Why choose this one? Factors in scientists' selection of
bioinformatics tools

        Joan C. Bartlett <#authors>, Yusuke Ishimura <#authors> and
        Lorie A. Kloda <#authors>
        McGill University, School of Information Studies, 3661 Peel St.,
        Montreal, QC, Canada H3A 1X1

Abstract

    Purpose. The objective was to identify and understand the factors
    involved in scientists' selection of preferred bioinformatics tools,
    such as databases of gene or protein sequence information (e.g.,
    GenBank) or programs that manipulate and analyse biological data
    (e.g., BLAST).
    Methods. Eight scientists maintained research diaries for a
    two-week period, and were then interviewed following a
    semi-structured interview schedule.
    Analysis. The diaries and interview transcripts were analysed
    using a content analysis approach to reveal the factors that
    affected the selection of the bioinformatics tools the scientists
used.
    Results. Some of the factors (e.g., ease of use, familiarity),
    were similar to those identified with respect to text-based,
    bibliographic resources, while others (e.g., interface, scalability)
    were specific to the bioinformatics domain. Particularly interesting
    was the variation in how a single factor was defined. Often what was
    preferred by one group of users was not preferred by another.
    Conclusions. The identification of the broad, and sometimes
    contradictory, range of factors preferred by scientists has several
    implications. These include the need to design and develop tools to
    accomodate all users, (e.g., with multiple interface options), and
    to devise means of recommending or selecting tools on the basis of
    preferred factors.

        Introduction

Among the challenges facing information seekers is to know which source or channel of information to use in order to meet their information needs. In the domain of traditional, text-based information, this task has historically been facilitated by information professionals such as reference librarians or archivists, who rely on their expertise as well as a range of resources such as published reviews and subject guides to support their evaluation, selection and recommendation of resources. For example, within the domain of health sciences information, there are defined and agreed-upon criteria for evaluating the strength of research evidence and in selecting information resources for use in clinical practice (e.g., Haynes 2006 <#hay06>). But, what happens in domains for which there are no established criteria for assessment? How do people choose which information sources to use?

One such area is within the realm of bioinformatics, which has been defined as '/the computer-assisted data management discipline that helps us gather, analyse, and represent [biological] information/' (Persidis 1999 <#per99>: 828). The advent of bioinformatics, enabling the collection and analysis of vast amounts of biological information, has had a tremendous impact on all areas of biomedical research. It has also led to a vast, new array of information resources. In addition to the already overwhelming range of print and electronic books, journals, bibliographic databases and Websites, biologists must consider bioinformatics tools among the information resources relevant to their work. These tools generally consist of databases of primary biological data, software to manipulate or analyse such data, or a combination of both. Unlike traditional, bibliographic resources in which primary and secondary information are generally found in different and distinct resources, bioinformatics tools often integrate these two types of information in one source. Compounding this challenge in choosing tools is the ever-growing number of tools appearing in this field. According to the 2010 edition of the /Nucleic Acids Research/ annual database issue, there were over 1,200 individual bioinformatics tools (Cochrane and Galperin 2010 <#coc10>). In contrast, in 2000 the number was approximately 200 (Baxevanis 2000 <#bax00>). These figures reflect only those tools that are freely and publicly accessible; therefore the actual total is underestimated. Many of these tools accomplish similar tasks, raising the questions: why are there so many duplicate tools and how do scientists select their preferred tools among these?

In recent years a small body of literature regarding users of bioinformatics tools has developed. Scientists have reported that bioinformatics tools are important and frequently used (Grefsheim /et al/.//1991 <#gre91>; Yarfitz and Ketchell 2000 <#yar00>). These studies considered which tools were used and included bioinformatics tools in the same context as bibliographic resources. Past research, however, has not addressed the question of how or why particular resources were selected. The oral tradition with respect to the use and application of bioinformatics tools has led to a lack of formal documented information about the selection and use of tools (Bartlett 2005 <#bar05>; Brown 2005 <#bro05>; Haines /et al/. 2010 <#hai10>). Another approach to studying bioinformatics has been to consider the tasks for which the tools are used. Researchers have analysed and classified the various types of

tasks in bioinformatics (Stevens /et al/. 2001, <#ste01>Tran /et al/. 2004 <#tra04>), but did not address the selection of tools to accomplish such tasks.

