

# **Bioinformatics Issues and Opportunities for Australia**

**Prepared for the Services and Emerging Industries Unit,  
Department of Industry, Tourism and Resources by  
BioLateral Pty Ltd and the Bioinformatics Industry  
Opportunity Task Force**

**Emerging Industries Occasional Paper No 15**

**January 2002**

This paper is issued for discussion and information purposes only. The views expressed do not necessarily reflect the views of the Australian Government or the Department of Industry, Tourism and Resources.

Copyright in ABS Data resides with the Commonwealth of Australia.

**THIS REPORT IS PART OF A SERIES OF EMERGING INDUSTRIES OCCASIONAL PAPERS. THE PAPERS CURRENTLY AVAILABLE IN THIS SERIES ARE:**

<b>Emerging Industries Occasional Paper 1</b>	Emerging Industries & Technologies Forum - Report of Proceedings, 17 August 1999
<b>Emerging Industries Occasional Paper 2</b>	Enabling Technologies for Australian Industry -A Pilot Study, November 1999
<b>Emerging Industries Occasional Paper 3</b>	A Study of Government R&D Expenditure by Sector and Technology, December 1999
<b>Emerging Industries Occasional Paper 4</b>	World Wide Work: Globally Distributed Expert Business Services, October 2000
<b>Emerging Industries Occasional Paper 5</b>	Skills Needs of Emerging Industries, October 2000
<b>Emerging Industries Occasional Paper 6</b>	Environment Management Industry Showcasing, "Being Heard Above the Crowd", December 2000
<b>Emerging Industries Occasional Paper 7</b>	Titanium Metal - a market analysis, February 2001
<b>Emerging Industries Occasional Paper 8</b>	A study of the USA and Canadian Flower Markets, March 2001
<b>Emerging Industries Occasional Paper 9</b>	Investing in Emerging Opportunities, Emerging Industries Forum 2000: Report of Proceedings, October 2000
<b>Emerging Industries Occasional Paper 10</b>	National Youth Entrepreneurship Attitude Survey May 2001
<b>Emerging Industries Occasional Paper 11</b>	Environment Industry Cluster Development- the South Australian Experience, May 2001
<b>Emerging Industries Occasional Paper 12</b>	The Impact of High Performance Computing Technologies on Australian Industry, August 2001
<b>Emerging Industries Occasional Paper 13</b>	Technology Planning for Business Competitiveness: a Guide to Developing Technology Roadmaps, August 2001
<b>Emerging Industries Occasional Paper 14</b>	Photonics: opportunities for Australia. 1998 international photonics benchmarking mission report update, January 2002.

Copies of these papers are available from the Services and Emerging Industries Unit of the Department of Industry, Tourism and Resources, GPO Box 9839, Canberra ACT 2601 or e-mail: [EmergingIndustries@industry.gov.au](mailto:EmergingIndustries@industry.gov.au)

Copies are also available from the Emerging Industries site accessed through <http://www.industry.gov.au>.

© Commonwealth of Australia (2002)

This work is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without prior written permission from the Commonwealth available through AusInfo. Requests and inquiries concerning reproduction and rights should be addressed to the Manager, Legislative Services, AusInfo, GPO Box 1920, Canberra ACT 2601.

ISBN 0 642 72131 9

ITR 2001/007

The views expressed in this report are not necessarily those of the Department of Industry, Tourism and Resources or of the Commonwealth of Australia.

# Contents

<b>1. Executive Summary</b>	<b>1</b>
<b>2. Introduction</b>	<b>3</b>
2.1 Background	3
2.2 Taskforce Membership	3
2.3 Key Objectives	4
<b>3. The bioinformatics industry</b>	<b>5</b>
3.1 Defining bioinformatics	5
3.2 Conceptual framework for the industry	5
3.3 Current players and projected value	6
3.4 Structure and growth of the bioinformatics industry	11
<b>4. Australia's bioinformatics capabilities</b>	<b>14</b>
4.1 Importance of bioinformatics in Australia	14
4.2 SWOT analysis of Australian bioinformatics	17
4.3 Education, training and skills development	18
4.4 Research	19
4.5 Australia's position in the global bioinformatics marketplace	21
4.6 Commercialisation	21
4.7 Industry role models, linkages and partnerships	24
4.8 International approaches	25
4.9 Encouragement and support of Australian start-up bioinformatics companies	26
<b>5. Bioinformatics Industry Development: issues and opportunities</b>	<b>27</b>
5.1 Training and skills development	28
5.2 Coordination of competitive and pre-competitive research	30
5.3 Building a strong bioinformatics infrastructure	31
5.4 Industry growth and development	32
<b>References</b>	<b>35</b>
<b>Appendix A: Bioinformatics synonyms and definitions</b>	<b>36</b>
<b>Appendix B: Stakeholders who would use/need bioinformatics</b>	<b>40</b>

### List of figures

<b>Figure 1.</b>	An integrated view of bioinformatics	6
<b>Figure 2.</b>	Worldwide sales of bioinformatics software and data	6
<b>Figure 3.</b>	Bioinformatics market segments in 2005	7
<b>Figure 4.</b>	Bioindustry market reach	8
<b>Figure 5.</b>	Biosciences Industry Growth	10
<b>Figure 6.</b>	Structure of the bioinformatics industry	11
<b>Figure 7.</b>	Bioinformatics publications	19
<b>Figure 8.</b>	A model for a bioinformatics cluster	33
<b>Figure 9.</b>	Interdisciplinary nature of bioinformatics	37
<b>Figure 10.</b>	Relationship of bioinformatics with biotechnology and ICT industries	38
<b>Figure 11.</b>	Intersection of bioinformatics with medical informatics and cheminformatics	39

### List of tables

<b>Table 1.</b>	Australian bioinformatics SWOT analysis	17
<b>Table 2.</b>	Current contacts for bioinformatics training in Australia	18
<b>Table 3.</b>	Comparison of undergraduate degrees available in Australia and the USA	19
<b>Table 4.</b>	The number of groups conducting 'traditional' bioinformatics research	20

# Executive Summary

## Executive Summary

Bioinformatics is the application of information technologies and sciences to the organisation, management, mining and use of life-science information. Bioinformatics plays a central role in the development of bioindustries in general, including biotechnology, pharmaceutical, agrifood, environment management, medical diagnostics, bioprocessing, bioprospecting and others. Indeed, it is impossible to imagine a modern bioindustry sector in Australia without a competitive bioinformatics sector.

There is a broadly accepted view that support for knowledge-based industries is critical to national economic growth, and bioinformatics is perhaps one of the best examples of such an industry.

The next three years will be crucial in establishing the bioinformatics industry in Australia. If intensive research and training efforts are initiated, then Australia can expect to see rapid spin-off effects in industry.

Globally the bioinformatics industry is expected to exceed US\$ 11B by 2004. In Australia, the bioinformatics industry is very small and pre-emerging. Even so, large opportunities in growing the industry exist here, with the most promising areas for growth being in services (~US\$ 2B globally by 2004), storage (~US\$ 5.5B globally by 2005) and servers (~US\$ 3.5B globally by 2005). The great challenge to the nation is how to capture a slice of this opportunity.

Bioinformatics technologies are needed to digitise, organise and mine Australian life-science datasets. These datasets span our unique clinical, biodiversity and specialised life-science knowledge base, information that will underpin future developments in biotechnology and related industries. A domestic bioinformatics capability will unleash the potential value of these datasets, and the deep domain knowledge of Australian life-scientists who have collected and curated these data will ensure both the effective organisation and exploitation of this information. Australia needs not only to develop a keen sense of how to 'prospect' this vast legacy of data, but also how to recognise the commercial value it holds, and then to convert it from pre-competitive (often undigitised) formats into a form ready for commercial exploitation using bioinformatics tools.

Three areas of major opportunity for such effort include Australia's extensive biodiversity datasets, animal and plant breeding data and substantial clinical, genetic and epidemiological datasets. Together, these would allow the capture of markets for bioinformatics products and services and the creation of downstream products.

While some bioinformatics technologies may be best accessed from off-shore sources (eg. computing hardware and other data sets accessed through the internet), many activities require a strong domestic capability. A local industry and bioinformatics skills capability would allow Australia to recognise and realise the value of its unique data sets and provide the opportunities to capture, develop and exploit these knowledge resources. A focus on bioinformatics would allow Australia to develop excellence in this niche area and earn a seat at the international science and technology table. Ignoring the opportunity to build a globally competitive bioinformatics sector would run the risk of Australia being a technology importer in this key area, stifling the growth of all bioindustries in Australia.

A strong domestic bioinformatics capability would accelerate the development of Australian bioindustries by ensuring that Australian firms have access to capabilities to help them develop biotechnology-based products and services. A bioinformatics industry must be located on-shore, as it would allow industry players to service these bioindustries at greatly reduced cost and would help to retain intellectual property and expertise within Australia. An Australian-grown bioinformatics capability would improve the international perception of Australia as a global bioindustry centre and, in turn, increase Australia's credibility to attract research and commercially related foreign direct investment.

There is great potential for growing the bioinformatics industry in Australia. It has many 'raw ingredients', including knowledgeable life-science and information technology researchers and professionals, good network infrastructure and a growing venture capital industry. However, while Australia has a nascent bioinformatics industry, without a national strategic approach it is unlikely to attract private investment at a scale and quality to make a significant difference to the science and industry base in Australia.

A vibrant bioinformatics industry would play an important role in the economic and social development of Australia. Bioinformatics, as an enabling technology for bioindustry

developments, would help to generate significant economic benefits such as high skill and high value jobs, growth of firms (eg. in contract research, equipment manufacturing and software development), additional private sector research and development spending, and tax revenues (eg. via property, sales and income).

Globally, bioinformatics is a relatively new sector, yet it is experiencing rapid growth and has healthy potential. With the right investment and skills base, the bioinformatics industry's economic and fiscal contributions are expected to increase at the same pace as those expected in the biotechnology and broader bioindustries.

Numerous studies point to the major returns Australia would gain from investments in innovation and knowledge industries. Australia needs to 'punch above its weight' in the global knowledge economy. By doing so, it would benefit from greater international technology transfer and engagement with global knowledge industries.

This report identifies opportunities and actions for developing Australia's bioinformatics capability. Action is required in the following critical areas:

- Education and training - increasing skills in bioinformatics;
- Research and development - producing new technologies and applications, both pre-competitive (basic, public funded) and competitive (applied, private funded) research;
- Facilities - major infrastructure required to process data; and
- Commercialisation - conversion of skills and innovation into jobs and wealth.

# Introduction

## 2 Introduction

This report is one of a series of Emerging Industries High Opportunity Studies, prepared by an industry-led Bioinformatics Industry Opportunity Taskforce (BIOT). BIOT was established through support from the Emerging Industries Section of the Commonwealth Department of Industry, Tourism and Resources (ITR). BIOT has engaged stakeholders involved in Australia's emerging bioinformatics industry, with a view to defining the industry and identifying opportunities and actions to stimulate commercial outcomes from Australia's nascent bioinformatics research and industrial capability.

### 2.1 Background

The Taskforce was formed with the knowledge that Australia has a track record and substantial potential in bioinformatics. In 1991, Australia established a national facility for Bioinformatics - the Australian National Genomic Information Service (ANGIS), which is a member of the European (EMBnet) and Asian (APBionet) consortia of bioinformatics facilities. Australia has a dispersed but growing bioinformatics research base, principally at the University of Sydney and the Australian National University, with activities also at several other locations including the West Australian Bioinformatics Consortium, the University of Queensland, the Victorian Bioinformatics Consortium and the CSIRO. It was not clear, however, how many bioinformatics companies exist in Australia or what is the growth potential of the industry.

The formation of an industry-led task force (BIOT) was considered the most effective and efficient way to investigate and explore the current and future opportunities for bioinformatics in Australia.

### 2.2 Task Force Membership

The Task Force comprised a core group of industry experts:

- **Chairman and bioinformatics industry expert**  
Dr Tim Littlejohn, Managing Director, BioLateral Pty Ltd.
- **Biotechnology/Pharmaceutical industry experts**  
Professor German Spangenberg, Plant Biotechnology Centre, Victoria; and Dr Deborah Rathjen, Managing Director and CEO, Bionomics.

- **Information and Communications Technologies (ICT) expert**

Mr Bob Mounic, RJM Services Pty Ltd; past executive of the Australian Information Industries Association (AIIA) and with Sun Microsystems Australia.

- **Investment community expert**

Dr Roger Buckeridge, Allen and Buckeridge Pty Ltd.

- **Research community representatives**

Dr John Curran, Entomology Division, CSIRO;  
Professor Simon Easteal, John Curtin School of Medical Research, Australian National University.

## 2.3 Key Objectives

The Task Force worked by bringing together or consulting different interest groups and relevant stakeholders to identify opportunities and actions for bioinformatics industry growth (considering that bioinformatics is a cross-cutting industry at the interface of biotechnology and Information Technology and Communications (ITC) industries).

The project included consideration of specific issues, including:

- defining the scope and parameters of the bioinformatics industry to reduce confusion and to focus attention on a specific industry segment;
- defining the role of bioinformatics in the growth of other bioindustries in Australia;
- identifying and defining industry development needs and those of the market;
- identifying specific goals and target outcomes needed to position Australia to tap into the large global bioinformatics industry and grow new companies;
- considering options for the establishment of an ongoing mechanism for the bioinformatics sector that stimulates networking and collaboration on matters concerning research, commercialisation and the exploitation of platform technologies.

# Bioinformatics Industry

## 3 The Bioinformatics Industry

### 3.1 Defining Bioinformatics

There is considerable confusion about the definition of bioinformatics: there is not even agreement within the community from which it arose as to what is the best terminology to define the area.

In its *broadest* definition, bioinformatics is the application of information technologies and sciences to the organisation, management, mining and use of life-science information.

The *narrow* (and typically undisputed) definition of bioinformatics is its application to the processing of molecular biology (DNA and protein) datasets. A broader definition of bioinformatics is required due to the very nature of biology and biological processes. Ultimately, biology is about the study of living systems, from biosphere to population to genome to gene to molecules. Accordingly, in this report, given the broad spectrum of life science issues to which it relates, the bioinformatics industry is defined as follows:

***"Companies that sell products or services that are either digitised<sup>1</sup> life science information or used by their customers to generate digital life-science information."***

Appendix A contains a detailed discussion of terminology, definitional issues surrounding bioinformatics and the reasons for using this definition.

