SEVENTH FRAMEWORK PROGRAMME

Research Infrastructures

INFRA-2008-1.2.2 – Scientific Data Infrastructures

Grant agreement for: Combination of Collaborative projects & Coordination and support actions (CPCSA)

Annex I - "Description of Work"

Project acronym: **4D4Life**
Project full title: **Distributed Dynamic Diversity Databases for Life**
Grant agreement no.: **238988**

Date of preparation of Annex I: 4 March 2009
Date of approval of Annex I by Commission: 7 July 2009

List of Participants

Beneficiary Num.	Beneficiary name	Benef. short name	Country	Enter project	Exit project
Main Members of the Work Packages					
1 (coord)	*(Coordinator)* The University of Reading	UR	UK	Mth 1	Mth 36
2	Botanic Gardens Conservation International	BGCI	UK	Mth 1	Mth 36
3	Museum National d'Histoire Naturelle	MNHN	FR	Mth 1	Mth 36
4	Narodni Muzeum	NARODNI	CZ	Mth 1	Mth 36
5	Stichting Expertisecentrum voor Taxonomische Identificatie	ETI	NL	Mth 1	Mth 36
6	Cardiff University	CU	UK	Mth 1	Mth 36
7	Stichting National Natuurhistorisch Museum Naturalis	NNM	NL	Mth 1	Mth 36
8	Species 2000	SP2000	UK	Mth 1	Mth 36
9	Vlaams Instituut Voor De Zee Vzw	VLIZ	BE	Mth 1	Mth 36
Members of the GSD Network (WP3)					
10	Museum für Naturkunde - Leibniz Institut für Evolutions – und Biodiversitätsforschung an der Humboldt-Universität zu Berlin	MfN Berlin	DE	Mth 1	Mth 36
11	Natural History Museum	NHM	UK	Mth 1	Mth 36
12	CAB International	CABI	UK	Mth 1	Mth 36
13	Royal Botanic Gardens Kew	RBGK	UK	Mth 1	Mth 36
14	Universitaet Wien	UNIVIE	AT	Mth 1	Mth 36
15	Universiteit Utrecht	UU	NL	Mth 1	Mth 36
16	National University of Ireland, Galway	NUI Galway	IE	Mth 1	Mth 36
17	Leibniz-Institut für Meereswissenschaften an der Universitaet Kiel	IFM-GEOMAR	DE	Mth 1	Mth 36
18	Deutsches Krebsforschungszentrum	DKFZ	DE	Mth 1	Mth 36
19	Museum and Institute of Zoology – Polish Academy of Sciences	MIZ-PAS	PL	Mth 1	Mth 36
20	Land Oberösterreich	Land OOE	AT	Mth 1	Mth 36

21	National Museum Wales	NMGW	UK	Mth 1	Mth 36
22	Institut de Recherche pour le developpement	IRD	FR	Mth 1	Mth 36
23	Agencia Estatal, Consejo Superior de Investigaciones cientificas	CSIC	ES	Mth 1	Mth 36
24	TSJ BVBA	TSJ BVBA	BE	Mth 1	Mth 36
25	Bayerische Staatsministerium fur Wissenschaft, Forschung und Kunst	SNSB, Munich	DE	Mth 1	Mth 36
26	Universita Degli Studi Di Padova	UNIPD	IT	Mth 1	Mth 36
27	Universiteit Van Amsterdam	UVA	NL	Mth 1	Mth 36
28	Institut Royal Des Sciences Naturelles De Belgique	RBINS	BE	Mth 1	Mth 36
29	The Chancellor, Masters and Scholars of The University of Oxford	University of Oxford	UK	Mth 1	Mth 36
30	Wageningen Universiteit	WU	NL	Mth 1	Mth 36
31	Royal Botanic Garden, Edinburgh	RBGE	UK	Mth 1	Mth 36
32	Københavns Universitet	UCPH	DK	Mth 1	Mth 36
33	International Trust for Zoological Nomenclature	ITZN	UK	Mth 1	Mth 36
Members of the Multi-Hub Network (WP4)					
34	Chinese Academy of Sciences,	CAS	CN	Mth 1	Mth 36
35	Centro de Referência em Informação Ambiental	CRIA	BR	Mth 1	Mth 36
36	Smithsonian Institution, National Museum of Natural History	National Museum of Natural History	US	Mth 1	Mth 36
37	Commonwealth Scientific and Industrial Research Organisation	CSIRO	AU	Mth 1	Mth 36
38	Landcare Research New Zealand Limited	Landcare Research	NZ	Mth 1	Mth 36

