow a database user to evaluate the analytical work, and (2) to provide guidance on the use of the data. Most of the analytical metadata should be considered essential, at least, all the information that is needed to recalculate data into a different notation or to reproduce a particular method of sample processing or analysis.

[28] The first metadata entry contains the analytical code given for each element in the main data table. The following entries in this section include the instrument type (using abbreviations listed in appendix D), the element or isotope (ratio) that this particular metadata set applies to and some general remarks about the technique. The next set of parameters includes information on the data origin, including the responsible analyst, the laboratory, the time of analysis, a citation regarding the laboratory and/or the analyst, and some remarks or keywords.

[29] Sample processing contains information regarding sample preparation before analysis. The first entry describes sample preparation, including in particular comments that help understand potential contamination issues. The second entry describes chemical treatments such as leaching, or ion exchange, drying or firing before analysis, and a column for remarks regarding sample treatment.

[30] The following section of analytical metadata is devoted to the reproducibility of the analysis for each geochemical parameter. Errors listed here should be exclusively “external” errors, that were determined on basis of repeat analysis of the reference samples. However, if internal errors are used (i.e., based on individual counting statistics for a particular sample) they should be given in the data table (Figure 1) as a separate column next to the respective column of analytical data.

[31] The next section is devoted to describe sample normalizations. Here, all information is compiled that is needed to transform data into different common normalizations. There are a variety of relevant normalizations that are considered in this section. Data may be normalized to 100%, relative to some specific reference samples, or to instrumen-

tal fractionation, such as due to mass fractionation. The first entry gives the reference sample name, the reference value to which the samples are renormalized. If data are normalized to 100% the original sums must be reported with the data in an additional data column so the absolute abundances can be retrieved for each element analyzed. There are also entries for a fractionation parameter, such as the mass fractionation its value and unit. These entries must contain all information necessary to allow a reader to re-normalize the published data such they fit any other commonly used CI concentration unit http://www.bipm.fr/enus/3_SI/base_units.html. All essential information has to be listed such that a “generalist” will be able to readily translate between different units used in geochemistry.

[32] Detection limits and blanks must be given for all analyses whereby the detection limit should be defined as the lowest calibrated value of an instrument. Blanks are described with a value, type of blank and its unit. Acceptable types of blanks include, for example, cumulate reagent blanks and procedural blanks. All blanks must be given so they can be correlated with a typical analytical procedure for the data reported.

6. References and Addresses

[33] In Figure 3 (appendix D), we have given a format for references that were used in the sample descriptions or analytical metadata, and for addresses as they may relate to personal references. For the references we have listed essential fields for the authors in citation form, authors, year, the Electronic Publication Identifier (DOI), the title, journal, volume and pages. Whenever appropriate, we also include information on a book, URL or remarks. For address information, we listed a minimum number of data that are necessary to get in touch with a person to get additional information, obtain sample splits and so on.

7. Concluding Remarks

[34] Any new comprehensive standardized format for data and metadata at first appears difficult and certainly labor intensive to implement. This is the

case particularly for data that are were obtained prior to an agreement on standards. Problems include conversion of metadata into new formats and extraction of metadata from maps and laboratory field notes. Many metadata may have been never collected or may not be recoverable at this stage. This should not prohibit publication of data because it is still better to have data with imperfect metadata than no data. Once data standards are established, their use in new studies does not impose any significant obstacle or additional burden. In fact, standard metadata sheets will be helpful as a checklist for comprehensive note keeping and as a filing system for metadata information gathered in the course of a study. Many steps in metadata acquisition and archival may be automated using laboratory instruments or handheld GPS receivers. For many studies, many metadata columns can be filled wholesale because the parameters do not change for the entire sample suite analyzed. It is obvious that metadata collection can be tedious, but it is also quite clear that they are one of the most efficient and transparent ways to keep track of information that is essential for scholarly scientific studies.