### 3.2 Conceptual framework for the industry

An Australian bioinformatics industry would ideally comprise four major inter-related elements:

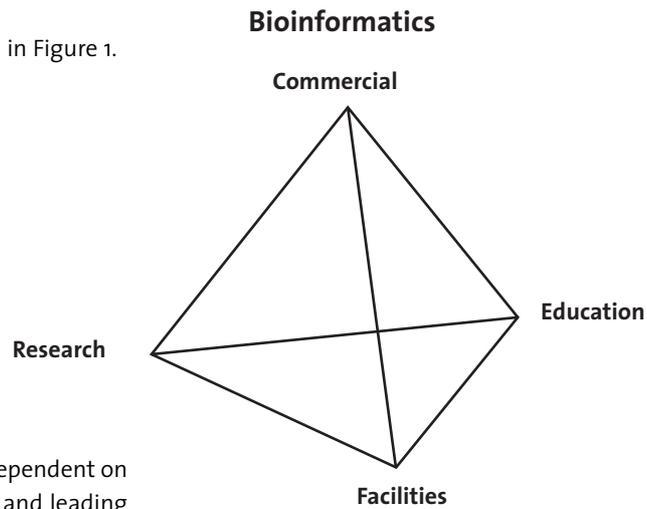
1. Education and training - increasing bioinformatics skills base.
2. Research and development - producing new technologies and applications, both pre-competitive (basic, public funded) and competitive (applied, private funded) research.
3. Facilities - major infrastructure required to process data.
4. Commercialisation - conversion of skills and innovation into jobs and wealth.

---

<sup>1</sup> "Digitised" and "digital" means computer based, network transmissible data

These four areas are interdependent, as shown in Figure 1.

**Figure 1. An integrated view of bioinformatics**

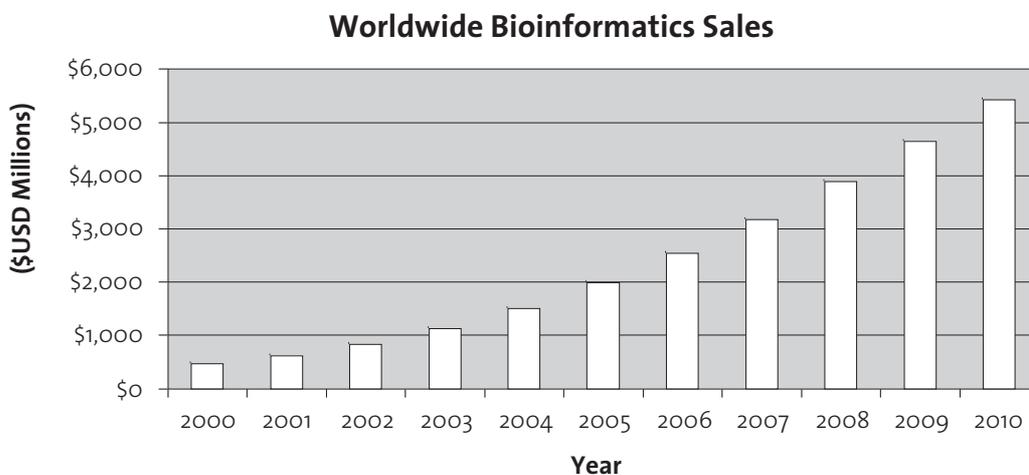


Each quadrant of the bioinformatics sector is dependent on the other. Education needs access to facilities and leading edge research, and is a service industry; Research needs access to facilities and education for re-skilling and providing graduates and results in commercially valuable IP; Facilities need novel technologies produced by research, learning content developed in the education sector, and commercial products; Commercialisation requires IP from research, education for skills acquisition and access to facilities, particularly for SMEs.

### 3.3 Current players and projected value

Bioinformatics industry opportunity has been analysed by companies such as Frontline Research (2000), using the narrow definition (bioinformatics data and software). Their data indicate that the market will be at least \$USD 2 billion by 2005 and \$USD 5 billion by 2010.

**Figure 2. Worldwide sales of bioinformatics software and data. (Frontline, 2000)**



The major bioinformatics market segments identified by Frontline are focused on perceived growth in needs principally by the biotechnology and pharmaceutical industries. These needs are partitioned into the following broad categories:

- enterprise system providers
- database (data) suppliers
- analytical software providers

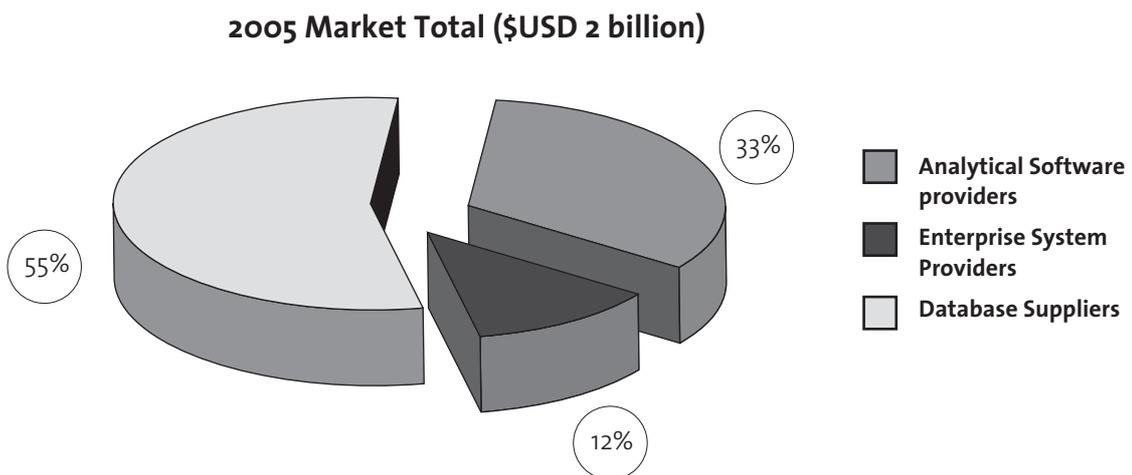
Enterprise solution providers include companies like Informax and NetGenics. Database suppliers include companies such as InCyte and Celera. These companies should more correctly be called *databank* suppliers, as they provide *content* not data management technologies. Database companies are more generic (and hence have larger markets that *include* bioinformatics, such as Oracle and Sybase). Analytical tool providers include companies such as Accelerlys (recently formed by Pharmacoepa after its recent acquisition of a number of players including GCG, Oxford Molecular and MSI). Interestingly, some companies

have products in more than one category, for example, InCyte sells its databanks and its 'Life Tools' Enterprise suite (used principally to access its databanks). Other companies (eg. Accelerlys) sell tools that span bioinformatics and chemoinformatics, showing the blurred lines between these industry spaces (reflecting the needs of the principally pharmaceutical companies they supply).

These segments are being driven by the rapid generation of genomic and proteomic data, including protein structure and function prediction, gene and protein expression data analysis, and SNP (single nucleotide polymorphism) detection and disease correlation.

In 2000, Frontline identified 24 major bioinformatics companies globally, with an additional 42 players. These numbers were expected to grow sharply as the industry grows over 10-fold in the next 10 years from its current base of \$USD 468 million in 2000 to \$USD 5,421 million in 2010. By 2005 the industry is predicted to be partitioned as follows (Figure 3).

Figure 3. Bioinformatics market segments in 2005 (Frontline, 2000)



Data are becoming available for the broader definition of bioinformatics that include:

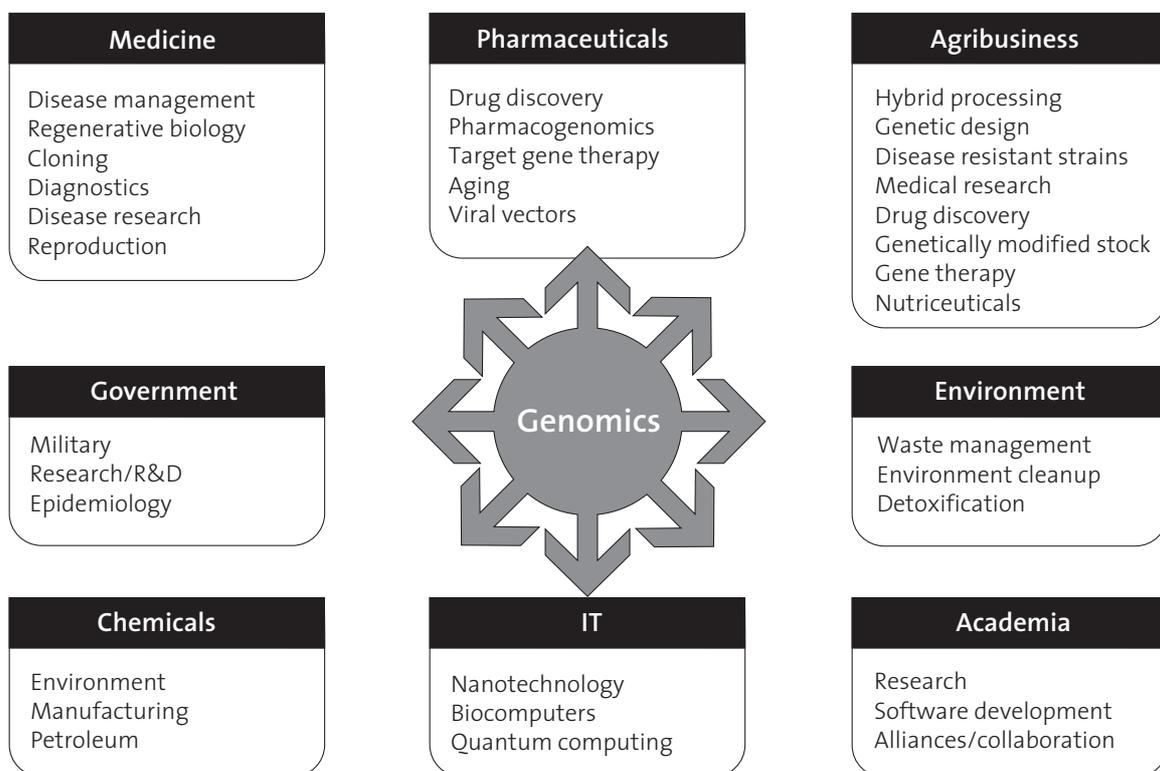
- computer hardware and equipment sales;
- pure bioinformatics services (support, training, custom application development);
- gene target identification and validation firms (eg. Entelos and Myriad);
- information services such as genetic testing; and
- non traditional bioinformatics products such as biomedical, biodiversity and neurobiology information services.

A market research report conducted by Frost and Sullivan for IBM estimated that life science companies will be spending substantially more than the amounts estimated by Frontline. Frost and Sullivan predicts the market will be worth \$USD 6.5 billion on computer-related services by 2004 and that the total spend for all bioinformatics activities might be as high as \$USD 9 billion in 2001 (Jasinski, IBM, 2001).

Data on the larger market potential for segments of this market are available from IDC (<http://www.idc.org>). For instance, in their report "The Convergence of Biology and IT: Market Impacts of the Bioscience Era" (<http://www.idc.com/El/eio42001.htm>), IDC says:

*"The scope of applications within the biosciences market ranges from bioinformatics and genome sequencing to proteomics and computational chemistry. Each one of these applications has many sub-disciplines that are markets in and of themselves. The applications influence a wide range of industries from pharmaceuticals for drug discovery, agriculture for developing disease-resistant crops, and chemicals for new product development to environmental management for waste clean-up. The common thread among all these applications is that they are using new methods, knowledge, and data derived from the sequencing of genomes and the analysis of proteins as the fundamental component of the research and development (R&D) process"* as shown in their figure (see Figure 4). Additional areas of strength in Australia that could be highlighted in this picture would include bioprospecting and bioprocessing.

**Fig 4. Bioindustry Market Reach. IDC, 2001. (<http://www.idc.com/El/eio42001.htm>)**



IDC adds: "At the heart of the bioscience revolution is data - data derived through experimentation and discovery. The volumes of data are immense and growing at nonlinear rates. To illustrate this point, we offer the example of Celera Genomics, the private sector company responsible for sequencing the human genome. The infrastructure that supports Celera's activities is impressive: 900 AlphaServer processors and 6 Paracel GeneMatcher systems (Paracel was an independent MPP (massively parallel processor) company that Celera purchased last year) when combined account for an aggregate 1.7 teraflops of installed processing power. The storage environment is equally striking, with a 70TB database growing at 15-20GB daily."

By way of comparison, a typical desktop PC runs today at ~200 mflops (<http://www.dl.ac.uk/TCSC/disco/Benchmarks/whetstone.html>) and contains 10 GigBytes. Thus in this example a single bioinformatics-driven company has the equivalent of almost 10,000 PCs. IDC continues to say that:

"Celera claims its data management and computational requirements will grow one thousandfold over the next 18 months. A perspective on the magnitude of this market opportunity says that proteomics is only one step, albeit an important one, in the evolution of this industry. Many other equally if not more demanding applications lie ahead on the horizon.

Although data is the key market driver, the key commodity - and ultimately the most valuable asset - is knowledge. Collecting the data is becoming 'relatively' easy. Being able to extract meaning and insight is much harder. To do this requires building an infrastructure to support the rapid growth and complexity of these data sets, coupled with an access and analysis environment that is simple to use.

The computing piece, however, is significant. Many of the applications in this segment demand ultrascalability - scaling for both throughput and complexity. This is a market with a considerable demand for systems with the ability to manage large amounts of data, to sort and compare multifaceted data and data types, and to do complex modeling and simulation.

In addition, an increasing number of these installations demand a mission-critical infrastructure, particularly the industrial concerns, requiring 24 x 7 availability. Down time is measured in both direct and indirect opportunity costs."

It is clear that IDC believes (as do firms like IBM) that bioinformatics is a vast business opportunity (see <http://www.idc.com/itforecaster/itf20010410.stm>).