(empty page)

Table of Contents

PART A

A.I. Budget breakdown and project summary

A.I.1 Overall budget breakdown for the project

Participant number in this project [9]	Participant short name	Estimated eligible costs (whole duration of the project)						Total receipts	Requested EC contribution
		RTD (A)	Coordination (B)	Support (C)	Management (D)	Other (E)	Total A+B+C+D+E		
1	UR	0.00	60,160.00	0.00	161,622.00	893,512.00	1,115,294.00	0.00	1,095,366.00
2	BGCI	0.00	179,051.00	0.00	0.00	0.00	179,051.00	0.00	179,051.00
3	MNHN	0.00	325,120.00	0.00	0.00	16,000.00	341,120.00	0.00	241,109.00
4	NARODNI	0.00	129,440.00	0.00	0.00	0.00	129,440.00	0.00	86,563.00
5	ETI	0.00	0.00	0.00	0.00	703,218.00	703,218.00	0.00	703,218.00
6	CU	370,612.00	0.00	0.00	0.00	0.00	370,612.00	0.00	277,959.00
7	NNM	0.00	22,560.00	0.00	0.00	0.00	22,560.00	0.00	20,116.00
8	SP2000	0.00	0.00	0.00	0.00	85,120.00	85,120.00	0.00	85,120.00
9	VLIZ	0.00	22,560.00	0.00	0.00	18,000.00	40,560.00	0.00	38,116.00
10	MfN Berlin	0.00	22,560.00	0.00	0.00	0.00	22,560.00	0.00	20,116.00
11	NHM	0.00	47,360.00	0.00	0.00	0.00	47,360.00	0.00	31,672.00
12	CABI	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
13	RBGK	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
14	UNIVIE	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
15	UU	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
16	NUI Galway	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
17	IFM-GEOMAR	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
18	DKFZ	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00

Participant number in this project [9]	Participant short name	Estimated eligible costs (whole duration of the project)						Total receipts	Requested EC contribution
		RTD (A)	Coordination (B)	Support (C)	Management (D)	Other (E)	Total A+B+C+D+E		
19	MIZ-PAS	0.00	22,560.00	0.00	0.00	0.00	22,560.00	0.00	20,116.00
20	Land OOE	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
21	NMGW	0.00	22,560.00	0.00	0.00	0.00	22,560.00	0.00	20,116.00
22	IRD	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
23	CSIC	0.00	40,232.00	0.00	0.00	0.00	40,232.00	0.00	40,232.00
24	TSJ BVBA	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
25	SNSB, MUNICH	0.00	40,232.00	0.00	0.00	0.00	40,232.00	0.00	40,232.00
26	UNIPD	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
27	UVA	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
28	RBINS	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
29	University of Oxford	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
30	WU	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
31	RBGE	0.00	20,116.00	0.00	0.00	0.00	20,116.00	0.00	20,116.00
32	UCPH	0.00	30,080.00	0.00	0.00	0.00	30,080.00	0.00	20,116.00
33	ITZN	0.00	17,760.00	0.00	0.00	0.00	17,760.00	0.00	15,836.00
34	CAS	0.00	8,800.00	0.00	0.00	0.00	8,800.00	0.00	5,885.00
35	CRIA	0.00	6,600.00	0.00	0.00	31,320.00	37,920.00	0.00	37,205.00
36	National Museum of Natural History	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
37	CSIRO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	Landcare Research	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TOTAL		370,612.00	1,409,167.00	0.00	161,622.00	1,747,170.00	3,688,571.00	0.00	3,300,000.00

A.1.2 Project summary

A coherent classification and species checklist of the world's plants, animals, fungi and microbes is fundamental for accessing information about biodiversity. The Catalogue of Life provides the world with a unique service: a dynamically updated global index of validated scientific names, synonyms and common names integrated within a single taxonomic hierarchy.

The Catalogue of Life was initiated as a European Scientific Infrastructure under FP5 and has a distributed knowledge architecture. Its federated e-compendium of the world's organisms grows rapidly (now covering well over one million species), and has established a formidable user base, including major global biodiversity portals as well as national biodiversity resources and individual users worldwide.

Joint Research Activities in this 4D4Life Project will establish the Catalogue of Life as a state of the art e-science facility based on an enhanced service-based distributed architecture. This will make it available for integration into analytical and synthetic distributed networks such as those developing in conservation, climate change, invasive species, molecular biodiversity and regulatory domains. User-driven enhancements in the presentation of distribution data and bio-data will be made.

