

Data Citationと そのIUGONETへの展開について

能勢正仁¹、小山幸伸¹、村山泰啓²、木下武也²、
家森俊彦¹、石井守³、篠原育⁴、門倉昭⁵

1 京都大学理学研究科地磁気世界資料センター

2 情報通信研究機構・世界資料システム国際プログラムオフィス(WDS/IPO)

3 情報通信研究機構・電離層及び宇宙天気世界資料センター

4 宇宙航空研究開発機構・宇宙科学衛星世界資料センター

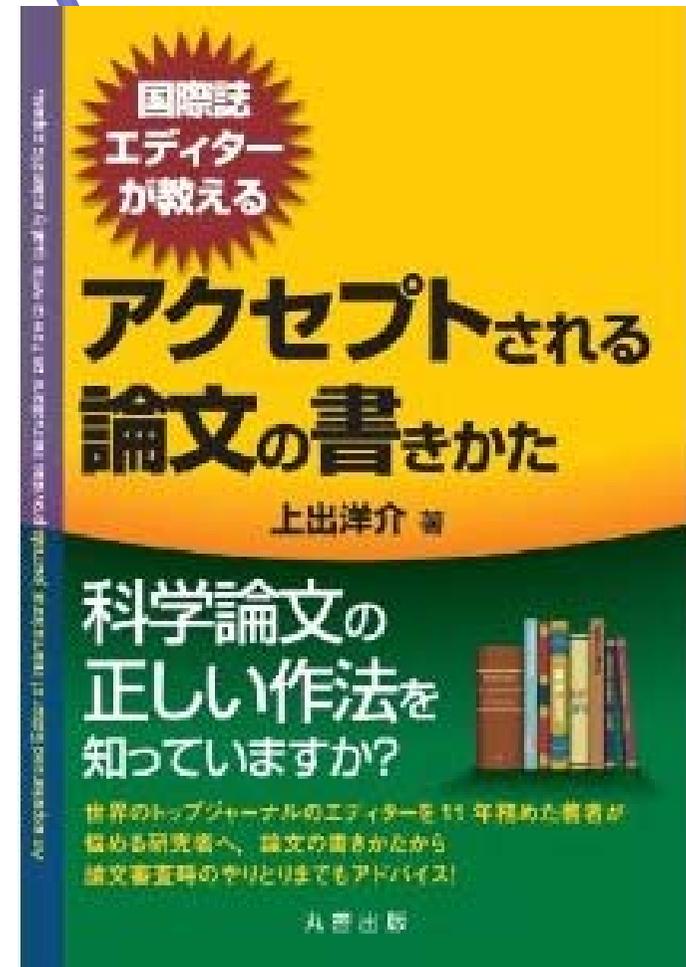
5 国立極地研究所・オーロラ世界資料センター

研究成果の発表

- ✓ 「論文発表は研究者の義務」
- ✓ 「研究の成果は論文で評価される」
- ✓ ...
- ✓ 納税者への還元
「とくに、大学や国公立の研究所では**国民の税金**を使って研究が行われているのですから、**研究成果を論文のかたちで報告することは義務にも等しいことなのです**。そのような研究活動を財政的に支えている国民への報告、あるいは還元が研究の基本要素です。」



- ✓ 国民の税金を使って得られたデータは国民の共有財産なのだから、データベースとして整理・保存し、広く一般に誰でも利用できるように整備することは義務にも等しいことなのです。



公文書・科学データ

✓ 公文書等の管理に関する法律

(平成二十一年七月一日法律第六十六号) … [福田首相の置き土産]

- 第一条 この法律は、国及び独立行政法人等の諸活動や歴史的事実の記録である公文書等が、健全な民主主義の根幹を支える国民共有の知的資源として、主権者である国民が主体的に利用し得るものであることにかんがみ、国民主権の理念にのっとり、公文書等の管理に関する基本的事項を定めること等により、行政文書等の適正な管理、歴史公文書等の適切な保存及び利用等を図り、もって行政が適正かつ効率的に運営されるようにするとともに、国及び独立行政法人等の有するその諸活動を現在及び将来の国民に説明する責務が全うされるようにすることを目的とする。

✓ 公文書だけでなく、(科学)データも「国民共有の知的資源」と捉えられるべき。

✓ NASAのResearch Announcement

- “Astrophysics Data Analysis, Solicitation: NNH14ZDA001N-ADAP”
Over the years, NASA has invested heavily in the development and execution of an extensive array of space astrophysics missions. …… To capitalize on this invaluable asset and enhance the scientific return on NASA mission investments, the Astrophysics Data Analysis Program (ADAP) provides support for investigations whose focus is on the analysis of archival data from NASA space astrophysics missions.