The large market identified includes the following categories:

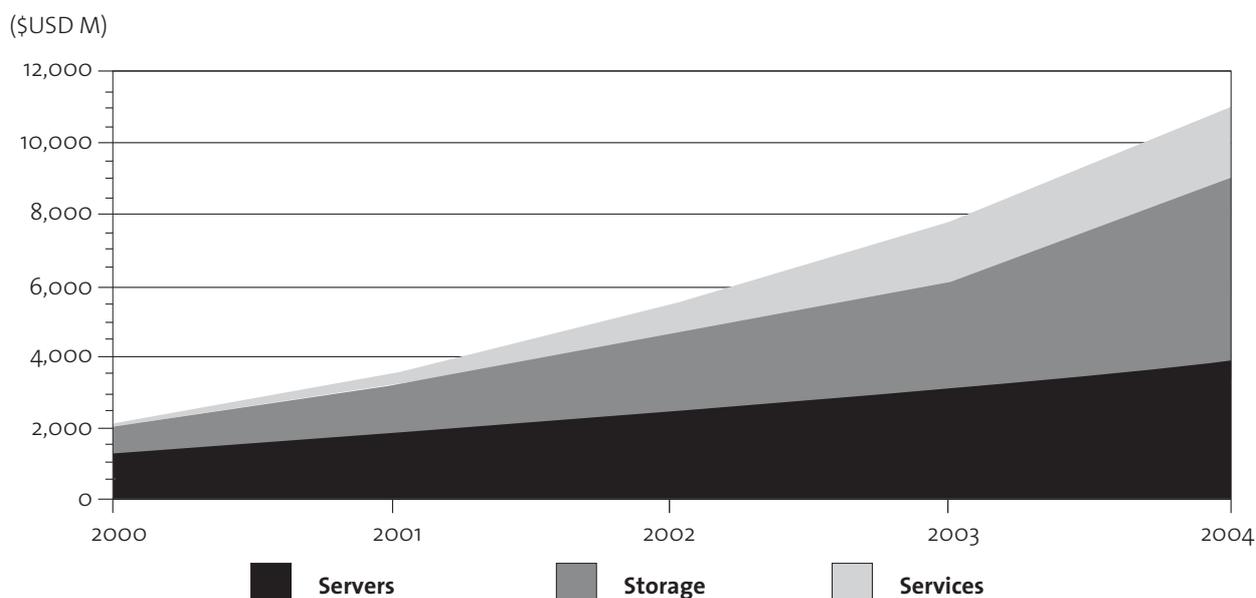
- Hardware: High-end servers, MPPs and other supercomputers, networking gear, workstations, and instrumentation.
- Storage: Disk storage of the order of PetaBytes (1000 TeraBytes), overall.
- Services: Hosting and application provision services for decoding genomes and analysing their functions.

In addition, IDC recognises key opportunities in areas already discussed:

- Software: Knowledge management, database systems, and applications for informatics, sequencing, proteomics, computational chemistry, visualization, and more.

These markets are projected to be collectively worth \$USD 11 billion by 2004, more than 10 times greater than the market for 'pure' bioinformatics companies as shown by Frontline (see above). The servers (computers), storage (disk, tape) and services are believed by IDC to grow in the way as shown in Fig 5 below:

**Fig 5. Biosciences industry growth. IDC, 2001 (<http://www.idc.com/El/eio42001.htm>)**



IDC claims that it is expected that companies that supply services will flourish owing to the fragmented nature of the bioinformatics industry and the need for companies that can help the consumers of these products maximise return on their investment in tools and data from disparate vendors.

It is anticipated that future market drivers of the bioinformatics industry will be initially around 'post genomic' and 'functional genomic' analyses, including:

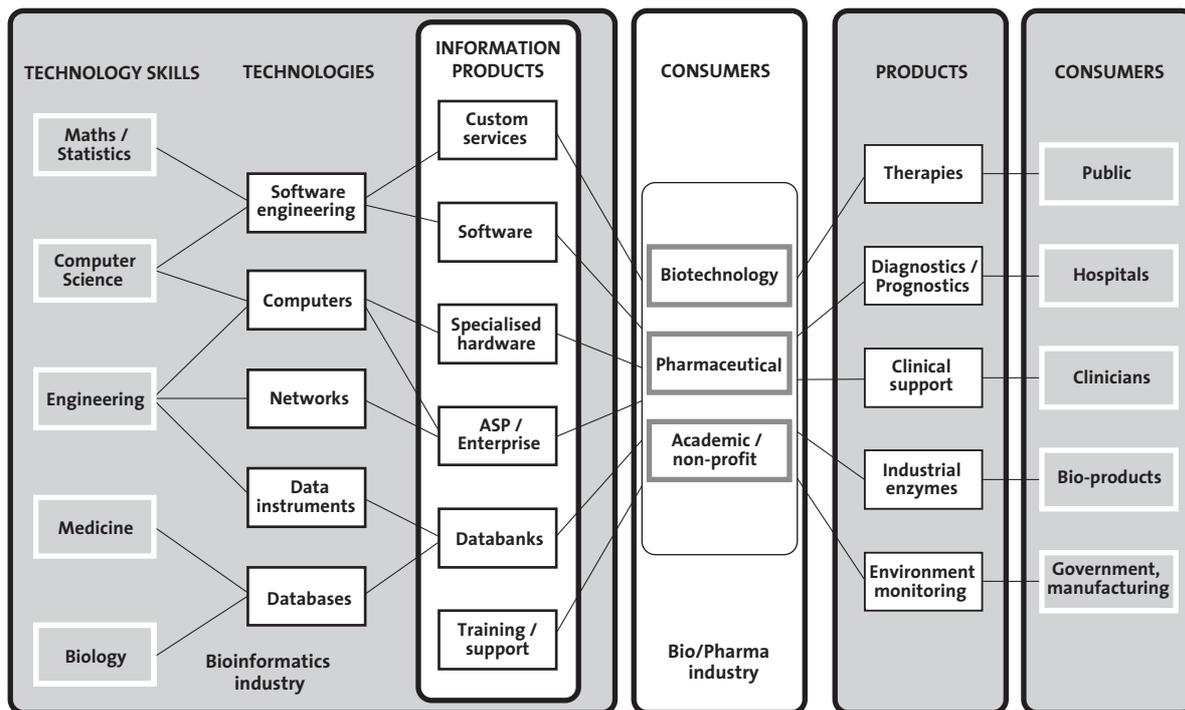
- gene expression analysis;
- population variation;
- proteomic data; and
- 'network' data - metabolism, physiology, development.

### 3.4 Structure and growth of the bioinformatics industry

#### Industry structure

The bioinformatics industry structure and key customers are shown in Figure 6:

Figure 6. Structure of the bioinformatics industry.



© Copyright Biolateral Pty Ltd, 2001

Figure 6 illustrates that the main consumers of bioinformatics industry products and services are the biotechnology, pharmaceutical and non-profit life-science R&D sectors.

The industry itself is composed of multi-skilled practitioners in statistics, computing, engineering, medicine, and the biological sciences. These people create bioinformatics developments that rely on ICT technologies (software, computers, networks, data generating instruments and databases) to make bioinformatics products (custom services, software, novel hardware, ASPs, enterprise solutions, databanks and training and support products). Today, the customers of the bioinformatics industry are small in number and niche: essentially the biotechnology, pharmaceutical, agriculture and academic/non-profit research sectors. However, these industries absolutely require bioinformatics products and services as a part of their product development cycles (eg. therapeutics,

diagnostics, and targeted breeding programs). These products have a vast marketplace (indeed, probably the biggest single market in the world ie. health). There is an opportunity for bioinformatics industries to create products and services that bypass the typical supply chain shown above and service consumers directly (eg. genetic testing services) and go from information to commodity consumption. See Appendix B for more information on stakeholders who would use bioinformatics.

#### Customer profiles

The key customers in the industry today operate in diverse functional categories, including:

- Biotechnology companies producing pre-clinical therapeutics, diagnostics and forensics;
- Biotechnology companies designing new industrial processes or enzymes (eg. thermostable enzymes from analysis of the genomes of extremophile bacteria);

- Agricultural genetics companies producing genetically modified crops/animals;
- Agricultural companies designing means to combat diseases (eg. viral, fungal);
- Pharmaceutical companies screening for new drug targets;
- Genomics companies screening for genes of commercial importance to any industry;
- Environmental organisations using gene based methods to manage eg. control of pests, biodiversity; and
- Government funded research organisations (eg. universities) conducting biological, medical, biotechnology, and bioinformatics research and education programs.

These customers generally represent 'B2B' (business to business) opportunities for bioinformatics firms. In addition, there is an opportunity for bioinformatics 'B2C' (business to consumer) markets such as bioinformatics services in such areas as genetic information resources (eg. databases of genetic diseases and associated health information) and ASP (application service provider) opportunities eg. 'software rental'. Indeed, the bioinformatics company Entigen, which originated in Australia, pioneered the B2C ASP bioinformatics industry.

### Future growth

Future growth will be around expanding the key functional positions identified above with new technologies and datasets (eg. proteomics, biodiversity), integration of disparate data sets, high performance computing (HPC), smarter statistically based analysis methods, and a clean up of 'messy' data to maximise their information content (much as is being done by the Gene Ontology Consortium or - in a broader context - the 'semantic web' initiative). Growth of the biotechnology industries will fuel demand for bioinformatics services and products and this in turn will fuel the industry. Indeed, services are an exciting part of the bioinformatics industry, expected to grow vigorously as a slice of the overall market. As researchers learn more and more about the complexity of living systems and the complex interrelationships that datasets have with each other, new sophisticated products will allow exploitation of that information and will be compelling products to industries which see them as ways of accelerating their own product development processes.

In addition to the substantial knowledge the Task Force members brought to this analysis, future growth was assessed by several processes:

- Analysis of the Australian and international government funded biotechnology and pharmaceutical industry market needs;
- Survey of existing Australian bioinformatics companies;
- Survey of the potential for bioinformatics industry growth as perceived by commercialisation arms of universities;
- Survey of the views of members of the Australian venture capital industry;
- Survey of the needs and growth in bioinformatics activities in the Cooperative Research Centres (CRCs); and
- Discussions with Australian bioinformatics firms.

The results of these discussions are summarised as follows:

- Domestic and international demand for bioinformatics products and services will probably outstrip Frontline's analysis, as the industry includes other domains (eg. biodiversity, neuroinformatics) and bioinformatics activities (eg. services). The broader definition of bioinformatics (see Appendix A) was adopted due to feedback from these industry groups.
- The few bioinformatics companies that are trading today see significant opportunity in the market, especially as their customers (other bioindustry players) grow in number and knowledge of the importance of bioinformatics.
- Pre-competitive research organisations (including universities) are optimistic that they will see growth in bioinformatics start-ups, spin-offs and investment.
- The venture capital industry sees the area as having growth potential: however, the feeling from the more ICT-focused investors is that the bioinformatics industry is more likely to appeal to the biotechnology investors. This view is shared by these investors.
- Some CRCs see an increasing demand for, and production of, bioinformatics technologies.
- Bioinformatics firms in Australia believe there is growth potential given the strong IP base that exists in biological information in Australia.

### **Supply characteristics**

Typically, bioinformatics companies sell their products to consumers through channels such as:

- direct sales to large companies;
- bundling with other products (eg. software sold with instruments such as Axon selling its GenePix software with its array scanners);
- ASP (Application Service Provider) 'pay per use' style models;
- shrink-wrapped software sold through channels such as reagent vendors (eg. John Morris in Australia); and
- via training courses around specialised problem spaces (eg. protein structure prediction courses).

Decisions on buying bioinformatics software, databanks and services vary depending on the cost and infrastructure requirements of the technology or service and the industry segment being sold to. Decisions are often made quickly by end users (eg. relatively cheap ASP or shrink-wrapped software packages). Others require long and slow involved sales cycles (eg. for large commercial databanks and enterprise software solutions sold to pharmaceutical companies).

### **Global markets**

The consumers of 'traditional' bioinformatics are few in number and truly a global group. It is estimated that there are around 400,000 potential consumers of such molecular-biology focused products globally in the research laboratories across all segments. There are, however, considerably more students who consume bioinformatics services and products as part of life-science education programs around the world and this number is rapidly growing as the uptake of bioinformatics education escalates. Further, as bioinformatics companies start to offer products and services that address larger-sized consumer markets (eg. integrated human genetic information resources that are of relevance to clinicians), global markets should grow dramatically.

# Stock-take

## 4 Australia's bioinformatics capabilities

### 4.1 Importance of bioinformatics in Australia

Australia has a major opportunity in the emerging bioinformatics industry, both through tapping into a substantial share of the 'narrowly defined' \$US 5 billion bioinformatics market (and also some of the more broadly defined industry) and then using this to leverage the much larger opportunity in the biotechnology industry.

In his report *A Chance for Change*, the Chief Scientist, Dr Robin Batterham, recognises this challenge, stating that, "*unless appropriate action is taken, Australia runs the risk of not being able to provide emerging industries like bioinformatics and nanotechnology with the required human capital.*"

Furthermore, considering that bioinformatics is a key technology for many bioindustries (pharmaceutical, biotechnology, health, agriculture), and given that these industries are becoming knowledge industries, it is important to consider the economic impact on the broader bioindustries of having a strong bioinformatics sector.

**Biotechnology:** The growth of the biotechnology industry has attracted much interest from State/Territory and Commonwealth governments. Biotechnology is widely seen as being the next great industrial revolution. The industry in Australia has been described in the *Australian Biotechnology Report* (2001) produced by Freehills and by Ernst and Young for the Department of Industry, Science and Resources. That report shows that Australia has a relatively small but vibrant biotechnology sector, a result supported by the Deloitte Biotech Index, which tracks publicly listed Australian Biotechnology companies.

Australia has approximately 35 listed and 190 unlisted core biotechnology companies, and a further 25 listed and 460 related unlisted biotechnology related companies employing more than 5700 staff, with the core companies collectively generating revenues of around \$AU 1 billion per annum. This contrasts with the US industry, which has revenues of \$AU 32 billion per annum. The Australian biotechnology industry is showing strong growth, and sectors with the

greatest number of products under development are human health, equipment and services and agriculture.

The report *Benchmarking Study of R&D Costs in Selected Segments of Australian Biotechnology* by Ernst & Young and Hay Group Strategic Industry Research Foundation notes, on the subject of Australia's general competitive Position in Biotechnology, that there are great opportunities if the nation provides "emphasis on high value Australian-based R&D platform technologies that have wide application eg. in reproductive technology, stem cells, "gene" shears, proteomics, bioinformatics, investing more in enabling technologies".

Further, the importance of bioinformatics is recognised in *The Biotechnology Australia Centre of Excellence Discussion Paper*. This paper states "Australia must develop and apply capabilities in key biotechnology platform technologies, such as functional genomics, bioinformatics and/or proteomics."

State/Territory governments are increasingly aware of the key role bioinformatics plays in the growth of biotechnology. For instance, the Victorian Government's paper *Biotechnology- strategic development plan for Victoria* notes that: "There are a number of 'platform technologies' that enable modern biotechnology. The key platforms, which are currently driving biotechnology advances, come from genomic research (research into the genetic material of organisms) and related disciplines including functional genomics, structural genomics, proteomics and bioinformatics." In response to this, the Victorian Government has established the Victorian Bioinformatics Consortium.

**Agri-industries:** Bioinformatics will also play a key role in growing Australia's agricultural industry, which has gross production worth around \$35 billion annually. Improved genetics and transgenic technologies are believed to be important to adding a sustained value added base to this production and essential for international competitiveness. Australia is an extremely efficient producer, with high-volume low-cost practices that are only likely to grow through agricultural biotechnology innovations that will need bioinformatics. In recognition of this need, the Australian Research Council (ARC) and the Grains Research and Development Corporation (GRDC) will jointly establish the Australian Centre For Plant Functional Genomics that will focus on key platform technologies. Bioinformatics is a key part of functional genomics.