In its Networking Activities 4D4Life will strengthen the development of Global Species Databases that provide the core of the service, and extend the geographical reach of the programme beyond Europe by realizing a Multi-Hub Network integrating data from China, New Zealand, Australia, N. America and Brazil.

Service Activities, the largest part of 4D4Life, will create new electronic taxonomy services, including synonymy server, taxon name-change, and download services, plus new educational and popular services, for instance for hand-held devices.

A.1.3 List of beneficiaries

- The University of Reading, (the "coordinator"), ("beneficiary n° 1"),
- Botanic Gardens Conservation International ("beneficiary n° 2"),
- Museum National d'Histoire Naturelle ("beneficiary n° 3"),
- Narodni Muzeum ("beneficiary n° 4"),
- Stichting Expertisecentrum voor Taxonomische Identificaties ("beneficiary n° 5"),
- Cardiff University ("beneficiary n° 6"),
- Stichting Nationaal Natuurhistorisch Museum Naturalis ("beneficiary n° 7"),
- Species 2000 ("beneficiary n° 8"),
- Vlaams Instituut Voor De Zee Vzw ("beneficiary n° 9"),

- Museum fur Naturkunde - Leibniz-Institut fur Evolutions- und Bioversitatsforschung an der Humboldt-Universitat zu Berlin ("beneficiary n° 10"),
- Natural History Museum ("beneficiary n° 11"),
- CAB International ("beneficiary n° 12"),
- Royal Botanic Gardens Kew ("beneficiary n° 13"),
- Universitaet Wien ("beneficiary n° 14"),
- Universiteit Utrecht ("beneficiary n° 15"),
- National University of Ireland, Galway ("beneficiary n° 16"),
- Leibniz-Institut fuer Meereswissenschaften an der Universitaet Kiel ("beneficiary n° 17"),
- Deutsches Krebsforschungszentrum ("beneficiary n° 18"),
- Museum and Institute of Zoology - Polish Academy of Sciences ("beneficiary n° 19"),
- Land Oberösterreich ("beneficiary n° 20"),
- National Museum Wales ("beneficiary n° 21"),
- Institut de Recherche pour le developpement ("beneficiary n° 22"),
- Agencia Estatal Consejo Superior de Investigaciones Cientificas ("beneficiary n° 23"),
- TSJ BVBA ("beneficiary n° 24"),
- Bayerische Staatsministerium fur Wissenschaft, Forschung und Kunst ("beneficiary n° 25"),
- Universita Degli Studi Di Padova ("beneficiary n° 26"),
- Universiteit Van Amsterdam ("beneficiary n° 27"),
- Institut Royal Des Sciences Naturelles De Belgique ("beneficiary n° 28"),
- The Chancellor, Masters and Scholars of The University of Oxford ("beneficiary n° 29"),
- Wageningen Universiteit ("beneficiary n° 30"),
- Royal Botanic Garden Edinburgh ("beneficiary n° 31"),
- Københavns Universitet ("beneficiary n° 32"),
- International Trust for Zoological Nomenclature ("beneficiary n° 33"),
- Chinese Academy of Science ("beneficiary n° 34"),
- Centro de Referencia em Informacao Ambiental ("beneficiary n° 35"),
- Smithsonian Institution National Museum of Natural History ("beneficiary n° 36"),
- Commonwealth Scientific and Industrial Research Organisation ("beneficiary n° 37"),
- Landcare Research New Zealand Limited ("beneficiary n° 38"),

PART B

Section B1: Scientific and technical quality

B.1.1 Concept and project objectives

The General Problem being addressed

Despite 250 years of effort in the taxonomic profession, there is still, in 2008, no complete catalogue of all presently known animals, plants, fungi and micro-organisms of the world. This is a critical problem for the scientific community, and for national, regional and global organisations that organise and regulate the exchange of biotic information and materials worldwide. The set of organisms known to science is a key dimension of human knowledge concerning global biodiversity, evolution, ecology, natural resources, and biotic response to climate change. It supplies a vital set of index terms needed to access most biodiversity knowledge. There is increasing public need and expectation, focussed through the UN Convention on Biological Diversity (CBD), to complete such a catalogue of all known organisms for international uses. Many commentators are surprised that a complete catalogue does not already exist. In fact it is a non-trivial task that is too large for the individual capabilities of even the largest taxonomic institutions, due to the distributed nature of the knowledge.

