Eagle Genomics Ltd. White Paper

Ten Steps to Successfully Outsourcing Industrial Bioinformatics

Richard Holland, Eagle Genomics Ltd., Cambridge, UK

www.eaglegenomics.com

Executive Summary

Outsourcing has long been the holy grail of companies trying to make cost savings and increase efficiencies, and never more so than in the current economic turbulence that is sweeping the globe. With reducing R&D budgets, revenue streams under threat from near-expired drug patents, and general loss of consumer confidence leading to reduced sales, every organisation in the biotech world from corporate to academic is faced with making difficult decisions about the future structure and purpose of R&D teams. Outsourcing is vital to the ongoing ability of bioinformatics teams to effectively support R&D activities within their organisations.

The Challenge

Bioinformatics has traditionally been the lowest priority in biotech R&D, with the bulk of funding and development money focused on wet-lab discovery and validation processes rather than constructing and supporting the IT systems required to analyse and fully understand the results that scientists produce.

Every bioinformatics manager will say that their group is under-resourced both in terms of staff and hardware, and that they are constantly asked to produce more from less. When times get hard for the organisation it is often bioinformatics teams that are the first to feel the pinch and come under pressure to downsize and reduce costs.

Difficult Decisions

> *"It's common for administrators to think that 'some graduate student must be free' to help out with bioinformatics projects."*

When cuts are made or requests for extra resources are denied, bioinformatics managers have to think long and hard about which services they will need to discontinue without negatively affecting their scientist clients.

Often the decision is impossible given that every scientist will claim that their project is vital and more important than all others and that their research couldn't possibly continue without the support that they are currently demanding or receiving. Tough decisions have to be made and inevitably projects will have to be cutback or cancelled, leaving scientists to suffer the consequences of the reduction of bioinformatics resources made available to support them.

With their limited resources coming under huge pressure to improve productivity and efficiency, bioinformatics managers need to start looking at the ways their mainstream IT colleagues deal with the same situations.

In the IT world it has long been best practice to hire in contractors for short-term projects with high resource requirements, or even completely outsource the entire development process to an external vendor, rather than hire permanent staff that morph into an expensive and difficult to down-size burden once the project ends.

Content Management

A useful metaphor for project outsourcing is corporate website and content management. It has been many years since any big organisation seriously considered directly recruiting individual graphic designers, web designers and content writers on its own staff to design and construct the company website.

The economic option dictates that websites are designed, built, maintained and even content-managed by specialist web design agencies that deliver high quality results for a fraction of the cost of hiring the equivalent permanent staff. The agency's expertise, proprietary tools and experience gained from repeating the maintenance and design process for so many clients is deeper, more extensive and more valuable than that which could be obtained by hiring employees to do the work in-house.

Up until now outsourcing has been rarely seen in bioinformatics departments. The reasons are various. An oft-quoted response to the question of outsourcing is concern over data security when involving third-party outsiders in the R&D process, but with careful consideration and planning this issue is never as insurmountable as it may first seem. Far more widespread however is that many bioinformatics managers have simply never seriously considered the benefits that outsourcing could bring to their projects.

Proof the Problem Exists

The financial pressures on large pharmaceutical companies to restructure and downsize are nothing new, and they look set to continue. Major announcements in recent times include GSK's 4,000 redundancies in early 2010[1] and Eli Lilly's layoff of 5,000 staff across the business a few months beforehand[2]. Eli Lilly has since entirely closed down its Singapore R&D centre with the loss of 130 jobs, a good number of whom were bioinformaticians[3].

Despite these large-scale downsizing activities, GenomeWeb's BioInform reported[4] that 'a number of scientists told [BioInform] that bioinformatics positions in their labs are hard to fill and in some cases must even be cancelled for lack of qualified applicants.' Interestingly they also noted that whilst it was hard to fill the positions that were advertised, the actual number of advertised positions was much lower than expected as 'budgets are being squeezed and administrators often do not readily see the value of hiring dedicated bioinformatics staff or do not have the funds to hire as many qualified bioinformatics scientists as they may actually need to support their research.'