研究活動・データベース整備活動に対する評価

- ✓ 研究活動は、論文に伴う客観的な数値により評価されやすい。
- ✓ データベース整備活動は、これまでは客観的な評価基準が不明確だった。
 - せいぜい、acknowledgementに記載されるだけ。
- ✓ データベースに対しても、論文出版のような考え方・仕組みが普及すれば。。。
- ✓ doiの導入によって、それが可能な段階に移りつつある。
 - doi=digital object identifier (デジタルオブジェクト識別子)
 - International DOI Foundation (国際DOI財団)が運営するサービス。
 - **本文到達性**・・・Handle Systemを利用。
 - **持続性**・・・永続する管理、運営ルール。
 - **一貫性**・・・識別子が指すものの保証。

	能動的	受動的
研究活動	論文数 [n編]	論文被引用数 [h指数, g指数]
データベース整備活動	データベース数 (Data Publication) [n個]	データ被引用数 (Data Citation) [h指数, g指数]

学術論文に付与されたDOI識別子

SPACE WEATHER, VOL. 10, S08002, doi:10.1029/2012SW000785, 2012

Wp index: A new substorm index derived from high-resolution geomagnetic field data at low latitude

M. Nosé,¹ T. Iyemori,¹ L. Wang,² A. Hitchman,² J. Matzka,³ M. Feller,⁴ S. Egdorf,⁴ S. Gilder,⁴ N. Kumasaka,⁵ K. Koga,⁶ H. Matsumoto,⁶ H. Koshiishi,⁶ G. Cifuentes-Nava,⁷ J. J. Curto,⁸ A. Segarra,⁸ and C. Çelik⁹

Received 27 February 2012; revised 6 June 2012; accepted 7 June 2012; published 1 August 2012.

[1] Geomagnetic field data with high time resolution (typically 1 s) have recently become more commonly acquired by ground stations. Such high time resolution data enable identifying Pi2 pulsations which have periods of 40–150 s and irregular (damped) waveforms. It is well-known that pulsations of this type are clearly observed at mid- and low-latitude ground stations on the nightside at substorm onset. Therefore, with 1-s data from multiple stations distributed in longitude around the Earth's circumference, substorm onset can be regularly monitored. In the present study we propose a new substorm index,

Citation: Nosé, M., et al. (2012), Wp index: A new substorm index derived from high-resolution geomagnetic field data at low latitude, *Space Weather*, 10, S08002, doi:10.1029/2012SW000785

データベースに付与されたDOI識別子

- ✓ 氷河時代の海面温度proxyデータ → doi:10.1594/PANGAEA.127383

de Vernal, A et al. (2004) x
doi.pangaea.de/10.1594/PANGAEA.127383

Not logged in (log i)

PANGAEA®
Data Publisher for Earth & Environmental Science

Always quote citation when using data!

Data Description

Citation: de Vernal, Anne; Eynaud, Frédérique; Henry, Maryse; Hillaire-Marcel, Claude; Londeix, Laurent; Mangin, Sylvie; Matthiessen, Jens; Marret, Fabienne; Radi, Taoufik; Rochon, André; Solignac, Sandrine; Turon, Jean-Louis; MARGO, SST (2004): Compilation of dinoflagellate cyst LGM SST data.
doi:10.1594/PANGAEA.127383.
In: Barrows, TT et al. (2011): Various paleoclimate proxy parameters compiled within the MARGO project. doi:10.1594/PANGAEA.760904

Project(s): Multiproxy Approach for the Reconstruction of the Glacial Ocean surface (MARGO) 🔍

Coverage: Median Latitude: 0.000000 * Median Longitude: 0.000000 * South-bound Latitude: none * West-bound Longitude: none * North-bound Latitude: none * East-bound Longitude: none

Event(s): MARGO_0000 🔍 * Device: Literary studies 🔍

License: Creative Commons Attribution 3.0 Unported

Size: 29.0 kBytes

Download Data

Download dataset

Data Citationの実例

ソリュートレ仮説の検証

2008 *Journal of the North Atlantic* 1:85–98

The Solutrean Atlantic Hypothesis: A View from the Ocean

Kieran Westley^{1,2,*} and Justin Dix³

Abstract - One current hypothesis for the Pleistocene peopling of the Americas invokes a dispersal by European hunter-gatherers along a biologically productive “corridor” situated on the edge of the sea-ice that filled the Atlantic Ocean during the Last Glacial Maximum (LGM). In this paper, we assert that critical paleoceanographic data underpinning this hypothesis has not yet been examined in sufficient detail. To this end, we present data which show that the corridor may not have existed, and that, if it did, its suitability as a migration route is highly questionable. In addition to demonstrating that the hypothesized migration was unlikely, this highlights the importance of integrating paleoceanographic and archaeological data in studies of paleo-coastal societies.