**Health and medical research:** The importance of bioinformatics reaches beyond biotechnology into health and medical research. The Wills Report, *Enabling the Virtuous Cycle- Implementation Committee*, notes "The specialist capabilities that are currently in short supply and should be developed for the post-genome era include ... bioinformatics ...". This report also recognises the importance of bioinformatics in discovery: "Using computers to match and analyse gene sequence data, one of the applications of the new science of bioinformatics, allows better understanding of disease causes and differences in effect across populations". The report recommends that the nation "develop stronger links between universities and external centres of excellence to stimulate PhD and undergraduate training in developing fields and methods such as genetics, molecular biology and bioinformatics."

**Bioprospecting and bioprocessing:** Bioinformatics also plays a key role in the evolving areas of bioprospecting and bioprocessing. Australia has great strengths in these bioindustry segments, which are often overlooked in analyses of their potential for growing the bioinformatics and broader bioindustries in Australia. The House of Representatives Standing Committee on Primary Industries and Regional Services Inquiry into development of high technology industries in regional Australia based on bioprospecting and bioprocessing makes a number of recommendations in this area.

The Standing Committee recommends that the "Commonwealth Government provide additional funding for digitising and networking information about all of Australia's biological resources." This was based on the observation reported to the Committee by CSIRO that, "in the area of biodiversity informatics, Australian science is at the forefront" and [is] technologically ... in a good position to keep up with advances in this area. For example, Australia is chairing the group that is establishing the Global Biodiversity Information Facility in Denmark."

It also notes that there was some concern for "Australia's capacity to undertake bioinformatics" and that "this is a serious issue affecting the national capacity to provide platform technologies for biotechnology". Further, it was recognised that there is a "critical shortage of people with the prerequisite skills and capabilities in bioinformatics." Although steps are being taken to address the shortage of skilled people, CSIRO suggested that more needed to be done. "Attempts to bring together all those with an interest

*in bioinformatics have failed so far [due to the fact that] each of the States is independently developing a capacity in biotechnology when sharing core capacities would be more effective. Such sharing could be achieved with funding from the Commonwealth Government. In such a situation, competitive funding models may not be the most effective way of providing national capabilities".*

The Standing Committee went on to recommend that, "*the Commonwealth Government, in consultation with State and Territory governments, industry and the research community, develop a national strategy for bioinformatics; and assist in funding its implementation so that the necessary infrastructure and skills are available to provide efficient access to information about Australia's biota.*"

Furthermore, the Committee noted that, "*while Australia is relatively well placed in relation to biodiversity informatics, the same is not true of molecular informatics*", molecular informatics being bioinformatics as narrowly defined in this report. In the Committee's report it is noted that, "*molecular bioinformatics are largely in private hands overseas. Australians need to be able to gain access to these data on favourable terms, and 'the bargaining chip that would be most effective would be collaborative arrangements in which we contribute to these databases with annotations relevant to our own biodiversity' from our own databases.*"

**Pharmaceutical Industry:** The 'Pharmaceutical Industry Action Agenda Discussion Paper' (2001) produced by the Pharmaceutical Industry Action Agenda Team in the Department of Industry, Tourism and Resources also identifies the pivotal role that bioinformatics will play in growth of the pharmaceutical industry in Australia. The pharmaceutical industry in Australia is a major industry segment: in 2000 it employed around 14,000 people, had a turnover of around \$AUD 7 billion and exports of nearly \$AUD 1.8 billion. Pharmaceuticals are a giant business: total worldwide sales of drugs in 2000 were \$USD 317 billion.

The Action Agenda paper describes the need for "*specialists in emerging disciplines such as bioinformatics developing analytical technology*" and "*...in the short term, the number of new products emerging from R&D each year is likely to decline for the next couple of years. However, product launches could reach as high as 50 per annum between 2007 and 2008 ... based on ... more efficient proving-up techniques offered by new scientific techniques such as bioinformatics...*". Bioinformatics already is a core component of major pharmaceutical companies' R&D programs: the

paper expands on this point by noting, "*there are a number of new technologies increasingly being used to improve pharmaceutical R&D. These technologies, sometimes referred to as "third-wave pharmaceutical R&D technologies" include genomics and pharmacogenomics, proteomics, combinatorial chemistry, bioinformatics and chemoinformatics, ultra-high-throughput screening, high-throughput DNA screening, large-scale expression profiling, and gene and cell therapy.*"

The opportunity to gain competitive advantages through focusing on key technology areas will be essential to the growth of the pharmaceutical industry in Australia. The discussion paper summarises this by stating, "*there are a number of emerging technologies, particularly in relation to biomedical technology, that could be essential for maintaining the general level of competency within the industry. Given our size, it is not clear whether it is feasible for Australia to be world class in all these disciplines (which include genomics, bioinformatics, cell technologies, x-ray crystallography, mathematics and relevant cross-disciplinary skill sets). If not, it must pick its areas of specialisation carefully.*"

We have a core capability in bioinformatics in Australia and it seems prudent, given the analysis of the Action Agenda Team, that we seriously consider bioinformatics as an area to grow in order to facilitate the development of the pharmaceutical industry.

**High performance computing:** Bioinformatics also has a role in other cross-cutting industry segments: for example, in the Emerging Industries Occasional Paper "The Impact of High Performance Computing Technologies on Australian Industry" (ISR, 2001), bioinformatics is seen as an important application of high performance computing. Indeed, genomics and functional genomics companies such as Celera and GenePro have some of the world's largest computing facilities in order to process the vast datasets they produce.

**Information Technology Industries:** Australia has strengths in the life sciences, biotechnology and information and communication technology (ICT) industries. Consequently, bioinformatics presents a latent opportunity for Australia over the next five years. As both an emerging industry and an enabling technology platform, bioinformatics will be critical if Australia is to benefit commercially from genomics information and capitalise on its position in medical and agri-food research.

Bioinformatics is a true knowledge industry: as many bioindustries will be information driven, they too will become information industries.

## 4.2 SWOT Analysis of Australian Bioinformatics

Australia has great potential for a vibrant bioinformatics industry as the following analysis indicates.

**Table 1. Australian Bioinformatics SWOT analysis**

<b>Strengths</b>	<b>Weaknesses</b>
<ul style="list-style-type: none"> <li>• Strong life science (biology and medical) research</li> <li>• Strong clinical, public health and epidemiological research</li> <li>• Early technology adopters</li> <li>• International recognition of scientific strengths</li> <li>• Good pool of IT professionals</li> <li>• Good pool of life-science researchers</li> <li>• Growing venture capital industry</li> <li>• Growing government support for biotechnology</li> <li>• Western culture implanted in Asia, allowing access to both markets</li> <li>• Large datasets (typically in universities and CSIRO) of commercial value</li> </ul>	<ul style="list-style-type: none"> <li>• Relatively weak track record in commercial biotechnology innovation</li> <li>• Paucity of entrepreneurs and individuals with CEO, CFO and other early-stage biotechnology/ICT company management experience</li> <li>• Few role models for bioinformatics commercialisation</li> <li>• Small local market</li> <li>• Poorly developed bioinformatics education and training</li> <li>• Poor bioinformatics research infrastructure for growing incubation</li> </ul>
<b>Threats</b>	<b>Opportunities</b>
<ul style="list-style-type: none"> <li>• Internet allows delivery of foreign bioinformation products and services into Australia with very low barriers to access</li> <li>• Continued 'brain drain' of Australian expertise (technical and business)</li> <li>• Local leadership moving to other regional hubs showing heavy investment (eg. Singapore with new investments in bioinformatics and genomics)</li> <li>• Poor IP protection, especially of biotechnology products that are based on information products</li> <li>• Reduced basic life-science research due to lack of local bioinformatics support industry</li> </ul>	<ul style="list-style-type: none"> <li>• Large expatriate community ready to come home to fuel the industry if conditions change</li> <li>• Willingness of venture capital community to invest in bioinformatics companies</li> <li>• Regional leadership that can be leveraged</li> <li>• Marriage of biotechnology and ICT industry growth to create an even stronger bioinformatics industry</li> <li>• Strong education infrastructure to tap international e-learning and foreign student markets</li> <li>• Huge existing data resources that could be commercialised as bioinformation products</li> <li>• Growth of domestic bioindustries (eg. biotechnology) will result in increasing demand for bioinformatics companies supplying products and services</li> <li>• Movement of medical research into using genomic and related information will see a convergence of medical informatics and bioinformatics, increasing the need for bioinformatics industry players</li> </ul>

### 4.3 Education, training and skills development

Australia has an internationally recognised world-class education sector. It has led the way in bioinformatics training of life-scientists through the small but effective ANGIS service since 1991. It has strong information, life science and engineering research and education programs in universities that are producing first-class graduates. It has also now a growing number of bioinformatics education programs as shown in Table 2.

**Table2. Current contacts for Bioinformatics Training in Australia (May 2001)**

	<b>University</b>	<b>State</b>	<b>URL</b>
<b>BSc (Bioinformatics)</b>	University of Sydney	NSW	<a href="http://www.usyd.edu.au">www.usyd.edu.au</a>
	University of New South Wales	NSW	<a href="http://www.unsw.edu.au">www.unsw.edu.au</a>
	University of Western Sydney	NSW	<a href="http://www.uws.edu.au">www.uws.edu.au</a>
	La Trobe University	Vic	<a href="http://www.latrobe.edu.au">www.latrobe.edu.au</a>
	Flinders University	SA	<a href="http://www.flinders.edu.au">www.flinders.edu.au</a>
<b>Graduate Certificate in Bioinformatics</b>	Curtin University of Technology	WA	<a href="http://www.curtin.edu.au">www.curtin.edu.au</a>
<b>Major Streams</b>	University of Melbourne	VIC	<a href="http://www.unimelb.edu.au">www.unimelb.edu.au</a>
	Australian National University	ACT	<a href="http://www.anu.edu.au">www.anu.edu.au</a>
	Monash University	VIC	<a href="http://www.monash.edu.au">www.monash.edu.au</a>
<b>Minor Streams</b>	Murdoch University	WA	<a href="http://www.murdoch.edu.au">www.murdoch.edu.au</a>
	Queensland University of Technology	QLD	<a href="http://www.qut.edu.au">www.qut.edu.au</a>
	University of Queensland	QLD	<a href="http://www.uq.edu.au">www.uq.edu.au</a>
	University of Western Australia	WA	<a href="http://www.uwa.edu.au">www.uwa.edu.au</a>
	University of Wollongong	NSW	<a href="http://www.uow.edu.au">www.uow.edu.au</a>
<b>Population / Genetics</b>	University of Newcastle	NSW	<a href="http://www.newcastle.edu.au">www.newcastle.edu.au</a>
	University of New England	NSW	<a href="http://www.une.edu.au">www.une.edu.au</a>

These strengths extend to postgraduate education, as shown in Table 3.

In comparison with the USA, Australia performs very well.

**Table 3. Comparison of Undergraduate Degrees available in Australia and USA. Degrees are only in their infancy and have restricted class sizes (<20 intake) in both countries.**

	Year of Inception	#grad 2000	#grad 2001
<b>Australia</b>			
University of Sydney	1998	13	12
La Trobe University	1999	0	0
University of New South Wales	2001	0	0
Flinders University	2001	0	0
University of Western Sydney	2002	0	0
<b>USA</b>			
Carnegie Mellon	1989	6	7
Rensselaer Polytech Institute	1998	0	2
University of California*	1998, 2001	0	0
University of Pennsylvania	2002	0	0
University of the Sciences in Philadelphia	2000	0	0
Indiana University	2001	0	0

\*Two campuses (Los Angeles, Santa Cruz)

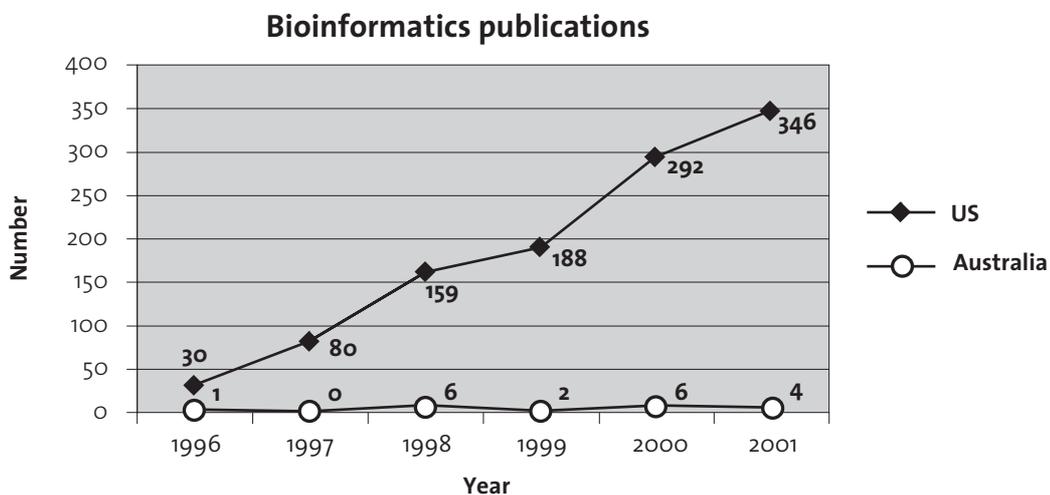
There is a growing need to increase the number of skilled bioinformatics researchers in Australian institutions. The best way to acquire these skills is to have Australian researchers spend significant time in world-class research centres overseas (such as Stanford University, the NCBI in Washington or the EBI in Europe). The Task Force has identified initiatives to ensure such skills and technology acquisition occurs rapidly.

#### 4.4 Research

Bioinformatics research in Australia, depending on the definition, is either considered to be almost non-existent or very small. This is clearly indicated in an analysis of publications in Medline (the major biomedical literature database) with the keyword 'bioinformatics' and shown in Figure 7 (data provided by Rustamzhon Turakulov, John Curtin School of Medical Research, ANU).

Around half the publications in Medline with bioinformatics as a keyword come from the USA. What is startling about these data is that while US (and global) publications in bioinformatics have grown in number dramatically over the last five years in the USA, the Australian publication rate has been low and stagnant.