The common misconception that bioinformatics is easy also comes out in the GenomeWeb article, quoting Mark Borowsky from the Massachussets General Hospital's department of molecular biology as saying 'It's common for administrators to think that 'some graduate student must be free' to help out with bioinformatics projects.'

"80 percent of one division's pre-clinical informatics staff has been cut in the last three years."

Downward Trends

The same article on GenomeWeb confirms the trend towards cutbacks in bioinformatics, particularly in drug discovery, by quoting an anonymous industry source as saying 'reorganizations and job cuts have hit drug discovery informatics hard. In one instance, 80 percent of one division's pre-clinical informatics staff has been cut in the last three years.'

A Bioinformatics LLC report in 2009[5] suggested that 65% of respondents in its survey of biotech scientists had indicated that their research had already been affected by the downturn in the economy.

Changing Attitudes

A Business Insights Ltd. report, also from 2009[6], shows that outsourcing of key bioinformatics tasks is catching on in the biotech world, and is growing at a remarkably rapid rate. Across the board of the bioinformatics services market grew by nearly 25% year-on-year during 2007–2009, a growth rate which is expected to continue until at least 2014. Total market value for bioinformatics services is expected to exceed US$1bn by this time, and the market for bought-in bioinformatics content and databases will exceed US$3.5bn.

A key driver of the trend towards outsourcing has been the provision of integrated services making easy the acquisition and integration of external data into internal systems. As the Business Insights report says, 'biotechnology companies are looking for suppliers that can offer total integration of data infrastructure, which includes data sharing, data security, customization, data searching and analysis. The total package is supported by after-sales services and consultancy, which suppliers provide as a core competency.'

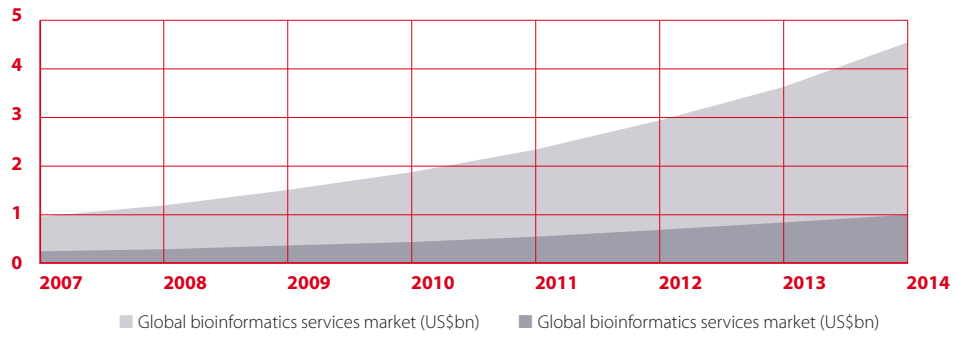


Figure 1 ***Predicted Bioinformatics Market Values***

1. **The Guardian**, January 31st 2010. *http://www.guardian.co.uk/business/2010/jan/31/glaxosmithkline-expected-job-cuts*
2. **Industry Week**, September 14th 2009. *http://www.industryweek.com/articles/eli_lilly_to_cut_more_than_5000_jobs_19970.aspx*
3. **Bio-IT World**, October 15th 2010. *http://www.bio-itworld.com/2010/10/15/lilly-singapore.html*
4. **GenomeWeb**, November 13th 2009. *http://www.genomeweb.com/informatics/bioinformatics-job-market-tug-war-heavy-demand-data-analysis-vs-tightening-budge*
5. **Bioinformatics LLC**, *'Prospering in a Down Market: Strategies for Life Science Suppliers'*, March 2009.
6. **Business Insights Ltd.**, *'The Global Bioinformatics Market'*, 2009.

Preserving the Corporate Memory

When bioinformatics budgets in large biotech organisations come under pressure, the degree of cutbacks necessary usually leads to one of two scenarios.