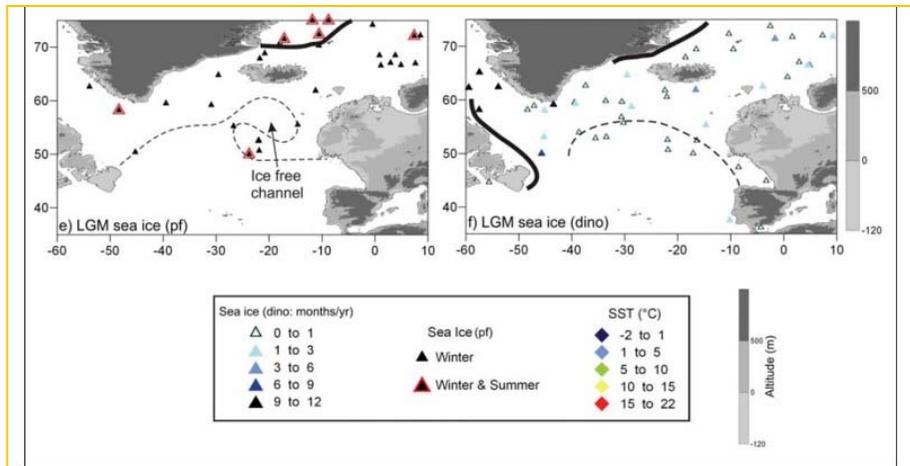


Figure 2. Quantitative reconstructions of LGM North Atlantic paleoceanography based on two different proxies: dinocysts and planktonic foraminifera. Data from De Vernal et al. (2004) and Weinelt (2004) (See also De Vernal et al. 2006, Kucera et al. 2005). a) Summer SSTs from planktonic foraminifera. b) Winter SSTs from planktonic foraminifera. c) Summer SSTs from dinocysts. d) Winter SSTs from dinocysts. e) Sea-ice extents from planktonic foraminifera: triangles show core sites with evidence of summer and winter ice. Heavy black line represents the extent of perennial ice, and dashed line is the maximum extent of winter ice (based on Sarnthein et al. 2003). f) Sea-ice extents from dinocysts: triangles show core sites with evidence of the duration of ice in months per year. Heavy black line represents the extent of perennial ice, and dashed line is the maximum extent of winter ice (based on De Vernal et al. 2006).

References

De Vernal, A., and C. Hillaire-Marcel. 2000. Sea-ice cover, sea-surface salinity, and halo-/thermocline structure of the northwest North Atlantic: Modern versus full glacial conditions. *Quaternary Science Reviews* 19:65–85.

De Vernal, A., and T. Pedersen. 1997. Micropaleontology and palynology of core PAR87A-10: a 23,000 year record of paleoenvironmental changes in the Gulf of Alaska, northeast North Pacific. *Paleoceanography* 12(6):821–830.

De Vernal, A., F. Eynaud, M. Henry, C. Hillaire-Marcel, L. Londeix, S. Mangin, J. Mattheissen, F. Marret, T. Radi, A. Rochon, S. Solignac, and J.-L. Turon. 2004. MARGO (SST) unpublished data: Compilation of dinoflagellate cyst LGM SST data. doi: 10.1594/PANGAEA.127383. World Data Center for Marine Environmental Sciences (WDC-MARE), Publishing Network for Geoscientific and Environmental Data (PANGAEA). Available online at <http://www.pangaea.de/>. Accessed June 2006.

De Vernal, A., F. Eynaud, M. Henry, C. Hillaire-Marcel, L. Londeix, S. Mangin, J. Mattheissen, F. Marret, T. Radi, A. Rochon, S. Solignac, and J.-L. Turon. 2005. Reconstruction of sea-surface conditions at middle to high latitudes of the Northern Hemisphere during the Last Glacial Maximum (LGM) based on dinoflagellate cyst assemblages. *Quaternary Science Reviews* 24:897–924.

Fiedel, S.J. 1992. Prehistory of the Americas. Cambridge University Press, Cambridge, UK.

Fiedel, S.J. 1999. Older than we thought: Implications of corrected dates for dates for Paleoindians. *American Antiquity* 64(1):95–115.

Fiedel, S.J. 2000. The peopling of the New World: Present evidence, new theories, and future directions. *Journal of Archaeological Research* 8(1):39–103.

Fladmark, K. 1979. Routes: Alternate migration corridors for early man in North America. *American Antiquity* 44(1):55–69.

Gamble, C., W. Davies, P. Pettitt, L. Hazelwood, and M. Richards. 2005. The archaeological and genetic foundations of the European population during the Late Glacial: Implications for “agricultural thinking.” *Cambridge Archaeological Journal* 15(2):193–223.