[35] We have compiled and organized geochemical data and metadata categories into a format that would allow effective publication of geochemical data in an electronic environment. We hope we have made the case that establishing such a format is an important step toward greatly improving data publication in geochemistry. We believe that adopting our formats will be a great step forward, while keeping the overall effort relatively small. Nevertheless, our key goal is not to impose a particular method of data publication but rather to contribute to the discussion of important technical publication issues. Resolution of these issues will have substantial impact on the scholarly quality of science publications and on the ease at which geochemistry will be able to take advantage of Information Technology. This paper is the result of discussions at the database subgroup session at the GERM 2001 workshop at La Jolla, CA. Internet access to this paper and future contributions to this topic can be found at "<http://earthref.org/metadata/GERM/>. Contributions or opinions to the metadata discussion are welcome.

Appendix A.

Table A1. Rock Sample Major Elements^a

Parameter	SiO ₂	Al ₂ O ₃	Fe ₂ O ₃ [t]	Fe ₂ O ₃	FeO	MgO	CaO	Na ₂ O	K ₂ O	H ₂ O [+]	TiO ₂	P ₂ O ₅	MnO	CO ₂
Unit	wt%	Wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%	wt%
Analytical code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
CY-19.7a	48.8	16.56	7.99	6.97	0.92	4.54	6.99	1.39	4.72	4.77	0.51	0.04	0.09	2.39
CY-19.7b	47.42	16.39	8			4.64	6.93	1.59	4.68		0.52	0.03	0.08	
CY-19.7c	47.1	16.2	7.57			4.5	6.56	1.49	4.58		0.48	0.02	0.09	
CY-23.2a	45.9	14.6	8.15	7.01	1.03	3.73	7.32	0.6	7.55	4.68	0.57	0.08	0.06	5.03
CY-23.2b	45.5	14.6	9	8.78	0.2	3.62	7.52	0	7.62	4.4	0.57	0.1	0.06	6.4
CY-23.3a	44.9	13.7	7.87			3.94	7	0.59	7.32		0.52	0.13	0.06	
CY-32.3b	47.5	17.2	8.48			7.07	3.28	1.42	4.18		0.58	0.03	0.08	
CY-32.3c	49.5	18.03	8.77	7.65	1.01	7.03	3.51	1.39	4.24	6.25	0.69	0.04	0.08	0.24
CY-32.3d	47.7	17.6	8.6	7.49	1	6.78	3.64	1.3	4.24	6.9	0.68	0.05	0.09	0.4
CY-32.3e	48.53	17.75	8.75			6.87	3.48	1.12	4.25		0.7	0	0.08	

^aExcel files of all appendix tables are available as auxiliary material in the HTML version of the article at <http://www.g-cubed.org>.

Table A2. Rock Sample Trace Elements

Parameter	Sc	Ti	V	Cr	Mn	Fe	Co	Co	Ni	Ni	Zn	Rb	Sr	Y	Sr	Nb	Cs	Ba
Unit	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
Analytical code	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
543A-10-3, 27–31	38.7	10440	357	284	2960	76650	37	39	67	69	90	23	130	37	112	6	0.4	35
543A-11-2, 34–36	37.8	10200	338	259	2680	74970	36	41	78	86	88	9	141	38	116	4	0.31	6.8
543A-12-2, 56–59	41.8	11880	391	258	3100	81340	35	38	65	70	92	18	135	46	134	5	0.5	10.3
543A-12-4, 99–101	38.6	9600	343	284	2400	70420	41	46	84	91	81	8	132	36	107	4	0.27	8.1
543A-13-2, 133–136	36.9	9300	327	268	2540	70050	41	41	95	96	93	5	128	36	111	5	0.1	
543A-13-5, 84–87	37.2	9300	328	258	2540	68950	40	44	90	100	80	5	130	35	107	5	0.19	2.8
543A-14-1, 33–36	37.5	8940	312	254	2680	68880	41	45	115	129	75	1	129	35	95	4	bdl	4.6
543A-15-3, 71–74	37.6	9060	323	233	2680	69580	44	48	106	121	74	1	127	35	95	4	0.02	bdl
543A-16-2, 34–37	37.4	9720	327	258	2820	73150	41	42	92	100	78	2	136	38	111	4	0.04	4.7
543A-16-3, 117–120	36.8	9120	337	283	2260	65380	49	53	107	115	90	1	131	31	96	4	0.26	