The first scenario is that new projects in the pipeline get cancelled or seriously reduced in scope, but staffing and hardware resource levels remain constant in an attempt to increase efficiency. In reality, the perceived efficiency improvements simply lead to existing staff working longer and harder hours. Although this does lead to an increase in man-hours available it can often be counter-productive because it creates unnecessary stress and fatigue amongst staff and can lead to loss of productivity, increased absenteeism, and even a rise in staff attrition.

Loss of Insight

The second scenario is that cuts are so bad that staff have to be laid off. With them goes the invaluable knowledge and experience of their field and of corporate infrastructure and ways of working that they have built up over the years. This experience is lost to the bioinformatics team just when they most need the expertise these laid-off staff could have brought to the table in developing innovative plans to improve efficiency and delivery prospects.

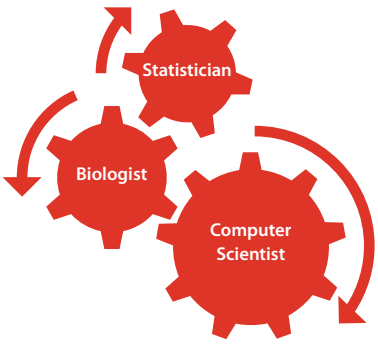
Not only that, but if another organisation in a similar field has taken an alternative approach to the restructuring issue and decided instead to increase their bioinformatics staffing levels, organisations run the risk of losing key experienced staff to a competitor. No matter what happens to the laid-off staff though, the remaining staff are usually demoralised by their colleagues departure and come under the same pressures for efficiency as under the first scenario above and therefore leading to the same problems.

Aside from budget cutbacks as a result of the current economic climate, the general historical lack of resources in bioinformatics groups can cause problems in other ways.

Asking the Postdocs

In academic groups and smaller biotech businesses, there is usually a complete lack of dedicated bioinformatics resources that can be realistically shared between all staff scientists. This leads to scientists writing their own code, or getting their postdocs to do it. Despite their best efforts, scientists trained in biological sciences are rarely the best programmers and the code produced is suboptimal. For those who delegate such tasks to their postdocs and PhDs, when these students move on in due course they take with them the only real in-depth knowledge about the software that exists in the organisation, essentially rendering it unsupported and open to rapid deterioration of service.

Types of Bioinformatician



Within bioinformatics groups, staff members usually fall into one of three categories – biologists, statisticians, and computer scientists. It is rare to find one person with in-depth expertise in two of these areas, let alone all three, rather they are usually highly trained and experienced experts in just one of these three areas with a passing knowledge in the others, enough to know the basics but not enough to be able to exploit the full range of possibilities that an expert training in that field would provide.

In a larger group the staff members with varying domain expertise can work together as a team to approach the problems of a project from all sides with equal aplomb, however in smaller groups and in particular in labs which only employ a sole bioinformatician, projects run the risk of excelling in one area (e.g. understanding the biological problem being solved) without necessarily delivering the best overall solution (e.g. incorrect selection of statistical techniques, or poorly crafted software code and user interfaces).

The Solution: Outsource!

When faced with the pressures of downsizing a bioinformatics group, or attempting to increase the productivity of an existing one without receiving any additional resources or support, then outsourcing of key tasks is a natural decision to make. Offloading standalone projects or long-term support and maintenance tasks to a trusted third-party vendor can allow the team to refocus on key projects that will make the biggest difference to the effectiveness of their organisation's R&D. Whilst it costs money to outsource, the value gained by doing so often outstrips the cost of having to carry out the equivalent work in-house.

However, as with all things informatics, care and attention to detail is vital when making the decision to outsource a bioinformatics project. Here are ten key points to consider when looking to outsource:

1 *Look for an established vendor with a track record in bioinformatics.*

A vendor that is well established and has a team of staff available to work on projects will have a wider breadth of knowledge and deeper insights into domain areas than a sole-trading consultant or contractor. The vendor's team must represent a broad spectrum of bioinformatics experience across a wide range of organisation types. Ideally, ask for references from a previous client.

Avoid generic IT outsourcing providers that do not have direct experience in bioinformatics. Any savings made in doing so may well be wiped out many times over by the expensive process of identifying and fixing bugs caused by their lack of domain knowledge.