Objectives

The objective of this research was to understand the factors that people use to distinguish among bioinformatics tools and to identify those features that are valued or preferred when a resource is selected.As the initial phase in a larger, ongoing research study, the approach was exploratory, with the intent of identifying as many factors as possible.Later stages of the research have the objective of refining our understanding of these factors, by identifying, for example, which are more important to scientists and also designing a system to utilise these factors to support scientists' decision making.

Background

This research is framed within the broader context of information behaviour,defined by Wilson as '/those activities a person may engage in when identifying his or her own needs for information, searching for such information in any way and using or transferring that information/' (Wilson 1999 <#wil99>: 249). One element of information behaviour, relating to information seeking, is the selection of preferred information sources and channels and it is this element that is foundational to our study.

Empirical studies of the selection of resources by life scientists have identified characteristics of resource selection. Currency was found to be a key factor (e.g., Grefsheim /et al./ 1991 <#gre91>), as was ease of access (e.g., Curtis /et al./ 1993 <#cur93>, 1997 <#cur97>). Information behaviour, including resource selection, was seen to vary depending on the discipline, task and experience of the scientist, with characteristic patterns associated with scientists of similar disciplines or experiences (Palmer 1991a <#pal91>, 1991b <#pab91>; Rolinson /et al./1996 <#rol96>, Rolinson /et al./ 1995 <#rol95>).

In a usability study of the set of bioinformatics tools at the National Center for Biotechnology Information, it was found that the persona of the participant (novice or expert) correlated with the assessment of the usability of the Center's tools (Javahery /et al./ 2004 <#jav04>). Experts were found to be satisfied with the suite of tools, while novices found the learning curve too steep.

Hogue (2001 <#hog01>) highlighted interesting factors affecting the selection of bioinformatics tools: the emphasis on either the algorithm itself, or its implementation. He argued that while many resource developers in bioinformatics place emphasis on the algorithm that underpins a tool and focus their attention on improving and developing the algorithm, the true value of a resource may actually lie in how well it can be implemented in terms of elements such as its scalability.

Bottomley (1999 <#bot99>) proposed a set of criteria, many of them system-related, for evaluating bioinformatics tools. However, they were not empirically determined or tested.

Methods

Recruitment

Purposeful sampling was used to select information-rich cases (Patton 2002 <#pat02>). Given that there are a range of people who use bioinformatics tools to support their work and that past research has found that both task and domain expertise are factors in the selection of information resources, we anticipated these might also be factors influencing our findings. As such, our sampling targeted participants with backgrounds in the different disciplines relevant to bioinformatics (e.g., biology, computer science), as well as those with a range of experience, from new graduate students to post-doctoral fellows or principal investigators.

Potential participants were identified using snowball sampling (Patton 2002 <#pat02>) and were contacted by email or telephone. Those who showed interest in participating in the research were asked to identify other potential participants. Recruitment and data collection continued until we reached data saturation.

Participants

We collected data from eight participants with a range of backgrounds, experiences and research tasks. All had used bioinformatics tools for at least one to two years and most for over five years. In general, they reported being satisfied or ambivalent about the tools they had used recently.Their research areas represented a range of sub-disciplines within biology.Three had their research centred on laboratory biology, for which bioinformatics tools were periodically used as part of their work, but were by no means the focus of the work.For another three, the use of bioinformatics analysis was central to their research.They were bioinformaticians working in silico, using and adapting bioinformatics tools to address a biological problem.The final two participants were computer scientists who designed and developed bioinformatics tools.

Data collection

Data were collected using the journal-interview method (Creswell 2008 <#cre08>) which was composed of two steps: 1) research journals and 2) individual interviews. Over a two week period, each participant filled in a research journal documenting each time they used a bioinformatics tool. Other information solicited through the research journal included: the name of the tool selected, the purpose of the activity, reasons for selecting the tool and evaluation of the usefulness of the selected tool. We conducted individual interviews after participants had

completed and returned their research journals. The journals were used as a starting point for interview questions, as they allowed participants to discuss recently selected bioinformatics tools and to elaborate on comments and judgements previously recorded. Sample questions from the interview appear below:

* What is the role of the tools you selected for your research?
* What were the criteria in selecting the tools?
* What is your opinion about the tool in terms of usability?
* What is your opinion about the tool in terms of quality?
* If you could select the perfect tool for each task, what criteria would be important to you?