Table A3. Water Analyses

Parameter	Potential Temperature	Salinity	Sigma Theta	Turbidity	Si(OH) ₄	PO ₄	Cl	Ar	K
Unit	degrees Celsius	psu	bar	ntu	μmol/l	μmol/l	ppm	ppm	ppm
Analytical code	33	34	35	36	37	38	39	40	41
OR1998-XY000129	18.234	3.21	1.234	0.023			6.7		0.25
OR1998-XY000130	18.108	3.22	1.232	0.034	0.2	6.5	8		0.31
OR1998-XY000131	18.345	3.2	1.229	0.035		bdl	9.6		0.32
OR1998-XY000132	18.321	3.21	1.22	0.039		bdl	10.1	0.567	0.56
OR1998-XY000133	18.289	3.22	1.198	0.045			8.8	bdl	0.52
OR1998-XY000134	18.239	3.25	1.222	0.089			8.9	bdl	0.49
OR1998-XY000135	18.239	3.22	1.239	0.088			9.3		0.45
OR1998-XY000136		3.23	1.223	0.086			10.3	0.667	0.32
OR1998-XY000137	18.499	3.23	1.222	0.083			9.5		0.35
OR1998-XY000138	18.411	3.23	1.244	0.066			8		0.29

Table A4. Gas Analyses

Parameter	In situ Pressure	Air Temperature	Air Humidity	CO ₂	Methane	Nitrous Oxide	Xe
Unit	bar	°C	%	ppm	ppm	ppb	ppb
Analytical code	42	43	44	45	46	47	48
CL2001-AIR00287		3.21	34.2				
CL2001-AIR00288		3.22	33.1	10.5	458.9	23.9	505.72
CL2001-AIR00289	12.1	3.31	32.9				
CL2001-AIR00290	12.4	3.29	32.8				
CL2001-AIR00291		3.28	32.7				
CL2001-AIR00292		3.27	32.7			bdl	
CL2001-AIR00293		3.28	32.7			bdl	
CL2001-AIR00294		3.28	32.7				
CL2001-AIR00295		3.28	32.8				
CL2001-AIR00296		3.27	32.9				

Table A5. Radiogenic Isotopes

Parameter	Sr	Nd	K	Rb	Cs	87Sr/86Sr	87Sr/86Sr	143Nd/144Nd	143Nd/144Nd
Unit	p7pm	ppm	ppm	ppm	ppb	ratio	1S-SE	ratio	1S-SE
Analytical code	49	50	51	52	53	54	ERR	55	ERR
26-4#8A	14.09	5.99	1821	0.23	1.7	0.702949	0.000012	0.513109	0.000011
32-1#1L	6.3	9.15	830	0.254	12.1	0.702933	0.000011	0.513108	0.000015
36-2#1C	17.93	7	795	0.046	1.1	0.702902	0.000009	0.513111	0.000009
45-2#4B	9.85	13.22	1583	0.123	0.2	0.702945	0.000018	0.513112	0.000009
840-7#1B	13.88	9.55	18.5	0.011	0.5	0.70323	0.000017	0.513112	0.000007
87-6#12	41.11	8.46	804	0.427	3.7	0.703349	0.000012	0.513142	0.000011
39-4#3D	30.41	4.16	9.24	0.0061	0.6	0.703343	0.000011	0.513097	0.000009
43-4#8A	19.71	7.05	540	0.232	0.7	0.70305	0.000015	0.513112	0.000011
66-2#6A	42.5	4.04	80.5	0.016	0.5	0.703647	0.000009	0.5131	0.000009
67-2#5	18.54	8.57	120.1	0.015	0.3	0.703553	0.000007	0.513106	0.000018

Appendix B.