This project was initiated by the Species 2000 consortium that has made substantial progress with resolving this problem. It has created, maintained and enlarged the Species 2000 Catalogue of Life to the point where it now covers 1.1 million species of plants, animals, fungi and micro-organisms, some 60% of the anticipated total of 1.8 million presently known species worldwide. It has done this by employing a radical architecture of federating global sectors of taxonomic expert knowledge from a growing array of supplier databases, and integrating these into a single taxonomic hierarchy and species checklist. The distributed system harvests taxonomic knowledge provided and maintained by a community of supplier organisations in the taxonomic profession, combining work by the major taxonomic institutions with that of smaller networks and individuals. This process was brought to production scale by the EC EuroCat project funded as a scientific infrastructure under FP 5 (2003 – 2006) and further developed since then with funding from other sources, including the EC EDIT Network of Excellence in FP 6.

Over the last two years the programme has concentrated on extending and improving the scientific content of the Catalogue of Life, which is now a unique and scientifically valuable resource. However, it has come as a bonus to see the rising and now substantial public usage in Europe and all over the world, including by GBIF and the Encyclopedia of Life, of what is presently an incomplete service. This project in the Capacities Programme provides a timely opportunity to develop a parallel focus on services. It will enable us to enrich the variety and technical sophistication of taxonomic services that are undoubtedly possible, exploiting the taxonomic resource that we are already building. The utility of these services will secure the sustainability of the whole programme into the future.

The Present Concept

The Species 2000 Catalogue of Life (henceforward ‘the Catalogue’) has a single purpose, to enable users throughout biological and biodiversity sciences, and across the many scientific

and non-scientific disciplines that use organism information, to access data about all organisms by means of a species checklist and a taxonomic hierarchy. It is already used to access data such as organism relationships, ecology, DNA sequences, protection status, invasive properties or information in any one of a myriad of other data domains. Such a Catalogue needs to be:

i) comprehensive: covering all known organisms in all groups;
ii) global: organisms of the whole world, in terrestrial, freshwater and marine environments;
iii) validated: a responsible, modern and professional globalised taxonomic view of the classification, supported by and embedded in the profession's activities;
iv) accurate: reflecting as accurately as is practical the detail of diversity of living organisms;
v) accessible to all: a clear view of the taxonomy, eventually in multi-lingual presentation;
vi) available to all: widely and freely available in a variety of forms; and
vii) dynamic: updated for taxonomic changes though time, either continuously or annually.

To be effective in the many applications in which it is used, the classification and the naming of species and higher taxa must be as close to 'agreed and correct' as is possible in taxonomy. This means for each taxon either using a consensus system, or selecting and using consistently one of the competing classifications where alternatives are in wide use. Because alternative classifications have been used both today and in the past, users must be able to locate species known by other names (or concepts) in the Catalogue, and discover alternative names under which to access data on the internet or in other resources. Consequently synonymy and common names must be included for each species. As much as possible should be 'concept-based', a precision provided by some of the databases.

The dream is simple - to create a Catalogue that contains an accurately maintained synonymic species checklist covering all known species, connected in a validated taxonomic hierarchy.

The Existing Programme

The present Catalogue of Life Programme, led by the global Species 2000 organisation based in Reading, and working with the N. American organisation ITIS (part of the Smithsonian Institution, National Museum of Natural History) was set up as an international programme at a UK-funded (BBSRC) workshop in 2001. Bringing the programme up to production scale was funded by the EC as one of its scientific infrastructures (EuroCat), with further funding by the Japanese Government, the US Government (through ITIS) and GBIF. Output is via the *Catalogue of Life Annual Checklists* on CD [2], and on the web [3], and the *Dynamic Checklist* on the web [4], both also available as web-services for electronic use.

In March 2007 an EC-funded 'Million Species Day' symposium was held to celebrate reaching one million species. The 2008 *Annual Checklist* now provides a quality species checklist of 1,105,589 species with unique identifiers and a hierarchy for all organisms (animals, plants, fungi, chromista, protozoa, bacteria, archaea, viruses). The estimate for the number of known extant species is currently 1.75 – 1.8 million [1]. The present Catalogue benefits from simplicity of structure incorporating minimal but standardised data for each species. These contribute to its success in providing a universal baseline needed by all biologists, and in making the project practicable. It consists of two knowledge structures, and software that enables the user to search or traverse them, and to toggle between them. i) The Species Checklist is a series of Species pages (Figure 1) that are located by name searches, with automated synonymic indexing. Each page gives the Standard data for a Species, including common names, the higher taxa it belongs to in the hierarchy, and geographical distribution. ii) The Taxonomic Hierarchy (Figure 2) is an expansible tree that can be

followed down through the classification to the 1.1 million individual species. Or it can be used to navigate upwards to the higher taxa containing the one that is viewed. By clicking on a higher taxon listed on a species page, the user can transfer to the tree for that taxon, and see all its daughters. Conversely, by clicking on a species at a twig in the tree, the user can visit the relevant Species page in the Checklist.