Data analysis

The intent of this first phase of our study was to identify the range of factors people considered (or identified as important) in their selection of tools and the breadth and diversity of how they defined and described the factors (i.e., the range of factors identified and the variability in how they were defined).As such, our data analysis took a very broad, open approach to capture this diversity. Three researchers openly coded each transcript to identify potential factors (both those that were perceived positively and negatively). We then revised these codes after reviewing one another's analyses, to create a more cohesive, refined list of potential factors. This process was repeated several times and the final list of factors was compared against several transcripts to ensure they reflected what participants said.

Findings

We gathered factors from discussing over thirty different and diverse bioinformatics tools specifically (by name) as well as tools in general.Several factors were identified by the participants, describing not only which factors they considered important in their choices of tools, but also how each person defined a given factor within their own particular context.A recurrent issue was that while participants might use the same name or label for a factor, it often held different meanings to different people.As an example, the ability to customise an analysis could mean the access to modify the source code, or a range of options presented in pull-down menus.

The boundaries between the factors were often fuzzy and unclear, with factors overlapping, or representing different aspects of the same characteristic.However, the reported factors did cluster around larger themes and we report them here in four broad categories:system-related factors, functionality, quality and personal preferences.The factors in the first three categories tend to be more objective, while those in the last category of personal preferences are more subjective.

System-related Factors

These are factors that relate to the computer system itself, including the platform on which the tool runs, whether it is locally installed or Web-based, the interface and cost.

Platform

There were several parameters relating to the platform on which a tool was built and run.Those who preferred the ability to write their own code had personal or system dependant preferences for programming language (linked to their preferred or known language) and operating system (depending on their own computer system).Open source software was preferred by some over proprietary systems because of the ability to both download the software and run it locally, as well as to modify or customise it as needed. However, there was no guarantee that they would receive necessary support from tool providers in a timely manner.

> The reason I chose it was, well, A, because it was free, B, because it was open source so I could modify the source code and C, it was written in Perl, which is a language I was familiar with.(P1)

Another factor relating to platform was whether the tool was locally installed or Web-based, with some participants favouring local and others preferring Web-based.Among the reasons given for those preferences were issues of the size of either the software or the data in relation to the available computer resources.Speed and availability of internet access was another factor.

> There is more data than most people's computer can hold on it... would be like twenty gigabytes to download and then you would have to write programs to search through it. Whereas they have a Web tool... and it will spit out a table for you. (P4)

> But, the problem is that it is online, so it's not accessible to me all the time. So if I go back home... and have a slow internet connection, I can't use it. I need everything to be local, so I can rerun analyses. (P2)

Interface

Participants expressed strong, but contradictory, preferences for their choice of interface format.Some wanted a graphical user interface, with features such as pull-down menus, tabs and pre-selected conditions for the analysis,

> Primer 3 Plus there is a drop down menu which says I want to do this, or I want to do that and then it fills in all the fields for you.(P4).

Others preferred a command line interface, with the user in complete control of how the analysis was specified and executed and the option existed to write customised scripts.

> It's handy, the command line... what we do, it needs to be
> scriptable.You need to be able to do things in large batches.And I'm
> used to working on a Unix command line... I don't really like point
> and click.(P2).

However important the interface was, its features should not come at the
expense of the overall function of the tool,

> Some of the other ones have more bells and whistles that look
> cooler, but that have less actual information.It's [the UCSC Genome
> Browser] less pretty than some of the other ones, but it's more
> functional. (P4)

Among those involved in the design and development of tools, one
expressed the opinion that the interface was almost an afterthought,
something added to the tool at the very end, instead of being integral
to the entire development process, '/And then the last step is always to
make it a nice interface... so that it can be used publicly/.' (P7)

Cost

Cost was a factor in the choice of tool, with preference given to free
tools.Apart from general thriftiness, participants were not always in
the position to make purchasing decisions for their lab, or may have
been reluctant to invest money on a tool with limited applications.

> Important to me would be ideally if it is free. Because most
> software that costs something, costs a substantial amount of money
> and our labs have limited budgets and it's not worthwhile for use to
> spend a big amount of money for a tool that I'm going to use a
> couple of times. Because, a lot of times when I use these tools,
> it's for a very specific questions and it's never used again. (P3)

Others considered the costs of a tool worthwhile if the tool was
superior to the free option in terms of quality or functionality.One
noted that it wasn't a personal expense,

> But I guess cost, I am most for fully accessible tools... 'cause
> every time I need a particular program, the money didn't come out of
> my pockets so it didn't hurt as much. (P6)

Functionality

Functionality factors relate to what the tool does and how it works.