Table B1 (Representative Sample). Rock Samples^a (The full Table B1 is available in the HTML version of this article at <http://www.g-cubed.org>)

Sample ID			Geographical Data					
Sample Number	Sample Description	Laboratory Identifier	lat	lon	External Accuracy	Elevation	External Accuracy	Depth in Core or Section
text	text	text	decimal degrees	decimal degrees	decimal degrees	meters above mean sea level	meters	relative meters
ALL26-D1-4	glassy rind		33.35	-37.5	0.01	-3203	12	
AVON2-D26-5	core of massive basalt flow	HOW-5	0.82	-177.34	0.01	-4845	9	
SP73-34	whole rock split		47.122	8.889	0.001	2521	10	
ODP-185 1149D 10R1 44-48		185BC12	28.124	-160.876	0.001	-3444	10	-234.5
CY-1a 225-230	ophiolite drill core		42.231	12.523	0.001	456	10	435.5

^aNote bold columns headings are essential.

Table B2 (Representative Sample). Water Samples^a (The full Table B2 is available in the HTML version of this article at <http://www.g-cubed.org>)

Sample Id			Geographical Data					
Sample Number	Sample Description	Laboratory Identifier	Lat	Lon	External Accuracy	Elevation	External Accuracy	Depth in Core or Section
Text	text	text	decimal degrees	decimal degrees	decimal degrees	meters above mean sea level	meters	relative meters
PH2O-03 450-465	thermal boundary water		34.533	-142.443	0.001	-450		22
OR1998-XY000129		OR1998-788a	-23.288	132.334	0.1	-10		0.1
OR1998-XY000130			-23.288	132.334	0.0001	-10		0.1
BSEA-198			24.8	31.4	0.02	-156		10
LW2000-DBH11			69.2991	-89.8242	0.05	-9.23		11

^aNote bold columns headings are essential.

Table B3 (Representative Sample). Gas Samples^a (The full Table B3 is available in the HTML version of this article at <http://www.g-cubed.org>)

Sample Id			Geographical Data					
Sample Number	Sample Description	Laboratory Identifier	Lat	Lon	External Accuracy	Elevation	External accuracy	Depth in Core or Section
text	text	text	decimal degrees	decimal degrees	decimal degrees	meters above mean sea level	meters	relative meters
PH2O-03 450-465	natural gas		34.533	-142.443	0.001	-450		2
68AK009	troposphere		89.1	128.3	0.1	3600		2
QQ01733-a1	fumarole gas		0.2356	156.9664	0.0001	3459		10
CL2001-AIR00287	air sample		-12.29	-78.35	0.02	10234		10
CL2001-AIR00288	air sample		-12.3	-78.37	0.05	10289		20

^aNote bold column heads are essential.

Appendix C.

Table C1 (Representative Sample). Analytical Techniques^a (The full Table C1 is available in the HTML version of this article at <http://www.g-cubed.org>)

Technique				Data Origin			
Analytical Code	Instrument	Measured Parameter	Technique Description	Analyst	Laboratory	Analysis Time	Time Zone
number 1	text XRF	text SiO2	Text	text S.S. Smith	text University of Minnesota	yyyy:mm:dd:hr 1979:02:09:00	text gmt
5	ICPMS	MgO		J. Bakker	NIOZ, Texel, The Netherlands	1965:03:23:00	gmt
54	SSMS	87Sr/86Sr		W. Clinton	University of Colorado, Denver	1998:09:23:02.12	pacific
55	SSMS	143Nd/Nd144	run in oxide mode	W. Oliphant	University of Colorado, Denver	1998:09:23:02.34	pacific
106	AR-AR	40Ar/39Ar	CO2 laser used for incremental heating	P.P Armstrong	Geochronology Laboratory of the University Bristol	1996:10:10:23.34	gmt
112	MS	18O		G. Vlinders	University of Colorado, Denver	2001:12:01:00	pacific

^aNote bold column heads are essential.

Appendix D.