2 *Don't disregard open-source solutions.*

Proprietary data and software licenses can be extremely expensive. Happily, in the bioinformatics field a vast amount of the data and software output comes from academic groups where the open source culture is very strong.

Although the quality and maintenance of the data or software can substantially vary, an experienced vendor with the right domain knowledge will know exactly which open-data and open-source options can be leveraged to produce scalable and reliable solutions at a fraction of the price of the proprietary equivalent, if indeed a proprietary equivalent exists at all.

3 *Consider data security.*

Will the vendor work on-site or off-site? If off-site, will they need VPN access or other remote login arrangements? If on-site, how will they be prevented from inadvertently exporting sensitive data on USB sticks or connecting their laptops to the internet without going through the corporate firewall? If they get to view unencrypted versions of proprietary and sensitive data, how can it be ensured that they will not use it for any other purpose than what is required for the project? How will these arrangements be enforced, and what provisions will be made for claiming losses if data should be leaked or misused?

Whilst the answers to these questions are different for every organisation, a good vendor should be able to work with clients to find operational solutions that will satisfy legal teams whilst still keeping the project viable.

4 *Know what the project is.*

Countless projects overrun because they were not fully specified, or requirements changed partway through. This is as much a problem for in-house development as it is for outsourced work, but when outsourcing it can lead to expensive costs inflations which negate the benefits of outsourcing in the first place.

Projects don't need to be specified in every detail at the outset, but bioinformatics managers should be aware of the key components and requirements from the start and ensure that these are communicated clearly to the vendor and that they are fully understood before work can begin. A good vendor with decent domain knowledge and extensive experience will know how to ask the client the right questions and will do so unprompted, but if they don't, ensure that information is provided early before any trouble starts. On a related note, make sure that the contract with the vendor explicitly states who is liable for any overruns.

5 *Share the goal with your vendor.*

A run-of-the-mill vendor will always be looking to run the project in the most efficient way they can, which usually means sticking to default configuration options for any analysis tools used, and returning raw datafiles as results in whatever format the tool that produced them usually churns out. Such default-configured default-formatted data is often no more useful to the client than nothing at all, and it can be expensive to try and untangle the mess and, if it's really bad, rerun the analysis.

A quality vendor will talk to the client in detail and in advance about every aspect of the analysis to be performed, including every configuration option, and try to gain an in-depth understanding of the research question the client is asking of their data so that they can produce the correct results first time. They will also take care to find out what specific format the results are needed in so that clients can feed them straight into reporting tools or in-house analysis pipeline without further modification. This extra time invested up-front will pay massive dividends in the long run.

6 *Set realistic goals.*

Bioinformatics projects are complex and lengthy. With the exception of pre-built pre-configured data processing services which execute identical series of analyses on demand across multiple input datasets of identical nature, then even with the best will in the world it is impossible to take the time to accurately specify, design, build and deliver a bioinformatics project in a period of anything less than weeks, or more usually months.

Vendors who can promise to deliver in impossibly short amounts of time are quite likely to be grossly underestimating the complexity of the task, indicating that they lack the domain knowledge and expertise to be able to fully comprehend the problem at hand. Likewise, if a vendor says it will take months longer than may seem necessary, it could be indicative of the same problem and they are playing on the safe side to avoid embarrassing overruns.

To get a realistic estimate of time, consider how long it would take in-house staff to perform the project – including training time on the assumption that they have never seen the relevant data or systems before – then add a small margin for negotiation/consultation time and use that as the benchmark for measuring vendor estimates. Remember that the more information that can be given to the vendor up front, the more accurate their quote is likely to be.

7 *Build a communication plan with your vendor.*

Weekly vendor progress reports or conference calls, or both, will help an outsourced project stick to time and budget. They can be used to prevent feature creep, to assess and prioritise development targets, and to check on and ensure relevance and accuracy of work done so far. They also provide an ideal forum for the vendor's developers to ask questions that they need detailed answers to in order to progress their work; the kinds of questions that can take days to explain and discuss by email but only take five minutes to sort out by phone.