Function

What a tool actually does is a factor in the choice of tool.Participants
indicated that they wanted something that did what they needed, which
was not always as straightforward as it might appear.

First of all it has to be able to do what I need it to do. And
that's not trivial. There's a lot of tools out there that will do
something sort of like what you want it to do, but not exactly. So
you end up having to jury-rig your data to retrofit it.(P4)

    Customisability

The ability to customise or personalise the analysis, output, or
interface was an important factor.Some participants favoured open source
software, which supported and enabled the modification of source code
and allowing them to optimise the performance of the software for their
specific needs.Correspondingly, it was important that the software be
downloadable and installed locally. This supported the ability to modify
the source code.

    If you're waiting for the Website, or they're changing things with
    it, or there's a bug in it, you know, you can't do anything. You
    have to sit and wait. I primarily write my own software and use
    little bits of UCSC [Genome browser <http://genome.ucsc.edu/>] and
    so on. (P2)

The ability to script or customise the analysis was also linked to the
scalability and ability to automate the analysis.Participants referred
to writing a script to support the automated analysis of large batches
(sometimes in the order of tens of thousands) of data, rather than
having to /point and click/ for each one individually.

    ...point and click, is generally not as easy to make into batch type
    stuff... so you can easily put that into a shell script and it will
    do the analysis... you can leave it to run overnight and come back
    in the morning and it's done.(P2)

For others, customizability was defined in terms of the interface.A
customizable interface was one that provided flexible options for things
such as setting the parameters of a search or analysis, specifying the
output format, or modifying display settings.Essentially, the interface
should provide the user with the range of options available through the
tool and the ability to easily identify and select the options of
choice, thus supporting a customised search or analysis,

    [In] Primer 3 Plus there is a drop down menu which says I want to do
    this, or I want to do that and then it fills in all the fields for
    you.(P4)

    Scalability

The concept of scalability relates to whether a tool can accommodate
very large data sets.Assessments of scalability included whether the
dataset could be accommodated on a local computer, the ability to
analyse multiple samples at once and the ability to batch process large
amounts of data (e.g., tens of thousands of input sequences).

Speed

Speed of analysis was specified as desirable factor in a tool, although
definitions of /speedy/ varied dramatically from fractions of seconds to
hours to days or even weeks, depending on the type and scope of the
analysis.Some mentioned that there could be a trade-off between speed
and accuracy or the rigour of the analysis.

> I think the big thing with BLAT [the *B*LAST-*L*ike *A*lignment
> *T*ool)] is the speed. It is the speed and the trade-off that you
> get, even though you are not getting the same percent identity, the
> cut-off is still acceptable and it is a lot faster. (P1)

Quality

When discussing the quality of a tool, participants also referred to
/reliability/ or /accuracy/, sometimes treating the various terms as
synonymous.There were a variety of approaches to determining quality.It
could be judged by manually evaluating the results of an analysis to
determine if they /looked right/, based on the researcher's knowledge of
the domain and/or past experience, /'Well, based from my experience, I
could see that the alignment that it has been doing is better'/(P 8), or
by benchmarking one analysis against another.

Popularity of a tool or site and its reputation could be taken as
indicators of quality, /'I mean, it wouldn't be popular if it wasn't
accurate ...'/ (P1); /'I'm pretty sure that the accuracy of NCBI
[National Center for Biotechnology Information] in general, but partly
aside from the very infrequent errors'/ (P6).

Some participants mentioned balancing quality against either speed or
cost, in that they would accept a slower speed if it were balanced with
higher quality.

> ...because there are some other tools which can do it faster. But
> suppose I have a lot of sequences and if I have to do it faster then
> I'll have to use them but they won't be as accurate as ClustalW [a
> general purpose multiple sequence alignment program for DNA or
> proteins]. So for me, accuracy is more important than doing it
> faster so I would rather use ClustalW itself and run it for longer
> time than compromising on the quality.(P8)

They would also consider incurring costs for a tool if there was a
quality improvement over a free option. '/accuracy is so important that
if it's the tool that cost money you would pay for it'./ (P8).