Table D1 (Representative Sample). References^a (The full Table D1 is available in the HTML version of this article at <http://www.g-cubed.org>)

Authors in Citation Form	Authors List	Year	Publication Number (DOI)	Title	Journal	Volume
text Rudnick & Fountain 1995	text Rudnick, R.L. and Fountain, D.M.	number 1995	text	text Nature and composition of the continental crust - a lower crustal perspective	text Reviews in Geophysics	number 33
Taylor & McLennan 1995	Taylor, S.R. and McLennan, S.M.	1995		The geochemical evolution of the continental crust	Reviews in Geophysics	33
Plank & Langmuir 1998	Plank, T. and Langmuir, C.H.	1998		The geochemical composition of subducting sediment and its consequences for the crust and mantle	Chemical Geology	145
Shaw et al. 1986	Shaw, D. M., Cramer, J.J., Higgins, M.D. and Truscott, M.G.	1986		Composition of the Canadian Precambrian shield and the continental crust of the Earth		
Wedepohl 1995	Wedepohl, K.H.	1995		The composition of the continental crust	Geochimica and Cosmochimica Acta	59
Weaver & Tarney 1984	Weaver, B.L. and Tarney, J.	1984		Empirical approach to estimating the composition of the continental crust	Nature	310

^aNote bold column heads are essential.

Table D2 (Representative Sample). Addresses^a (The full Table D2 is available in the HTML version of this article at <http://www.g-cubed.org>)

Last Name	First Name	Middle Initial	Organization	Address	City	State	Zip Code
text Arculus	text Richard	text J.	text Department of Geology, The Australian National University	text Universe Lane 8855	text Canberra	text	text ACT 0200
Blichert-Toft	Janne		Ecole Nationale Supérieure de Lyon	46 Allée d'Italie	Lyon		69364 Lyon Cedex 7
Derry	Louis	A.	Geological Sciences, Cornell University	Snee Hall	Ithaca	New York	NY 14853
Elliott	Tim		Department of Earth Sciences, Bristol University	Wills Memorial Bldg	Bristol		UK BS8 1RJ
Elthon	Donald		Department of Chemistry, University of Houston	Nobel Drive 1122	Houston	Texas	TX 77004
Francois	Louis		Institut d'Astrophysique, Université de Liege	5, Avenue de Cointe	Liege		B-4000

^aNote bold column heads are essential.

Appendix E.

Table E1. Instrument and Techniques

Method	Method Explanation
AAS	ATOMIC ABSORPTION
ALPHA	ALPHA COUNTING
ALPHA-ID	ALPHA COUNTING ISOTOPE DILUTION
ANC	ANION CHROMATOGRAPHY
AR-AR	⁴⁰ AR- ³⁹ AR AGE DETERMINATION
CALC	CALCULATED
CHN	CHN
CHN-G	CHN GAS CHROMATOGRAPHY
COL	COLORIMETRIC
COUL	COULOMETRICAL ANALYSIS
CSA	CARBON SULFUR ANALYZER
DCP	DIRECT CURRENT PLASMA
DROES	DIRECT READING OPTICAL EMISSIONS SPECTROSCOPY
EMP	ELECTRON MICROPROBE
ENAA	ENAA
ES	EMISSION SPECTROMETRY
FIS	FISSION TRACK
FL-ES	FLAME EMISSION SPECTROSCOPY
FPHOT	FLAME PHOTOMETRY
FTIR	FOURIER TRANSFORM INFRARED SPECTROMETRY
GC	GAS CHROMATOGRAPHY
GIO	GRADIENT ION- CHROMATOGRAPHY
GRAV	GRAVIMETRY
HPLC	HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY
HR-ICP-MS	HIGH-RESOLUTION ICP-MS
ICP	ICP
ICPAE	ICP ATOMIC EMISSION SPECTRO
ICPMS	ICP-MS
ICPMS-ID	ICP-MS ISOTOPE DILUTION
IGN	IGNITION
IMP	ION MICROPROBE
INAA	INSTRUMENTAL NEUTRON ACTIVATION ANALYSIS
IR-SP	INFRA-RED SPECTROSCOPY
ISE	ION SENSITIVE ELECTRODE
K-AR	K-AR AGE DETERMINATION
LA-ICP	LASER ABLATION ICP-MS
MANO	MANOMETRY
MC-ICP-MS	MULTI-COLLECTOR ICP-MS
MS	MASS SPECTROMETRY
MS-ID	ISOTOPE DILUTION MASS SPECTROMETRY
NAA	NEUTRON ACTIVATION ANALYSIS
NN	UNKNOWN
OES	OPTICAL EMISSION SPECTROMETRY
OPS	OPTICAL SPECTROSCOPY