**Figure 7. Bioinformatics publications.**



The number of groups conducting 'traditional' bioinformatics research is shown in Table 4:

**Table 4. The number of groups conducting 'traditional' bioinformatics research**

Site	Activity	URL
ANGIS	Australian National Bioinformatics Facility	<a href="http://www.angis.org.au">www.angis.org.au</a>
Entigen	Australian spin off bioinformatics company	<a href="http://www.ebioinformatics.com">www.ebioinformatics.com</a>
Murdoch University	Bioinformatics research	<a href="http://www.murdoch.edu.au">www.murdoch.edu.au</a>
WEHI	Bioinformatics research	<a href="http://www.wehi.edu.au">www.wehi.edu.au</a>
Monash University	Bioinformatics research; VBC member	<a href="http://www.monash.edu.au">www.monash.edu.au</a>
CBIS- ANU	Bioinformatics research	<a href="http://cbis.anu.edu.au">cbis.anu.edu.au</a>
CSIRO CMIS	Bioinformatics research; VBC member	<a href="http://www.cmis.csiro.au">www.cmis.csiro.au</a>
CSIRO Entomology	Biodiversity bioinformatics	<a href="http://www.ento.csiro.au">www.ento.csiro.au</a>
IMB, University of Queensland	Bioinformatics research	<a href="http://www.imb.uq.edu.au">www.imb.uq.edu.au</a>
LaTrobe University	VBC member	<a href="http://www.latrobe.edu.au">www.latrobe.edu.au</a>

There is an increasing number of regional consortia coordinating bioinformatics activities, including the Victorian Bioinformatics Consortium (VBC), the West Australian Bioinformatics Consortium (WABC) and the NSW Bioinformatics Consortium (NBC). Recent initiatives include support from the Commonwealth Government for the Peter Wills Bioinformatics Centre at the Garvan Institute and the \$30M investment in the new bioinformatics and functional genomics institute by the University of Sydney and the Medical Foundation of the University of Sydney.

The Task Force recognises the importance of State and Territory initiatives, however, it is clear that all stakeholders would benefit if State/Territory and Commonwealth initiatives were better coordinated to reduce unnecessary duplication of effort and infrastructure.

The difficulty in growing bioinformatics research in Australia is exacerbated by the fact that none of the funding agencies has a speciality in bioinformatics, in stark contrast to the expert panels they have in other disciplines. This lack of coalesced expertise and dedicated funding has stifled bioinformatics research, which needs to grow through careful nurturing by experts.

In the UK, the Biotechnology and Biological Sciences Research Council (BBSRC) and the Engineering and Physical Sciences Research Council (EPSRC) identified the need for a bioinformatics granting scheme and cross-agency support for this fund. As a consequence, BBSRC now supports bioinformatics through a joint Thematic Initiative with EPSRC. The Initiative seeks interdisciplinary applications that demonstrate a blend of information technology and biological sciences. Collaborations with industry are particularly encouraged and applications are invited through an annual call for proposals.

The Taskforce believes that the importance of bioinformatics to underpinning the life sciences in Australia necessitates not just the coordination of all government initiatives, but also the creation of a special research centre that would have bioinformatics as its primary focus. This would allow critical mass to be created in the discipline, resulting in enhanced research outputs and production of skilled staff to fill positions in bioinformatics and other bioindustries. The poor publication rate of Australia in bioinformatics is an indication that Australia must rapidly invest in this important area if it is to create the skills, intellectual property and companies that will allow the growth of bioindustries in Australia.

## 4.5 Australia's position in the global bioinformatics marketplace

As identified in the SWOT analysis above, Australia has a number of opportunities in the global bioinformatics marketplace. However, it is still a major importer of bioinformatics products and services and its competitive position in research and education is weak. Universities such as Stanford in the USA have established bioinformatics research and education programs (see for instance the Stanford medical informatics group at <http://smi.stanford.edu>). It has an e-learning program and has spun off a number of commercial bioinformatics entities.

## 4.6 Commercialisation

Commercialisation of bioinformatics innovation in Australia has been very small. A summary of bioinformatics companies and those with a significant bioinformatics component are shown below. Note that this list does not include companies such as genetic testing companies that are biological information providers as discussed above.

- Desert Scientific
  - <http://www.desertsci.com>
  - Protein modelling software
- BioLateral
  - <http://www.biolateral.com.au>
  - Bioinformatics R&D consulting & contracting
- PKS - Pacific Knowledge Systems
  - <http://www.pks.com.au>
  - eHealth, diagnostics
- ID+PLUS -
  - <http://www.idplus.com.au>
  - BioMetrics
- Examination Services for Psychology Ltd (ESP)
  - <http://www.cogstate.com>
  - novel psychological tests designed specifically for the internet
- SciCon
  - <http://www.scicon.com.au/>
  - Scientific, including bioinformatics, training and consulting

Companies with large bioinformatics components:

- Cerylid
  - <http://www.cerylid.com.au/>
  - Discovery of novel lead compounds for development as new medicines, genomics for target identification
- PSL- Proteome Systems Limited
  - <http://www.proteomesystems.com/>
  - Proteomics
- Bionomics
  - <http://www.bionomics.com.au>
  - Biotechnology company focused on discovery of genes associated with epilepsy, breast cancer and angiogenesis
  - cited in Invest Australia's bioinformatics in biotechnology brochure as a bioinformatics company
- Bio-Gene Bioprospecting Ltd.
  - <http://www.bioprospect.com/>
  - Natural compound screening
- Axon
  - <http://www.axon.com>
  - principally an instrument company, also a bioinformatics software developer

### Barriers to industry development

The barriers that face the commercialisation of bioinformatics products and services fall into two categories:

- Bioindustry-wide issues; and
- Bioinformatics-specific issues.

Bioindustry-wide issues have been addressed in detail in other documents (eg. Commonwealth and State discussion papers on biotechnology). These issues typically are:

- lack of an entrepreneurial skills base;
- small pool of executives with knowledge of the bioinformatics industry;
- lack of track record of commercialisation of innovation from public sector research institutes; and
- lack of pre-seed funds and proof-of-concept support.

Specific impediments to the success of bioinformatics companies include:

- lack of skilled bioinformaticians in Australia;
- lack of adequate facilities, targeted research funds and a centre of bioinformatics excellence in Australia to encourage pre-competitive research;
- too small a local customer base to prove/develop our capability; and
- lack of pre-seed funds targeted to bioinformatics.

The Task Force recognises these challenges faced by the bioinformatics industry in particular and the broader bioindustries in Australia. One way to address these impediments is to increase the skills base in Australian firms in the short term by initiatives that facilitate the hiring of bioinformatics specialists into Australia.

Training, research and other activities create the preconditions for developing Australia's expertise in bioinformatics. This should have a major impact on industry, either as companies integrate this expertise for business development purposes or as new firms spring up to exploit it directly.

The next three years will be crucial in establishing the bioinformatics industry in Australia. If intensive research and training efforts are initiated, then as the Task Force suggests, Australia can expect to see rapid spin-off effects in industry. In fact, this is why the Task Force believes that industry involvement is critical in determining the orientation and design of strategic training and research components.

All possible measures to facilitate the integration of bioinformatics into Australian industrial practice should be explored. Attracting off-shore skills can be addressed through it being made a priority area for immigration and by perhaps targeting these areas in promotion by Austrade and Invest Australia and through our embassies. In addition, marketing campaigns can be carried out at biotechnology conference trade fairs and other such venues. Specialists could be brought in to impart skills in Australia through initiatives of the Commonwealth Education, Science and Training portfolio. Companies in bioinformatics could be targeted to set up Australian operations and to collaborate with Australian firms.

Other possible measures to stimulate industrial demand in bioinformatics could include targeted tax rebates and other incentives for hiring of bioinformatics personnel. This would have substantial benefits, as bioinformatics expertise is rare and expensive. Throughout the world, practitioners in the field receive numerous job offers, and Australia must take into account that fact that Australian companies seeking to recruit bioinformatics specialists have to compete with numerous international groups (commercial and non-commercial) looking for the same personnel, many of whom have a much greater capacity to pay competitive salaries. Such an initiative would allow Australian firms to recruit top bioinformatics staff from overseas.

The small local commercial customer base is a serious barrier to growth of the bioinformatics industry. Investment strategies that grow domestic biotechnology and other bioindustry firms would help reverse this situation, and trends in State/Territory and Commonwealth government policies in recent months and years have helped in this process. All the bioindustries would benefit greatly by strategic government programs that attract firms that would accelerate bioindustry sector growth through providing skilled people, key technologies, and customers for Australian bioindustry firms, including those in the bioinformatics sector. Such firms could be attracted from off-shore with investment incentives, and would accelerate industry growth in Australia. Such companies fall into some/all of the following categories:

- Platform technology companies - genomic, functional genomic and others. These firms would be able to work in partnerships with Australian firms to accelerate their product developments, increasing global market shares, while local firms would benefit by early access to technologies and would be able to influence developments towards their specific needs.
- Data companies - firms selling data and technologies for data analysis and integration. These firms would allow channels for commercialisation of Australian databanks and allow early access to large and sophisticated datasets.
- Large biotechnology and pharmaceutical firms. These companies often represent the end market for small bioindustry firms, and having these companies (with global access to North American and European markets) located on-shore would allow fast and efficient business development and channel relationships to be developed, accelerating growth of Australian bioindustry firms.

### **Commercial value of bioinformatics research**

There is great commercial value in bioinformatics research in the broad: this includes all of the product and service areas identified above. Furthermore, a vibrant bioinformatics industry will be essential for the growth of Australian bioindustries such as the biotechnology, pharmaceutical and agriculture industries.

Australia has vast amounts of pre-competitive (publicly funded, usually 'basic') research that is ready to be commercialised. This includes large quantities of unique life-science data in areas such as:

- Forestry
- Agriculture
- Veterinary Science
- Biotechnology
- Biodiversity
- Clinical data
- Public health data

These datasets have substantial value. Using bioinformatics approaches, they relate to the possibility of discovery of new products and services such as:

- Improved breeding strategies for economically important plants and animals relevant to veterinary, agriculture, forestry and other markets;
- Genetic based diagnostics and identification of plants and animals as they relate to quarantine, horse-racing, domestic pets and other markets;
- Development of novel therapeutics targeted to well defined diseases with large integrated datasets of clinical, epidemiological and genetic data;
- Novel diagnostics for human diseases based on molecular genetic methods, underpinning a potentially vast molecular pathology (human tissues and pathogen characterisation) market;
- Biodiversity assessment, environment monitoring and conservation; and
- Information services (eg. sales of access to these and other datasets) and related services (eg. ASP access to tools, high performance computing, training and other bioinformatics value-add services).

### **Barriers to the commercialisation of bioinformatics research**

There are several barriers to the commercialisation of bioinformatics research: many of these have been described above. The greatest barrier is the fact that much of the valuable IP in the datasets collected by Australian researchers is still in a 'raw', pre-competitive state. Other barriers are the lack of skills in the community in bioinformatics and awareness of the commercial value of this information. The major barriers are thus digitisation, organisation, education, commercialisation and infrastructure.

One of the major difficulties small to medium enterprises have in the development of bioinformatics products and services is access to world-class facilities. Bioinformatics developments require access to expensive facilities, often beyond the means of small firms. Product development frequently needs access to expensive high performance computing, databanks, software tools and bioinformatics experts: these should be provided through rationally organised, coordinated national facility infrastructure. This would save significant amounts of duplicated effort in small firms as they start up and engage in bioinformatics product development. In addition, such facilities would provide the much-needed concentration of expertise required in Australia to help grow the industry.

### **Extent of commercialisation of bioinformatics research**

As discussed above, Australia has a poor record in the commercialisation of bioinformatics research. A good example of a previous success was Entigen, which was spun out of research conducted at the University of Sydney in running the ANGIS service. In addition, PKS has links with the University of New South Wales and the Garvan Institute, and SciCon has strong links to Murdoch University.

The Task Force recognises that one of the most significant barriers to the commercialisation of bioinformatics research is the lack of pre-seed funds to take innovations from the pre-competitive (often publicly funded) environment onto the next (earliest) stage of commercialisation. A bioinformatics pre-seed fund would help meet this gap, although the current Commonwealth Government pre-seed fund initiative and programs such as the Biotechnology Innovation Fund (BIF) and Innovation Investment Fund (IIF) have to some extent addressed this need.

### **Strengths and weaknesses in Australia's commercialisation capabilities**

Australia has great opportunity to commercialise innovations in bioinformatics. Although linkages between institutions that have pre-competitive research (universities, hospitals and CSIRO) and industry are growing through CRCs and other activities, the recognition and exploitation of intellectual property is still largely lacking. Some Australian bioinformatics companies have attracted financing from consortia of venture capital firms from both biotechnology and ICT industries sectors (eg. Entigen had Rothschild and Allen & Buckeridge as founding investors). In general, the biotechnology investment community is more aware of the opportunities in the bioinformatics industry, and the ICT investor community is only beginning to become aware of the opportunities.

The strengths of Australia's commercialisation capabilities include a growing number of well-resourced investors (due to increases in funds through the Commonwealth Government's Innovation Investment Fund and other schemes). Australia has a strong entrepreneurial culture, however, this has yet to arise in force in the bioindustries in Australia in general. The lack of industry-experienced CEOs and experienced managers are some of the major problems: this has been identified in recent analyses of the biotechnology industry in Australia, and applies equally to the bioinformatics industry.

The opportunity for growth of the bioinformatics and related bioindustries would be supported by several key activities: an improved research base in bioinformatics, enhanced skills development; pre-seed funding to take these innovations towards commercialisation; and formation of clusters or a centre of excellence in bioinformatics.

The general innovation-commercialisation skills in Australia need to be enhanced in order to capture the maximum opportunity from the emerging bioinformatics industry. Governments could assist by looking at ways to enhance the business enterprise skills of promising SMEs in the sector so that they can better access existing funds such as the pre-seed fund, R&D START Program, COMET and private sector investment. A number of Australian and international observers have noted the difficulties of securing venture capital for start-up firms in this field. Many IT venture capitalists believe that biotechnology focused venture capital firms are best equipped to grow bioinformatics firms, and many biotechnology venture capital firms do not have

the expertise to judge the investment worthiness of bioinformatics enterprises. Some existing venture capital firms have shown a willingness to invest in bioinformatics companies. If Australia can rapidly build a research network of high quality resources in genomics, proteomics, and bioinformatics, then this expertise should attract investors.

Australia is uniquely positioned to capture regional (principally Australasian and South-East Asian) markets. In order to do this, Australia needs to promote its strengths in bioinformatics, both pre-competitive and competitive-based activities. One disadvantage is that Australia currently has no coordinated and single point of access regarding its bioinformatics activities and capabilities.

A further strength of Australia is its robust and innovative ICT sector. Bioinformatics firms need business people, bioinformaticians, biologists and software engineers. Australia has an abundance of biologists and software engineers, and the cost of development (due to comparatively lower salaries internationally) is highly competitive in Australia (around half that of the USA). This strength provides Australia with a highly competitive position for the growth of bioinformatics firms.