Personal Factors

While many of the above factors could be objectively defined (e.g.,
interface format, cost, etc.) personal factors tended to be quite
subjective, with considerable variability in how participants defined

the factors.

Usability

Usability was a common theme among participants, but when they defined it, participants had different concepts of what features made a tool usable.One aspect of usability was how easy it was to learn. This could be interpreted as having had been taught to use the tool, or having been able to figure it out for oneself.Another facet of usability related to the function of the tool and the extent to which it did what the participant required.

> It's not a very difficult tool to use by itself. You can teach it to anyone... But, what I realised if you ask a biological question, something very specific, the answer is not very easy. You can spend days and days to look for one [piece of] information.(P5)

Usability was also related to features of the interface such as data entry mechanisms, data visualization and navigation.

Easy to use

The term /easy/ came up in various contexts such as /easy to use/ or /easy to learn/.Ease of use could relate to the amount of time it took to complete an analysis, with one participant having a one-hour timeframe in which a tool should give results.Another definition related to whether a laboratory had implemented a standard operating procedure, which would direct a novice through the analysis, rather than them having to work with the manual and documentation and figure it out themselves.The number and scope of the steps required to complete the analysis related to ease of use, with a preference for both fewer and smaller steps.

Easy to learn and documentation

Participants referred to the documentation in or about a tool, including help features, tutorials and support (online or by phone, email, etc.) services as a desirable feature of a tool.This was cited by some in the context of either learning about a tool, or learning to use a new tool, /'It had a really nice tutorial that explained [to] me how to do it'./ (P1).

> ...the support group is actually better for UCSC [University of California, Santa Cruz], because the people there are actually paid to be there and answer your questions... and all these five people do is answer questions and fix the browser.(P1)

Choice of tool

In addition to describing the factors they valued in a bioinformatics

tool, participants also described how and why they chose to use the tools they did.This could involve balancing factors such as quality against others such as speed or cost, with the choice influenced by the specific task - some tasks required a more stringent analysis, favouring quality, while others were tolerant of less stringency with faster or less expensive analysis.

> I guess the main reason I would say is just the speed. Obviously there are trade-offs though... it is fast but for its own reasons, like it requires minimum factors 99% identity between two sequences... so for the application that I'm after things like BLAT are more ideal because they're faster. (P6)

Some considered the history and reputation of a tool and chose it because it was a /classic/,

> I mean, like it's a pretty old-fashioned too, I've been using it since I started and it seems to be the class one that everybody uses.(P3)

Familiarity with a tool (or similar ones) was a factor, as was the tendency to continue using a tool once it had be chosen and learned,

Yeah, it's [GenBank, the National Institutes of Health genetic sequence database] good, because it's the same thing as NCBI's PubMed.It looks very familiar so it's easy to use. (P3).

The choice could be influenced by the equipment being used or the experimental protocol being followed.

> I use the ABI software [for gene sequencing analysis] to design primers. That's like when I'm going to use an ABI machine, specifically and I really want very specific conditions that are optimised for them. I guess they have a different algorithm that will really only look for primers that will fulfill those extremely, extremely, specific criteria... (P3)

In some cases, the choice was pre-determined by the fact that a decision had been made to purchase a specific type of equipment or software for the laboratory, /'Because that's come with the machine we buy from them . . . so the choice was already made, when this lab went with ABI.'/ (P5).

Discussion

The more than seventy different factors identified by the participants clustered into several categories. Some were similar to those previously identified and were not unexpected. These include issues such as ease-of-use, familiarity, speed, cost and quality. Others, such as the interface, scalability, customizability and functionality appear to be novel or have a novel definition in the context of bioinformatics. Even for those factors that were similar between bibliographic and bioinformatics information resources, there was variation in how the

participants operationalised the factors. This highlights the need to understand the information behaviour of the full range of people who use bioinformatics tools and to clearly understand not only which factors they value in a resource, but also how they define those factors.

We noted patterns forming in the data, with participants with similar academic background and work task reporting similar preferences.The interface was a key example.Laboratory biologists tended to prefer graphical user interfaces, with features such as pull-down menus to show the various options available to them and simple data entry mechanisms requiring minimal or no customization or manipulation of data.For bioinformaticians, the general preference was for command line interfaces that allowed them to freely manipulate the data and have full control over the customization and refinement of the analysis they were conducting.For computer scientists developing tools, the emphasis was more on the algorithm and the continuous refinement of the analysis, with the interface often an afterthought.This finding supports Hogue's assertion that algorithm may have been over-emphasised by tool developers, at the expense of user-centred factors such as the interface.