Heaton, T.A., S.L. Talbot, and G.F. Shields. 1996. An Ice Age refugium for large mammals in the Alexander Archipelago, Southeastern Alaska. *Quaternary Research* 46:186–192.

Hemming, S.R. 2004. Heinrich events: Massive Late Pleistocene detritus layers of the North Atlantic and their global climate imprint. *Review of Geophysics* 42:RG1005.

Henshaw, A. 2003. Polynyas and ice-edge habitats in cultural context: Archaeological perspectives from

Building a Culture of Data Citation





Data Citationを推進するには？

- ✓ データ提供者・データセンターへのメリット説明
 - 引用doi数を統計処理することにより、データ作成やデータセンターの寄与を明らかにできる。
 - データ作成やデータセンターの“visibility”, “professional recognition”, “reward”を向上させることになる。(被論文数と同様の意義)

- ✓ データ利用者へのメリット説明
 - データにアクセスしやすくなる。
 - doiをたどっていくことにより、これまでに存在を知らなかったデータを発見する機会が増え、研究の発展が期待される。

- ✓ 出版社・学会への働きかけ
 - データ引用の推奨
 - 論文投稿のルールにおける義務付けなど

- ✓ IUGONET・WDSのような集団的データ提供者からのdoi付与一斉開始

データにdoiを付与するには？

- ✓ 国際DOI財団が指定した“Registration Agency”に対して、doiを登録する。
- ✓ 日本では、Japan Link Center (JaLC)がRA業務を請け負っている。
- ✓ データセンターが用意しなければならないもの
 - データベース
 - データベースへのアクセス入り口となるLanding Page (URL)
 - doi (suffix)
 - 上記の情報が示されたメタデータファイル
- ✓ doi登録をJaLCに申請する (メタデータファイルの送付)
- ✓ URL-doiマッピングのメンテナンス



IUGONETにおけるdoiの展開について

- ✓ doi prefixを取得
 - 年間2-30万円
- ✓ メタデータスキーマの拡張
 - IUGONET 1.0.3 (現時点)
 - doi入力項目を含める
- ✓ Landing pageの用意
- ✓ メタデータスキーマ間の変換
 - IUGONET MDスキーマ ↔ JaLCが指定するMDスキーマ
- ✓ 「新たな手間・人力の消費」と「doi導入に伴うメリット」のトレードオフ

IUGONET
数値データセット NumericalData カテゴリーのメタデータの例

```

<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet href="http://www.iugonet.org/catalogs/schema/iugonetxsd.xsd" type="text/xsl"/>
<NumericalData>
  <ResourceID>
    <ResourceID>
      <ResourceHeader>
        <AccessInformation>
          <InstrumentID>
            <MeasurementType>
              <TemporalDescription>
                <ObservedRegion>
                  <SpatialCoverage>
                    <Parameter>
          ...
          
```

Resource ID
全研究リソースに対して unique に割り振られる URI 形式の ID
space://IUGONET/NumericalData/STEL/SuperDARN/HOK/HFradar/sd_hok_common_ergscdf

Resource Header 部
リソースの名前、説明、発行日時、コンタクト情報など

Access Information 部
実データベースの場所、データファイルのフォーマット、アクセス権限などの情報

InstrumentID, MeasurementType
このデータを取得した観測器に関する情報。観測器自体のメタデータへリンクするようになっている

Temporal Description 部
データが存在する時間範囲、またデータの時間分解能に関する情報

Observed Region
観測領域に関する情報。この場合は地球電離圏のF層

Spatial Coverage 部
観測領域の範囲の緯度・経度・高度等の情報(時刻→観測範囲での検索に利用)

Parameter 部
データとメタデータに関する情報。パラメータ

1



まとめ

- ✓ データベースに対しても、“Data Publication”, “Data Citation”といった論文出版・引用と同様の仕組みが考えられる。
- ✓ doiの導入によって、それが実現可能な段階に移りつつある。
(一部では実現が始まっている。)

- ✓ 特に、データ提供者・データセンターへのメリットは大きいと思われる。
 - “visibility”, “professional recognition”, “reward”

- ✓ IUGONET・WDSからのdoi付与一斉開始はデータ引用文化の潮流を作るうえで重要(?)。
 - 「新たな手間・人力の消費」と「doi導入に伴うメリット」のトレードオフ
 - doi prefixを取得 (年間2-30万円)
 - メタデータスキーマの拡張 (doi入力項目を含める)
 - Landing pageの用意
 - メタデータスキーマ間の変換 (IUGONETスキーマ ↔ JaLC指定スキーマ)