## **4.7 Industry role models, linkages and partnerships**

There are a number of professional organisations that are relevant to the growth of the Australian Bioinformatics industry:

1. Australian Biotechnology Association (AusBiotech)  
<http://www.ausbiotech.org/>
2. Australian Information Industry Association (AIIA)  
<http://www.aiia.com.au/>
3. Australian Venture Capital Industry (AVCAL)  
<http://www.avcal.com.au/>
4. Australasian Tertiary Institutions Commercial Companies Association (ATICCA)  
<http://www.aticca.com/>

In addition, the nascent Australian Association for Bioinformatics (AAB) is being formed, and will help draw together key players in research, bioinformatics facilities, education and industry. The Australian Association for

Bioinformatics (AAB) is a non-profit organisation that is a peak bioinformatics industry body that is providing a national approach in the field of bioinformatics by building connections between bioinformatics professionals, local and regional centres of bioinformatics and national IT infrastructure, to coordinate existing strengths, add value and build new capability.

Specifically the AAB's goals are:

1. To promote a managed and cooperative approach to research and education in the field so as to maximise the benefits from that research and education and assist the development of Australia's academic and industrial strengths in bioinformatics.
2. To encourage industry development and to provide incentives for commercialisation in such a manner as to ensure that the maximum benefit accrues to Australia, including Australian industry and the Australian economy generally, provided that any such commercialisation is ancillary to and supportive of the other objectives.
3. To enhance access to national and international bioinformatics resources, and improve public knowledge of bioinformatics and awareness of its importance in biotechnology and biomedicine.

The Task Force recognises the need for a mechanism to bring together all the necessary stakeholders in Australia. The Task Force believes that a peak industry body for bioinformatics is needed to help break down the isolation of industry players, build strengths in bioinformatics, and represent and champion the industry. Existing peak bodies such as Ausbiotech could help champion the needs of the Australian bioinformatics in the biotechnology sector. However, the cross-cutting technology and emerging nature of the bioinformatics industry means that there is a need for a peak industry body to represent all aspects of the industry, including research, education, national bioinformatics facilities and industry growth.

The Australian Association for Bioinformatics (AAB) has formed recently to achieve these goals. Modelled after APAC (the Australian Partnership for Advanced Computing), AAB could take the role of the required peak body and could sit shoulder to shoulder with, and handle cross-cutting issues of, other peak bodies such as Ausbiotech (the Australian Biotechnology Association) and the Australian Information Industry Association (AIIA). It could be responsible, in

cooperation with other industry bodies, for overseeing the pursuit of opportunities for the ongoing growth of the industry.

#### **Industry role models, case studies and leaders**

There are many international, and a handful of Australian, role models that provide good case studies and show industry leadership. Australia's first bioinformatics company, Entigen, was formed in 1998 with financing from Australian and international venture capital and corporate investors. However, due to finance impacts flowing from the 11 September 2001 events in New York, Entigen is in the process of ceasing operations. Other Australian bioinformatics firms have established in recent times, including the bioinformatics company BioLateral. BioLateral's business model focuses on generating revenues through consulting, contracting and professional training, and the company has an internal R&D program with support from the Commonwealth Government's BIF (Biotechnology Innovation Fund). Other Australian companies have a significant bioinformatics component to their operations, such as PSL (Proteome Systems Limited). Both BioLateral and PSL have identified the need to partner with large firms to grow and to reach global markets, and indeed both companies established relationships with IBM in 2001 to help achieve these goals.

## **4.8 International approaches**

There are emerging numbers of international studies on the importance of bioinformatics - as an industry itself, as a facilitator of bioindustries in the broad and as a means of growing the ICT industry.

The Québec Government in Canada (in the report 'Bioinformatics in Québec: a cornerstone of bioindustry development') has identified bioinformatics as an emerging industry that is intimately linked with the growth of their bioindustries and for the need to remain competitive in a global information economy. Québec has well-developed biotechnology and pharmaceutical industry sectors.

The recommendations made by the Task Force have drawn on this analysis of the role of bioinformatics in the bio-knowledge economies, and the essential role that bioinformatics plays in the growth of these industries. For more information, see <http://www.cst.gouv.qc.ca/reBioinf.html>.

The bioinformatics industry in the USA is perhaps better established than in any other part of the world. The USA has a major bioinformatics centre, the National Centre for Biotechnology Information (NCBI- <http://www.ncbi.nlm.nih.gov>) and a number of bioinformatics training programs (see Table 3) as well as established firms (eg. Accelrys and Informax) and associated companies (eg. Celera and IBM). Many research groups lead the world in the USA, such as those based at Stanford University and around the many genome centres in the USA.

The situation is similar in Europe, where the European Bioinformatics Institute (EBI) is the European centre for bioinformatics. This facility serves European nations research needs, and houses numerous bioinformatics research groups, many with commercial outcomes. The EBI shares a common campus with the Sanger Centre, a large genomics and bioinformatics centre, and the Human Genome Project Resource Centre (HGMP-RC), making this site in Hinxton (near Cambridge) a powerhouse of European bioinformatics. Many bioinformatics companies have operations nearby, such as Lion Biosciences and BioWisdom. Bioinformatics education in Europe is strong, with key groups in the UK (eg. Manchester University) and Germany (eg. Bielefeld University).

Bioinformatics in Asia is strongest in Japan and Singapore - Japan hosts the DDBJ DNA sequence database, equivalent to the GenBank database maintained by the NCBI in the USA, and there are numerous world class bioinformatics research groups in Japan (eg. the National Institute Of Genetics in Mishima). Several firms in Japan have bioinformatics activities, including IT companies such as Fujitsu. Singapore sports two bioinformatics centres - the established Bioinformatics Centre (BIC) at the National University of Singapore and the newly established Bioinformatics Institute (BII). Bioinformatics has been targeted as a major growth industry in Singapore.

#### **4.9 Encouragement and support of Australian start-up bioinformatics companies**

Bioinformatics firms need much of the encouragement and support that is needed by other bioindustry firms. There is a need for mentoring and for financing through pre-seed, proof-of-concept and venture capital funds. In addition, the unique nature of the industry needs input from both ICT and biotechnology industry experts, and an increased awareness of value of bioinformatics IP. These challenges are exacerbated by the small number of bioinformatics firms in Australia, resulting in few mentors and well-informed investors and advisors. In addition, the relatively small biotechnology industry sector in Australia poses challenges to the industry's growth, as this results in few local firms which can become collaborators, partners and customers.

# Issues and opportunities

## 5 Bioinformatics Industry Development: issues and opportunities

Discovery in the life-sciences has changed dramatically in the last decade, which was in many ways the 'genome era'. In 1991, Nobel Laureate Walter Gilbert predicted a paradigm shift in the discovery process when he said:

"The new paradigm, now emerging, is that all the 'genes' will be known (in the sense of being resident in databases available electronically), and that the starting point of a biological investigation will be theoretical. An individual scientist will begin with a theoretical conjecture, only then turning to experiment to follow or test that hypothesis." (Towards a paradigm shift in biology. *Nature*, 349:99).

In February 2001 the complete human genome was published, joining the ranks of many other complete genomes of other organisms, and 10 years after this prediction was made, biology has transformed into an information science with bioinformatics providing the 'glue' and technology to drive discovery.

The critical actions for growth of the bioinformatics industry are to build capabilities in bioinformatics that will underpin bioindustries growth in Australia. Actions need to be in four major areas:

- Education and skills development
- Research
- Infrastructure requirements
- Commercialisation

Bioinformatics technologies are needed to digitise, organise and mine Australian life-science datasets. These datasets span our unique clinical, biodiversity, and specialised life-science knowledge base - information that will underpin future developments in biotechnology and related industries. A domestic bioinformatics capability will 'unleash' the potential value of these datasets, and the deep domain knowledge of Australian life-scientists who have collected and curated these data will ensure both the effective organisation and exploitation of this information. Australia needs not only to develop a keen sense of how to 'prospect' this vast legacy of data, but also how to recognise the commercial value it holds, and then to convert it from

pre-competitive (often undigitised) formats into a form ready for commercial exploitation using bioinformatics tools.

A strong domestic bioinformatics capability would accelerate the development of the bioindustries generally in Australia and introduce a cascading series of benefits by ensuring that Australian firms have access to capabilities to help them develop biotechnology-based products and services. Growing a domestic capability in bioinformatics would provide national benefits arising from flow-on benefits to all bioindustries. A domestic bioinformatics industry would be able to service these bioindustries at greatly reduced cost and would help to retain intellectual property and expertise within Australia. An Australian-grown bioinformatics capability would improve the international perception of Australia as a global bioindustry centre and in turn, increase Australia's credibility to attract research and commercially related foreign direct investment.

There is a broadly accepted view that support for knowledge-based industries is critical to national economic growth, and bioinformatics is perhaps one of the best examples of such an industry. Bioinformatics represents an industry heavily reliant on knowledge and intellectual property (and capitalising on these) as the basis for economic growth.

Numerous studies point to the major returns Australia would gain from investments in innovation and knowledge industries. Australia needs to aim to 'punch above its weight' in the global knowledge economy. By doing so, it would benefit from greater international technology transfer and engagement with global knowledge industries. All of these desirable outcomes can only be achieved through a coordinated national strategy.

The following section proposes specific industry development opportunities and actions based on the preceding discussion and analysis based around the four major areas of education and training, research, infrastructure and commercialisation identified in section 3.2 above.

## 5.1 Training and skills development

### Issues

The Taskforce has identified that the global shortage of human resources in bioinformatics is also endemic in Australia. A vibrant bioindustry sector in Australia will depend on biologists who are skilled in information technologies for managing and analysing large data sets and information scientists (computing, mathematics and statistics) and software engineers who are aware of the needs and opportunities for discoveries in the life sciences. Graduate programs at the university level are being established, and there is a small number of professional intensive and postgraduate training programs coming on line. These initiatives will start to redress the critical shortage of bioinformatics skills in Australia.

Strong institutional partnership between government and universities will be vital for the success of such programs. International competition for bioinformatics courses is growing, and institutions such as Stanford University in the USA are already offering such courses. Without a local, equivalent quality content, graduates will leave Australia for training and most likely will gain immediate employment on graduation overseas. However, formal courses are relatively slow to come on line and industry demand is acute and strong. Thus, in the short term, fast-track solutions are required. In the medium term, universities with their undergraduate and postgraduate programs will take over this need.

The need for bioinformatics training has been recognised by various state governments - the Victorian Bioinformatics Consortium has a strong emphasis on bioinformatics training and the Queensland Government has recently recognised a similar need. (The Queensland Government's policy fact sheet on skills, training and education says a key issue is, "the development of skills suitable for the new emerging industries, such as bioinformatics and nanotechnology"

<http://www.sd.qld.gov.au/dsdweb/htdocs/global/publications.cfm?ID=6565>).

### Options for action

The options for filling the skills gap include:

- Immediate term
  - Rapid intensive professional training programs of Australian researchers (weeks to months in duration)
  - Hiring of skilled staff from overseas
- Medium term
  - Post-graduate training programs to re-skill biologists and information scientists in bioinformatics (12-24 months in duration)
  - International post-doctoral fellowships (36 months in duration)
- Long term
  - Undergraduate training programs in bioinformatics (ideally 4 years plus honours to cover the required materials, and all with specialist bioinformatics units from year one)

The Task Force noted that bioinformatics training initiatives must create practitioners from sources other than computer science programs, as companies in communications, multimedia and other sectors already aggressively recruit these graduates. It may be more productive to train bioinformatics specialists by providing extra training to people with a specialty in the life sciences rather than trying to convince computer science students to specialise in the field. In the longer term, however, growing bioindustry demand for bioinformatics specialists and the rising salaries that result will draw increasing numbers of computer scientists per se. Growth in the number of skilled IT professionals in the marketplace is also required to fill key positions for bioinformatics systems engineers.

Industry must be associated with these educational and training initiatives, especially companies already active in bioinformatics that could serve as locations for trainee internships.

In the immediate term, intensive training programs could rapidly train researchers to a skill level where they are capable of sufficiently mastering the main aspects of bioinformatics in order to meet the most urgent needs of research and industry teams in Australia. Recruitment of skilled staff would also fill immediate staffing needs, although this would be difficult given the internationally competitive nature of the bioinformatics labour marketplace.

In the medium term, post-graduate programs should be developed and targeted fellowships could be used to train the core group of research leaders we will need in the years ahead to develop bioinformatics research and training in Australia. Fellowships for training in bioinformatics in centres (eg. bioinformatics centres such as the NCBI and EBI) outside Australia would aid in skills acquisition. Many of these fellows would return to Australia with world-best-practice skills in bioinformatics that would help accelerate the industry's growth in Australia. On face value this strategy involves an element of risk, as there is the possibility that the most promising specialists end up being recruited by companies or universities outside Australia. However, until Australia develops a solid PhD program in bioinformatics, graduate and postgraduate fellowships remain an excellent means of training scientists in the field. The best way to ensure that they come back to stay is to stimulate strong demand for their services in Australia's university and industrial communities. Even after the creation of graduate programs in bioinformatics, it could still be in Australia's interests to send students abroad to train or upgrade their skills.

In the long term, the growing number of bioinformatics undergraduate programs will help produce a pool of skilled graduates who will be able to fill bioinformatics positions. However, the instructors of these courses should be the targets of focused and intensive retraining to ensure their skill base is at world's best practice levels.

### Implementation considerations

The above options for action each have considerations for their implementation:

- Immediate term
  - Intensive professional training programs of Australian researchers would ideally be a public-private partnership arrangement between bioinformatics firms and institutions of higher education.
  - Recruitment of skilled staff from overseas will require higher compensation levels and flexibility in employment arrangements (eg. joint appointments).
- Medium term
  - Post-graduate training programs are needed, and these will require re-skilling of academics in bioinformatics (life and information scientists alike) to run these programs effectively

- International post-doctoral fellowships would ideally be funded through a targeted program that ensures reasonable numbers of individuals are skilled up in this way
- Long term
  - Undergraduate training programs in bioinformatics, much like postgraduate programs, should see the academics running these programs formally reskilled in bioinformatics and with degrees that are ideally longer than a standard 3 year degree, reflecting the deeper content of the subject area (spanning both information and life science materials).

## 5.2 Coordination of competitive and pre-competitive research

### Issues

Several issues face the development of a vibrant bioinformatics research base that will fuel industry growth:

- Pool of skilled practitioners;
- Coordinated national approach, in cooperation with State and Territory initiatives;
- Focused granting agency panel(s) on bioinformatics; and
- Access to infrastructure.

An active bioinformatics research base needs a pool of expert practitioners: this issue is addressed in the preceding section.

Currently the bioinformatics research activities across the nation are not well coordinated. Much of the bioinformatics and bioindustry growth in Australia is occurring at the State/Territory level. Many States and Territories now have major biotechnology and ICT activities, and bioinformatics cuts across these sectors. Some have targeted bioinformatics initiatives. It is important that all relevant Commonwealth Government initiatives engage and leverage State/Territory initiatives and vice-versa.