In addition, establish a chain of command within the organisation that can respond to requests for information throughout the duration of the project. Key staff members at the client organisation being off sick or on annual leave are a frequent cause of delays during outsourced projects where no backup contacts have been assigned. A responsible vendor will include plans for maintaining regular contact in their project proposal.

8 *Start small, think big.*

When deciding to outsource, don't do it all at once, and especially not with a newly sourced vendor. Instead, divide up the projects in mind into smaller chunks and outsource them one at a time until confidence in the process is built up and future projects can be aggregated together into single contracts. This allows inferior vendors to be eliminated at an early stage without risking large-scale failures, whilst the quality vendors can slowly build up their portfolio one project at a time until they become a trusted and integral part of the outsourcing procedures.

As an example, when considering the outsourcing regular sequencing analysis projects to a new vendor, draw up a project involving just one sample to test the water, then bring the vendor back on a larger-scale longer-term contract only if their work ethic, communications and output quality meets or even better exceeds expectations.

9 *Check with the lawyers.*

Legal issues and intellectual property (IP) considerations surrounding software and data can be highly complex, and there are a number of issues that must be considered when hiring a vendor to work on outsourced projects.

Who will own any IP that is developed by the vendor? If the organisation retains the IP, will the vendor be licensed to reuse it on other projects with other customers? If the vendor retains the IP, what licence will the organisation receive it under, and what would happen if the vendor ceased trading? Would it be easier to open-source it and if so under what licence? What compensation mechanisms have been set up to deal with overruns and project failures? Does the vendor have enough insurance, of the right type, and in the right jurisdictions? How will the organisation recover damages if the vendor misuses data or disrupts key systems?

There could be many more scenarios and questions which may need to be addressed, but the selection of questions given here are a good starting point and should always be asked of every vendor at the earliest possible stage of negotiations.

10 *Maintenance and support.*

After a project is delivered, who will be responsible for carrying it forward in-house after delivery? If it is a software product, who will train the users and man the helpdesk, and who will provide updates and patches? If it is a data service, who will keep the data updated, mapped to the most current reference dataset, or integrated with ever evolving proprietary data? Answers to these questions are vital before a project even begins, because the answers can determine the way in which the project is implemented. For example, software intended to be maintained by in-house developers must be documented in great detail by the vendor to enable the developers to understand the inner workings, whereas software to be maintained by the vendor may not require developer documentation to be delivered.

Data that will be maintained by the vendor must have an agreement negotiated to ensure the vendor provides those updates in a timely manner over an agreed support period, with appropriate renewal clauses to ensure continuing service availability. Software or data that will remain closed source or licensed to the client by the vendor must always have an escrow clause that will allow the client to gain access to the original source and retain the right to continue using the product in the event that the vendor ceases trading. The best vendors will account for all these situations in their proposals and discuss them during the quote process.

*The **ElasticEagle**™ Solution*

Eagle Genomics are experts in genome content management. The company's flagship ***ElasticEagle***™ provides a timely, cost-effective and appropriate bioinformatics outsourcing solution that suits all organisations using genomic data from big pharma to the smallest startup biotech companies, and across all sectors from drug discovery and crop improvement through to animal health and food safety.

Track Record

Eagle Genomics' track record of delivering outsourced projects of all sizes on time and on budget to a number of the world's largest pharmaceutical and agri-biotech organisations stands testament to our ability to understand and clearly define client projects from the outset and get it right first time. Close collaboration with key open-source academic projects enables us to gain access to the latest cutting edge bioinformatics research software and data, ensuring that our clients also benefit from knowing that everything Eagle does is going to be state of the art.

ElasticEagle™ is everything organisations need to successfully outsource bioinformatics projects.

www.eaglegenomics.com

For more information about ***Elastic Eagle***™ services and about Eagle Genomics Ltd. in general, please visit our website or contact Richard Holland on:

Email: holland@eaglegenomics.com
Tel: +44 (0)1223 654481 ext. 3