Customizability was another such factor.While most participants reported that customizability was a factor they considered in choosing a tool, biologists tended to prefer a system that presented the available options and allowed them to choose.Bioinformaticians expressed a preference for open-source software that they could manipulate and customise at will.These factors in turn related to issues such as platform; locally installed resources were considered more likely to be open-source and programmable, while Web-based resources were considered more likely to have menu-driven, graphic, user-interfaces.

These results are consistent with findings in the bibliographic domain that educational background, domain of expertise and work task may correlate with information behaviour in general and selection of resources in particular.In our findings, it was striking that the various options reported by the participants were often very dissimilar, even to the point of being polar opposites.The distinction between a graphical user interface and a command line interface is but one example; Web-based versus locally installed is another.These findings pose interesting challenges for areas such as user-centred design.It raises the question of who is the user.For whom should a tool be designed?Perhaps designers of bioinformatics tools must become aware of the range of users their resources may serve and work to built variable options into the system to accommodate that range.In the same way that many bibliographic search systems may employ both a command line and a graphic user interface, bioinformatics tools may also require multiple access options to support the range of users; most tools currently have only a single access option.

   Conclusions

We have identified a series of factors that people use to distinguish among bioinformatics tools and to select a preferred resource.Some of these were similar to those factors identified in the bibliographic

context; others were specific to the bioinformatics domain.Particularly interesting was the variability in how participants operationalised the factors they considered significant.Since often what was preferable to one group of users was not so for another, tools should be designed to support a range of preferences, not with a /one-size-fits-all/ approach. This has clear implications for the design, development, evaluation and recommendation of bioinformatics tools, particularly under a user-centred paradigm.

Future Work

The work reported here is the first phase of a larger study.To date, we have identified a wide range of factors involved in the selection of bioinformatics tools.In the next phase, we plan to use these findings as the foundation for a survey of a larger population of scientists.The larger sample size will permit quantitative analysis in addition to further qualitative findings.We will ask participants to not only identify but also to rank their selection factors, in order to be able to determine those that are most significant.We will also continue to investigate the relationship between background and task and the preferred characteristics of bioinformatics tools.Again, the larger sample size will allow us to consider the statistical correlation of these parameters.

In the final phase of the work, we plan to evaluate and annotate a test sample of bioinformatics tools according to the factors we have identified.We will also build and test a prototype system to filter and/or rank tools from the test sample according to a user's preferences.

Acknowledgments

About the authors

Joan Bartlett is an assistant professor in the McGill University School of Information Studies, where she teaches in the areas of health information, bioinformatics and information literacy.Her current research interests revolve around information behaviour and information interaction, particularly in the context of biomedical and bioinformatics information. She can be contacted at joan.bartlett@mcgill.ca <mailto:joan.bartlett@mcgill.ca>

Yusuke Ishimura is a doctoral candidate in the McGill University School of Information Studies. He has an M.L.I.S. from Dalhousie University. His Master's thesis and doctoral dissertation investigate international

students' information behaviour and information literacy skills in Canadian universities. She can be reached at yusuke.ishimura@mail.mcgill.ca <mailto:yusuke.ishimura@mail.mcgill.ca>

Lorie Kloda is a doctoral candidate in the McGill University School of Information Studies and a librarian at the McGill Life Sciences Library. Her interests include the information needs of health professionals, expert searching for systematic reviews and evidence-based practice. Lorie is also Associate Editor of the journal, Evidence Based Library and Information Practice. She can be contacted at lorie.kloda@mcgill.ca <mailto:lorie.kloda@mcgill.ca>