One of the challenges facing the industry's growth is that there are no panels of experts in any of the funding agencies that have a specific expertise in bioinformatics. There is a strong need for increased investment in basic bioinformatics (pre-competitive) research. The challenge here is to create

panels that can assess bioinformatics research applications. In addition, if each funding agency were to create their own targeted bioinformatics panel, there would be few expert researchers to sit on all these panels, and so, the quality of assessment would be low and/or the same experts would need to sit on multiple panels.

Bioinformatics research needs good access to infrastructure: high performance computing, bioinformatics software, biological databases, good networks and other capabilities. These needs are addressed below.

### Options for action

Australia would ideally establish a joint thematic initiative along the lines of the UK BBSRC and EPSRC funding agencies, to encourage bioinformatics research in Australia. The ARC, NHMRC and RDC's (especially the GRDC) would be ideal agencies to manage this process. This need was also identified in 2000 by an analysis of bioinformatics needs in Australia, led by the NHMRC with the participation of the ARC and CSIRO. The ARC and GRDC jointly established 'Australian Centre For Plant Functional Genomics' could be a good model for such a cross-agency initiative.

### Implementation Considerations

The creation of a targeted bioinformatics funding stream should not require additional resources. Indeed, many funding agencies (eg. NHMRC, ARC, GRDC) already support bioinformatics research. Rather, what is needed is a panel of bioinformatics experts to judge research grant applications. This panel could be shared between the agencies to save each body establishing their own and drawing too heavily on the small number of experts in the nation.

Recruitment and training of bioinformatics researchers for public-funded research should be coordinated with other initiatives that need bioinformatics experts (eg. biotechnology industry, the higher education sector).

Bioinformatics is by its very nature interdisciplinary. In order to facilitate research, collaborations and partnerships between life scientists and information scientists should be encouraged so that bioinformatics innovations can grow at the interface.

### 5.3 Building a strong bioinformatics infrastructure

#### Issues

The needs for growing the pool of skilled bioinformatics personnel and for stimulating bioinformatics research are addressed in the section on bioinformatics research above. Research, education and commercial bioinformatics operations need access to world-class bioinformatics infrastructure that includes:

- High performance computing facilities;
- Large and secure data storage capabilities;
- Up-to-date and maintained industry standard bioinformatics software and databanks;
- Support for use of these facilities; and
- Concentration of bioinformatics excellence to allow interactions, synergies and commercialisation.

#### Options for action

In order to build a focal point for bioinformatics research and industry development in Australia rapidly, a centre of bioinformatics excellence (COBE) should be established.

The COBE could be structured as a bioinformatics resource within a network or as a single, concentrated resource at a special centre or institute. Although these two options are based on different organisational strategies, they share common components. Either way, it is essential to create an organisation with the necessary critical mass to allow rapid bioinformatics developments as happens at other bioinformatics centres (such as the NCBI and EBI) as quickly as possible.

A network model is perhaps best suited to the geography of excellence in bioinformatics in Australia. It may even be possible to strengthen a network focal point gradually and move toward a more concentrated option. The goals of this centre would be to:

- create an environment that will attract and retain bioinformatics specialists;
- support graduate training in bioinformatics through advanced research;
- establish a foundation for active university-industry cooperation in bioinformatics (eg. a CRC);

- provide a skills incubator for industry (eg. through technology incubators);
- serve as a business incubator by creating spin-offs;
- place Australia on the international map for its bioinformatics capabilities;
- attract large companies (biotechnology, pharmaceutical and IT) with a specific need (to use or develop products) for such capabilities;
- establish the basic core of a future bioinformatics industrial cluster; and
- provide a co-location facility for shared infrastructure (computing, storage, software, databanks).

Developments in bioinformatics generally occur in conjunction with genomics and functional genomics projects drawing upon bioinformatics applications. The COBE would need to be located near to, and collaborate with, leading life-science public-research groups and firms.

Start-up and small to medium enterprises (SMEs) in the bioindustries need access to world class infrastructure to conduct bioinformatics related research and development. This infrastructure is composed of high performance computing, networks, software, genome and functional genome databanks and specialised facility staff for bioinformatics research. Without such infrastructure, industry will be forced to create its own at greatly increased costs and reduced efficiencies. Australia has a national bioinformatics facility (ANGIS). One option for providing this necessary infrastructure is to grow the ANGIS facility and expand its capabilities. As a part of the COBE, this capability could be centralised or distributed, much along the model of the ANGIS 'Node' concept, where facilities exist in each State and Territory.

#### Implementation Considerations

Bioinformatics facilities in Australia, both public and private, are proportionately small in comparison with other developed nations. We have no facility equivalent to the NCBI or EBI, and no firms with bioinformatics capabilities even a fraction of those established by firms such as Celera. It is hard to imagine a vibrant bioinformatics base, and indeed a strong bioindustry sector, without such investment.

A COBE could be created by a number of steps:

- Attracting overseas firms to establish major bioinformatics activities in Australia;
- Encouraging Australian firms to invest in data-mining equipment (bioinformatics) to move product development activities into this lucrative and promising space;
- Establishing a major bioinformatics centre in one or several locations, as a loose federation or a single organisation, centralised or distributed. Either way, investment in building 'nodes' or a core-facility will be critical; and
- Generally increased investment in bioinformatics infrastructure through facilities such as ANGIS.

## 5.4 Industry growth and development

### Issues

Growth of the bioindustries in Australia will be key to the future economic health of the country. Bioinformatics will be a key component of such growth, as such industries increasingly become information-driven.

Issues facing the industry include:

- Trade
  - Improving the biological-information trade balance
  - Improving balance of payments for the bioindustries
- Investment
  - attracting foreign investment to Australia
  - Awareness raising in the investment community of the opportunities in bioinformatics
- Commercialisation
  - Building entrepreneurship
  - Developing strategic international linkages / partnerships
  - Intellectual Property management
  - Regulatory issues
- Industry leadership
  - Need for a peak bioinformatics body
  - Coordination with other peak bodies eg AusBiotech
  - Adoption of strategies by other professional societies
- Managing bioethics issues

### Options for action

Australia has a unique position in the region: its cultural linkages and history give it strong ties with the USA and Europe, yet it is part of Asia. As a consequence, Australia could be a gateway for many industries, but especially emerging industries such as bioinformatics. Australia needs to grow its internal bioinformatics capabilities rapidly and then use these to strengthen regional alliances and become the focus of the bioinformatics industry. Already other countries such as Singapore are investing heavily in bioinformatics. However, Australia has strong ties with western countries, a highly innovative culture and strong economy that means it could take the regional lead in bioinformatics.

One mechanism to achieve regional leadership is to host the secretariat for the regional bioinformatics body, APBioNet (Asia Pacific Bioinformatics Network). Engagement of Austrade and Invest Australia would be key to opening up industries in the region and attracting new resources into this emerging industry base.

The Task Force proposes careful monitoring of developments in the field over the next three to four years, as much from a scientific point of view as an economic or even social one. Bioinformatics industry growth in Australia needs to have a whole of government approach, with input from peak bodies such as AAB, Ausbiotech and AIIA.

These monitoring and technology watch initiatives must not focus exclusively on bioinformatics. The entire field of genomics and functional genomics and their related applications deserve careful monitoring in order to understand how bioinformatics is developing in Australia, assess progress and identify difficulties. Areas requiring special attention include the rapid development of bioinformatics training programs, venture capital reactions to business start-ups, infrastructure and equipment needs, and the intervention of other players such as the funding agencies, Commonwealth agencies, and Australia's university research base.

Through all of this, it is also important to keep the public informed and allow people to express opinions regarding the social and ethical concerns related to the development of bioinformatics. Because these technologies have implications for human health as they are applied to medical problems, the ethical use of these technologies must be considered.

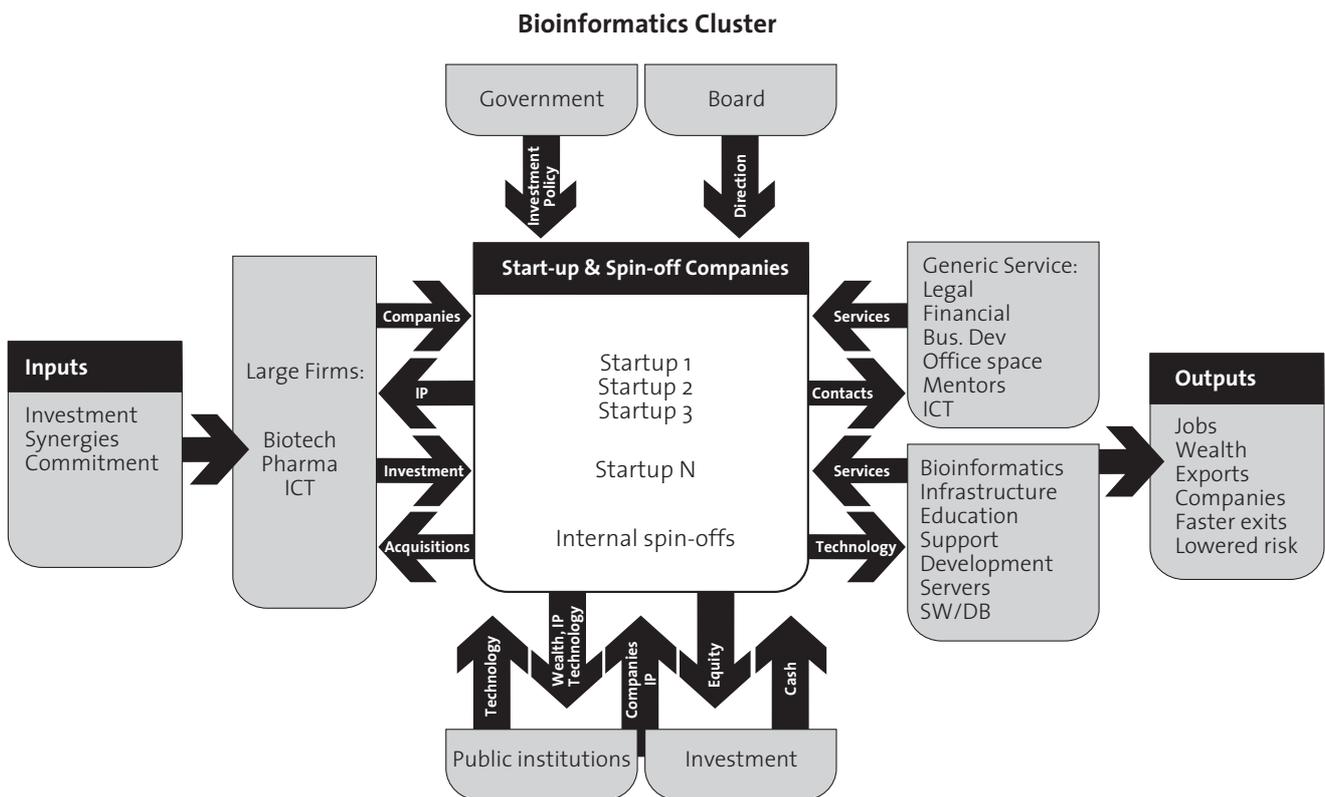
### Implementation Considerations

Achieving the commercialisation of bioinformatics innovation and growth of the bioindustries around this growth will require a number of steps. Many of the issues facing the successful growth of the bioinformatics industry are common to high-technology industries in general.

One of the best mechanisms for creating a vibrant bioinformatics industry might be to create a bioinformatics industry cluster. The cluster could be co-located with the COBE so that infrastructure is close to industry growth.

The key ingredients of a bioinformatics cluster are shown in Fig 8.

**Fig 8. A model for a Bioinformatics Cluster. Derived from discussions with the ‘Genesis Program’ - Matthew Griffiths.**



The key ingredients of the cluster are:

1. **Bioinformatics firms** - Typically small and start-up or spin-off in nature, these companies would interact with each other, synergising and leveraging from each other's strengths, IP and business knowledge. In addition to core bioinformatics companies, molecular biotechnology, instrumentation, medical informatics and other related firms should also be attracted to the cluster, as they would provide 'relevance' to bioinformatics developments through providing inputs to bioinformatics developments. They would also be consumers of bioinformatics technologies.
2. **Service industries** - These include firms providing legal, financial and other services typically needed by start-up companies. These industries will gain access to a marketplace, and the bioinformatics firms gain access to economies of scale. The service industries could also provide a critical mentoring program to the early-stage firms as they find their way into the commercial world. In addition, bioinformatics infrastructure is a critical part of the cluster, and this could be provided by the COBE.
3. **Large firms** - these would provide investment into the cluster companies, use the cluster as a home for their own spin-offs, have contract research provided by the cluster companies and use the cluster as a source of IP licensing and acquisitions. Ideally, both life-science (biotechnology and pharmaceutical) and ICT firms would be engaged in this role.
4. **Investment community** - would use the cluster as an ideal investment focus, being able to reduce risk by investing in firms that have themselves a lower chance of failure due to the services and relationships provided by the cluster. The relationships with the big firms and their interest in technology licensing and acquisition also speeds the time to achieve an exit for the investors.
5. **Public institutions** would be a key part of the cluster, as they would feed the cluster with IP (licensing), spin-offs, services (eg. education, contract research) and gain by having access to a safer environment for their innovations to be commercialised in a supported, lower risk environment.
6. **Government involvement** in the cluster would be crucial to ensure effective leverage of government support programs and to use the cluster as a 'hot-house' of

technology incubation that would expose any policy impediments needing change to ensure that the cluster was globally competitive.

7. **Board of Management** - composed of key individuals with knowledge of the industry and commercialisation processes, the Board would ensure that the cluster managed inputs (investment, synergies between players, commitment of participants) to maximise outputs (wealth, jobs, IP, exports, lowered risk to all participants, new firms). The Board could also take an active role in the many early-stage companies in the cluster, adding value to each in turn.

Clusters are recognised as an ideal mechanism to grow many emerging industries. Australia has a small number of successful clusters, probably the best well known being the photonics centre in Sydney. The experiences of the photonics centre would be valuable in the formation of a bioinformatics cluster.

Regional market access can be captured through regional bioinformatics industry leadership and developing trade and investment links. This could be achieved at a small cost (eg. a secretariat), and could be a role of the COBE.

To monitor and represent bioinformatics industry needs and developments on an ongoing basis, a mechanism needs to be put in place, preferably through an industry representative body. The nascent Australian Association for Bioinformatics (AAB) would be an ideal body to take on such a task.

# References

## References

Batterham R. (2000), The Chance to Change - a discussion paper by the Chief Scientist, ISR, Canberra.

Bioinformatics in Québec: a cornerstone of bioindustry development. January 2001.

Brotchie, J. et. al. (2001), The Impact of High Performance Computing Technologies on Australian Industry, ISR, Canberra.