Table E1. (continued)

Method	Method Explanation
PEN	PENFIELD METHOD
PMP	PROTON MICROPROBE
POT	POTENTIOMETRICAL
RA-TH	226RA-230TH AGE DETERMINATION
RNAA	RADIOANALYTICAL NEUTRON ACTIVATION
SIMS	SECONDARY ION MASS SPECTROMETRY
SMA	DUPONT SOLID'S MOISTURE ANALYZER
SP-PH	SPECTROPHOTOMETRIC
SPEC	SPECTROGRAPHIC ANALYSIS
SS-ID	SPARK SOURCE MASS SPECTROMETRY - ISOTOPE DILUTION
SSMS	SPARC SOURCE MASS SPECTROMETRY
TITR	TITRATION
U-PA	235U-231PA AGE DETERMINATION
U-TH	238U-230TH AGE DETERMINATION
U-TH-AGE	U-TH-HE AGE DETERMINATION
UV-ES	UV EMISSION SPECTROGRAPHY
VOL	VOLUMETRIC
WET	WET CHEMISTRY
XRF	X-RAY FLUORESCENCE
XRF-EDS	ENERGY-DISPERSIVE X-RAY FLUORESCENCE

Table E2. Errors and Uncertainties

Error	Error Explanation
1S-SE	1 sigma of the mean as a standard error
1S-SE-REL	1 sigma of the mean as a relative standard error
1S-SE-PERCENT	1 sigma of the mean as a relative standard error expressed in percentage
1S-SE-PERMIL	1 sigma of the mean as a relative standard error expressed in permil
1S-SE-PPM	1 sigma of the mean as a relative standard error expressed in parts per million
1S-SD	1 sigma of the mean as a standard deviation
1S-SD-REL	1 sigma of the mean as a relative standard deviation
1S-SD-PERCENT	1 sigma of the mean as a relative standard deviation expressed in percentage
1S-SD-PERMIL	1 sigma of the mean as a relative standard deviation expressed in permil
1S-SD-PPM	1 sigma of the mean as a relative standard deviation expressed in parts per million

Table E2. (continued)

Error	Error Explanation
2S-SE	2 sigma of the mean as a standard error
2S-SE-REL	2 sigma of the mean as a relative standard error
2S-SE-PERCENT	2 sigma of the mean as a relative standard error expressed in percentage
2S-SE-PERMIL	2 sigma of the mean as a relative standard error expressed in permil
2S-SE-PPM	2 sigma of the mean as a relative standard error expressed in parts per million
2S-SD	2 sigma of the mean as a standard deviation
2S-SD-REL	2 sigma of the mean as a relative standard deviation
2S-SD-PERCENT	2 sigma of the mean as a relative standard deviation expressed in percentage
2S-SD-PERMIL	2 sigma of the mean as a relative standard deviation expressed in permil
2S-SD-PPM	2 sigma of the mean as a relative standard deviation expressed in parts per million

Acknowledgments

[36] We thank Bruce Deck for his gas metadata suggestions, and Rick Carlson and Marjorie Wilson for insightful reviews and the National Science Foundation for funding this work (EAR0000998, DUE-027684, NSF 01-21684).

References

- GERM Steering Committee, Electronic data publication in geochemistry: A plea for “full disclosure,” *Geochem. Geophys. Geosyst.*, 2, Paper number 2001GC000234, 2001.
- Helly, J., H. Staudigel, and A. Koppers, Scalable models of data sharing in Earth sciences, *Geochem. Geophys. Geosyst.*, 4(1), 1010, doi:10.1029/2002GC000318, 2003.
- Swift, J. H., and S. C. Diggs, Improved exchange format for WOCE hydrographic program data, version 12 Sep. 2001. (Available at http://whpo.ucsd.edu/exchange/exchange_format_desc.htm).