References

  * Bartlett, J.C. & Toms, E.G. (2005).Developing a protocol for bioinformatics analysis: an integrated information behavior and task analysis approach. /Journal of the American Society for Information Science and Technology/, *56*(5), 469-482.
  * Baxevanis, A.D. (2000). The molecular biology database collection: An online compilation of relevant database resources. /Nucleic Acids Research/, *28*(1), 1-7
  * Bottomley, S. (1999). Bioinformatics: guide for evaluating bioinformatics software. /Drug Discovery Today/, *4*(5), 240-243.
  * Brown, C. (2005). Where do molecular biology graduate students find information? /Science and Technology Libraries/, *25*(3), 89-104.
  * Cochrane, G. R. & Galperin, M. Y. (2010). The 2010 Nucleic Acids Research Database issue and online database collection: a community of data resources. /Nucleic Acids Research/, *38*(Database issue), D1-D4.
  * Creswell, J.W. (2008). /Educational research: planning, conducting and evaluating quantitative and qualitative research/ (3rd ed.). Upper Saddle River, NJ: Pearson Education.
  * Curtis, K.L., Weller, A.C. & Hurd, J.M. (1993). Information-seeking behaviour: a survey of health sciences faculty use of indexes and databases. /Bulletin of the Medical Library Association/, *81*(4), 383-392.
  * Curtis, K.L., Weller, A.C. & Hurd, J.M. (1997). Information-seeking behavior of health sciences faculty: the impact of new information technologies. /Bulletin of the Medical Library Association/, *85*(4), 402-410.
  * Grefsheim, S., Franklin, J. & Cunningham, D. (1991). Biotechnology awareness study, part 1. Where scientists get their information. /Bulletin of the Medical Library Association/, *79*(1), 36-44.
  * Haines, L.L., Light, J., O'Malley, D. & Delwiche, F.A. (2010). Information-seeking behavior of basic science researchers: implications for library services./Journal of the Medical Library Association/, *98*(1), 73-81.
  * Haynes, R.B. (2006). Of studies, syntheses, synopses, summaries and systems. The "5S" evolution of information services for evidence-based healthcare decisions. /Evidence Based Medicine/, *11*, 162-164.
  * Hogue, C. (2001). Implementation or algorithm. /Drug Discovery World/, *2*(2), 61-63.
  * Javahery, H., Seffah, A. & Radhakrishnan, T. (2004). Beyond power:

    making bioinformatics tools user-centered. /Communications of the
    ACM/, *47*(11), 58-63.
  * Palmer, J. (1991a). Scientists and information: I. using cluster
    analysis to identify information style. /Journal of Documentation/,
    47(2), 105-129.
  * Palmer, J. (1991b). Scientists and information: II. Personal factors
    in information behaviour. /Journal of Documentation/, *47*(3), 254-
275.
  * Patton, M. Q. (2002). /Qualitative research and evaluation methods/
    (3rd ed.). Thousand Oaks, CA: Sage.
  * Persidis, A. (1999). Bioinformatics. /Nature Biotechnology/,
    17(8), 828-830.
  * Rolinson, J., Al Shanbari, H. & Meadows, A.J. (1996). Information
    usage by biological researchers. /Journal of Information Science/,
    22(1), 47-53.
  * Rolinson, J., Meadows, A. J. & Smith, H. (1995). Use of information
    technology by biological researchers. /Journal of Information
    Science/, *21*(2), 133-139.
  * Stevens, R., Goble, C., Baker, P. & Brass, A. (2001). A
    classification of tasks in bioinformatics. /Bioinformatics/,
    17(2), 180-188.
  * Tran, D., Dubay, C., Gorman, P. & Hersh, W. (2004). Applying task
    analysis to describe and facilitate bioinformatics tasks. /Studies
    in Health Technology and Informatics/, *107*(2), 818-822.
  * Wilson, T.D. (1999). Models in information behaviour research.
    /Journal of Documentation/, *55*(3), 249-270.
  * Yarfitz, S. & Ketchell, D. S. (2000). A library-based bioinformatics
    services program. /Bulletin of the Medical Library Association/,
    88(1), 36-48.

How to cite this paper

Bartlett, J.C., Ishimura, Y. & Kloda, L.A. (2010). "Why choose this one?
Factors in scientists' selection of bioinformatics tools" /Information
Research/, *16*(1) paper 463 [Available at
http://InformationR.net/ir/16-1/paper463.html]

Find other papers on this subject

Check for citations, using Google Scholar
<http://scholar.google.co.uk/scholar?hl=en&q=http://informationr.net/ir/1
6-1/paper463.html&btnG=Search&as_sdt=2000>

logo Bookmark This Page <http://del.icio.us/post>

Hit Counter by Digits <http://www.digits.com>