Commonwealth Government of Australia et.al. (2001), Australian Biotechnology Report, Canberra.

Commonwealth Government of Australia (2001), Enabling the Virtuous Cycle - Report of the Implementation Committee for the Health and Medical Research Strategic Review, Canberra.

Deloitte, Touche Tohmatsu (2001), Deloitte Biotech Index - A Review of Key Biotechnology Stocks Listed on the Australian Stock Exchange, Sydney.

Frontline Strategic Management Consulting Inc (2000) Bioinformatics: A Strategic Business Analysis Report.

Gilbert, Walter. Towards a paradigm shift in biology. Nature, 349:99.

House of Representatives Standing Committee on Primary Industries and Regional Services (2001), Bioprospecting: Discoveries Changing the Future, The Parliament of the Commonwealth of Australia, Canberra.

ISR (2001), Backing Australia's Ability: An Innovation Action Plan For The Future, Canberra.

ISR (2001), Biotechnology Centre of Excellence - Discussion Paper, Canberra.

ISR (2001), Pharmaceutical Industry Action Agenda - Discussion Paper, Canberra.

Littlejohn, T.G. (2000) Bioinformatics in Australia, Bioinformatics 16: 849-850.

# Appendix A

## Appendix A

### Bioinformatics synonyms and definitions

In 1996 the OECD convened a megascience forum on 'biological informatics' to look at areas of 'neuroinformatics' (computer based information on neurological and brain biology) and 'biodiversity informatics' (computer based catalogues of biodiversity information). The OECD made the decision to omit molecular biology (DNA and protein) informatics as a focus as they believed it was sufficiently mature an area to not require attention. They felt the need to define the activity as 'biological informatics' as they believed that 'bioinformatics' was narrowly thought of as being molecular biology specific.

The professional society that best represents the bioinformatics research community is known as the International Society for Computational Biology (ISCB- see <http://www.iscb.org>). The ISCB "is dedicated to advancing the scientific understanding of living systems through computation" which is a broad definition but is then qualified with "our emphasis is on the role of computing and informatics in advancing molecular biology".

The official journal of the ISCB is called Bioinformatics. Published by Oxford University Press, (<http://bioinformatics.oupjournals.org>) the journal aims to publish scientific papers in the fields of 'computational molecular biology, biological databases and genome bioinformatics'. The journal changed its name from "Computer Applications in the Biosciences" in 1997 to reflect the growing and broader use of the term.

The bioinformatics community meets at conferences such as the Intelligent Systems for Molecular Biology (ISMB - <http://ismbo1.cbs.dtu.dk>), the Pacific Symposium of Biocomputing (PSB- <http://psb.stanford.edu/>) and Recomb (<http://recomb2001.gmd.de>), a conference on 'Computational Molecular Biology'.

Research groups around Australia have adopted an even wider set of terms: the national bioinformatics facility ANGIS stands for the Australian National Genomic Information Service ([www.angis.org.au](http://www.angis.org.au)), and the bioinformatics research group at the ANU goes by the name CBIS, the Centre for Biological Information Sciences.

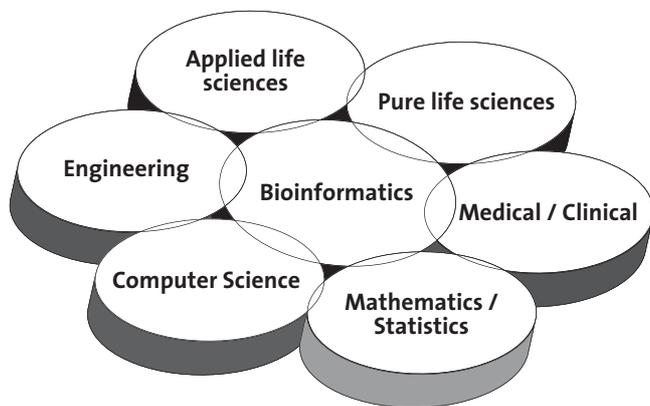
Additionally the term has grown to include even wider concepts of computer-based modelling of living processes and the term 'in silico biology' has also emerged with conferences (eg. <http://www.healthtech.com/conference/ooisb/index.htm>) and journals (eg. <http://www.bioinfo.de/isb/>) using this terminology.

For the purpose of this document the following are considered to be synonyms for, or fully included by, bioinformatics:

- Biological informatics
- Computational biology
- Biocomputing
- In silico biology
- Information biology

This confusion in terminology for bioinformatics has made defining the industry difficult, including identifying ways it underpins other industries and can itself grow.

**Figure 9. Interdisciplinary nature of bioinformatics**



Bioinformatics is most commonly defined as the use of computing systems to store, manage, and analyse biological information as it applies to molecular biology information, although this definition is rapidly expanding to include other information (eg. biodiversity). It is considered to draw on a variety of computing technologies including databases, algorithms, networks, high performance computing and domain knowledge in biology (gene and protein structure

and function, metabolism, physiology, population biology, evolution, biodiversity) by recruiting approaches in mathematics, statistics, engineering and computer science to map and model life's code: genomes and the information they carry. There would be no human genome project without bioinformatics - bioinformatics is the means by which the three billion bits of human genome data each of us carry can be understood.

Bioinformatics has most commonly been used to describe the storage and analysis of genomic information (DNA, genes, proteins). This reflects the history of the field and the recent focus on genome projects and their relevance to biotechnology and pharmaceutical industries. Indeed, the amount of data generated by the human and other genome projects is enormous, and doubles approximately every year (Genbank- [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). The Internet has become the major method of access to this data for life scientists, and it also allows the biologist to gain access to the substantial computing power required to analyse this information.

Bioinformatics has grown as a discipline and industry due to the large amounts of data that have been produced by genomics and functional genomics programs. Revolutions in data production instrument technologies have allowed biologists to produce vast quantities of life science data in very short times at increasingly lower costs. Improvements in DNA sequencing (driven by companies such as Perkin Elmer) have underpinned genome projects. Functional genomics instruments are producing even larger and more complex datasets: DNA array data is produced by companies such as Arrymetrix, proteomics data is produced by firms such as MDS Proteomics, and genome variation data produced by firms such as Sequenom. Bioinformatics is all about converting data to information and information to knowledge, using computing systems that implement sophisticated mathematical and statistical methods into algorithms for efficient mining of biological data.

In an analysis of the bioinformatics industry, Frontline Research identified several major bioinformatics technologies that help define the traditional view of bioinformatics:

- Bioinformatics Databases
- Genomic Data Analysis
- Proteomic Data Analysis
- Protein 3D Structure Analysis

- Clinical and Pharmacological Data Analysis (pharmacogenomics)
- Internet-based Bioinformatics Technologies and Business-to-business Developments

However, many see this definition as being too restrictive, as discussed below.

### A broader definition of bioinformatics and its industry

In its broadest definition, bioinformatics is the application of information technologies and sciences to the organisation, management, mining and use of life-science information.

The narrow (and typically undisputed) definition of bioinformatics is its application to the processing of molecular biology (DNA and protein) datasets. A broader definition of bioinformatics is required due to the very nature of biology and biological processes. Ultimately, biology is all about the study of living systems, from biosphere to population to genome to gene to molecules.

It is often difficult to look at one level of living systems without taking into account the interactions these levels have with each other: mutations in genes can, for instance, result in altered proteins and changed gene expressions which affect metabolic pathways, development and can ultimately result in disease. The frequency of these mutations in populations creates population genetics information, and the changes in gene frequencies information is used in the analysis of evolution. All this information may be applied to practical outcomes in biotechnology, pharmaceutical, health and agricultural research and product development - industries where ultimately the mode and mechanism of biomolecule functions are exploited.

Because of the complexity of living systems and the richness of biological data, bioinformatics is a separate industry in its own right. It does not only serve molecular biology related industries (eg. molecular biotechnology and pharmaceutical industries) but, through common and integrated information sources, also serves a range of other activities such as environmental monitoring and biodiversity assessments.

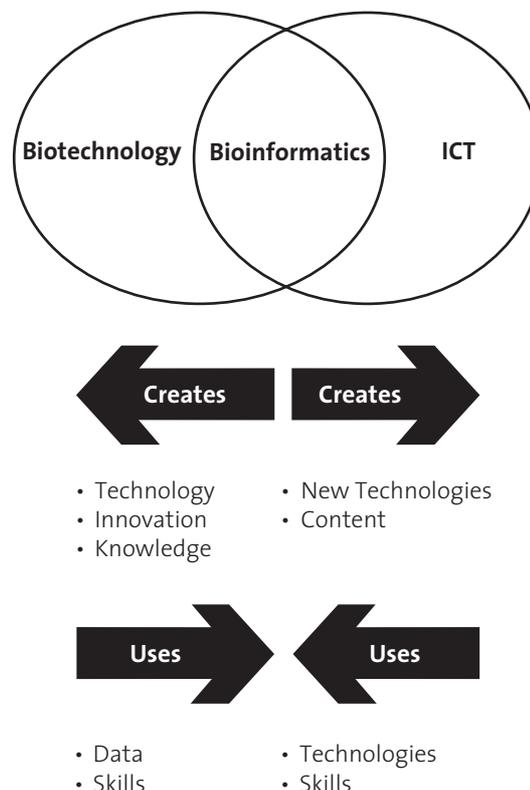
**Given the broad spectrum of life science issues to which it relates, the bioinformatics industry is defined as follows:**

**"Companies that sell products or services that are either digitised<sup>2</sup> life science information or used by their customers to generate digital life-science information."**

This definition includes companies that make and sell (i) tools (computers, software, database technologies) used to manage, analyse and visualise life-science information; (ii) data and derived information products (eg. genome databanks, protein structure repositories, human genetic variation collections); and (iii) services (Application Service Providers, custom tool development, databank creation, training & education).

This definition distinguishes bioinformatics very clearly from biotechnology. The biotechnology industry has products or services that are organic molecules (eg. therapeutics) or non-computing instruments (eg. medical devices, data generating instruments). By this definition some companies that have information as their final product (eg. genetic testing companies) could be reclassified as bioinformatics companies.

**Figure 10. Relationship of bioinformatics with biotechnology and ICT industries**



<sup>2</sup> "Digitised" and "digital" means computer based, network transmissible data

The bioinformatics industry is also clearly delineated from the ICT industry (see Figure 10) although it overlaps with this industry in that it is a consumer of ICT technologies (eg. computers), an application of ICT, and a producer of ICT (novel applications, content and hardware). Indeed, some traditional ICT companies (eg. IBM) have active life science groups. Due to the large and growing market opportunity, it is not surprising that companies like IBM are investing heavily in capturing a slice of this market. For example, IBM has budgeted \$USD 200 million to gaining the business of companies working with genome data. Other computer giants such as HP/Compaq are investing similar amounts, and firms such as Sun Microsystems have heavy commitments to growing their market share in the industry. However, in the narrow definition of the bioinformatics industry, these companies are principally computer hardware or general service providers and so should not be considered bioinformatics companies. Nevertheless, they are significant players in the industry landscape.

When companies have a portfolio of services and products that may include bioinformatics, unless this is the core business of the company, the company should not be considered part of the bioinformatics industry.

#### Related informatics areas

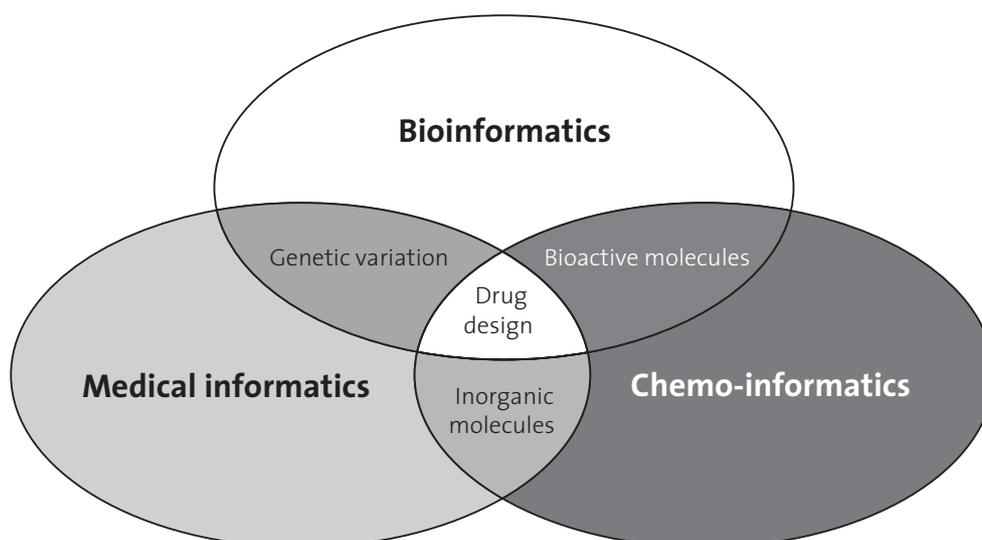
Related information-science based industries include the medical informatics and chemoinformatics industries. These industries are quite separate from bioinformatics in a number of distinct ways.

Clinical or medical informatics is focused on all areas of the application of ICT to the delivery of health outcomes. This can be as simple as the collection and storage of clinical records through to high-bandwidth telemedicine and other distance health delivery technologies. Medical informatics might be seen as overlapping with bioinformatics in that it supplies critical data used for the design of new therapeutics based on genomic information.

Chemoinformatics is another established industry: it is principally concerned with the simulation and modelling of chemical structures and reactions on computers. This industry also overlaps with bioinformatics in, for example, the modelling of small molecule (ligand) interactions with proteins in the area of bioinformatics focused on protein structure and function.

All three areas interact in complex process such as the computer aided design of new drugs based on computerised information of gene products (proteins which are often the 'targets' for drugs), the atomic-level chemical interactions of these small and large molecules, and the clinical relevance of such interactions, as shown in Figure 11 below.

**Figure 11. Intersection of bioinformatics with medical informatics and chemoinformatics.**



# Appendix B

## Appendix B

### Stakeholders who would use/need bioinformatics

AusBiotech

<http://www.ausbiotech.org/>

Australian Pharmaceutical Manufacturers Association

<http://www.apma.com.au/>

Cooperative Research Centres

<http://www.isr.gov.au/crc/centres/index.html>

Dairy Research Development Corporation

<http://www.drdc.com.au/>

Grains Research and Development Corporation

<http://www.grdc.com.au/>

Rural Industry Development Corporation

<http://www.rirdc.gov.au/>

Australian Information Industry Association

<http://www.aiia.com.au/>

Australian Medical Association

<http://www.ama.com.au/>

Genetics Society of Australia

<http://gsa.angis.org.au/>

Australian Society for Biochemistry and Molecular Biology

<http://www.asbmb.org.au/>

The Australian Society for Medical Research

<http://www.asmr.org.au/>