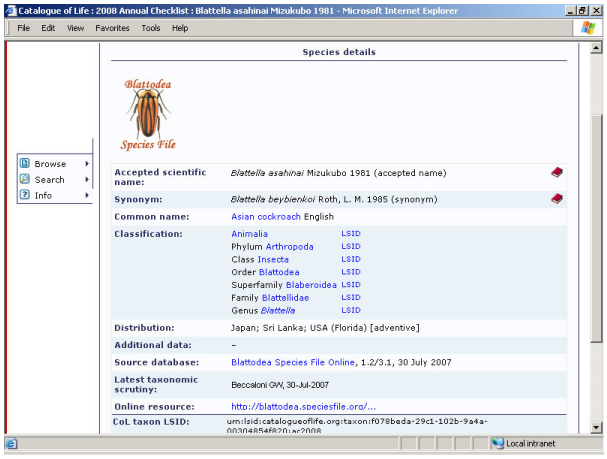


Figure 1. Species page in the Catalogue

A comprehensive checklist cannot be made simply by adding together regional or single-country lists. Different classification and naming schemes mean that a simple additive list would be massively duplicative and of little use. The current system is a successful development of the original BBSRC SPICE project. It federates the sector checklists provided by an array of distributed databases (GSDs), which are globalised checklists of a whole taxon, harvested across the Internet, and fitted together 'end-to-end' within a single overall classification. When enough sectors are fitted this process can eventually create a complete list. The number of GSDs contributing one or more taxonomic sectors to the Catalogue reached 52 for the 2008 edition, including 37 based in Europe, 11 in the USA, and 4 in Russia, Japan, Australia and the Philippines and approaches 66 for the 2009 edition. The model ensures that sectors are enhanced taxonomically by the supplier databases, and ca. 3,000 experts globally contribute to these databases.

Each GSD sector is attached at its 'top point' (its highest ranking taxon) in the hierarchy, and in addition to harvesting the checklist, the system also harvests branches of the tree beneath this top point for the hierarchy leading down to the species in that sector of checklist. The checklist and hierarchy created from a growing array of GSDs in this present-day architecture ('Architecture 1') is referred to as the 'Global Hub'.

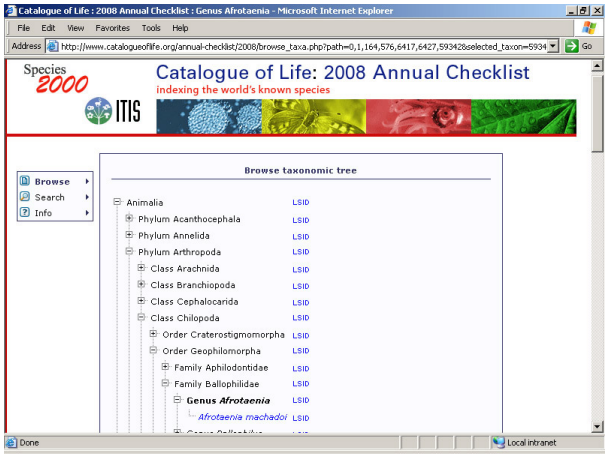


Figure 2. The Taxonomic Hierarchy

Despite the evident success of Architecture 1 in permitting the rapid build up of the Catalogue to its present point, its limitations have been evident for some time. The difficulty is simply that no-one anywhere in the world is creating global species databases for some of the least known taxonomic groups, so by this model these would be destined always to remain as gaps in the Catalogue. In the EC EuroCat project (2003 – 2006), we additionally experimented with making a Regional or 'Euro-hub' with a further set of European regional databases, and versions of SPICE that could handle multiple hubs, and the first steps towards integrating their contents using the LITCHI 2 taxonomically intelligent integrity tracking. In 2005 we started to plan an 'Architecture 2', in which an array of Regional Hubs might be connected to the Global Hub, this providing linkage to regional databases from many parts of the world, but also the

potential for the Global Hub to harvest data or checklist sectors from Regional treatments for the species groups that were missing from the Global Hub. Good progress is being made with initiating these Regional hubs now – with the test N. American Hub now ready for connection in addition to the European Hub, and plans in WP 4 of this project to develop a unified concept and specification for this Multi-Hub Network working with the designated centres for China, New Zealand, Brazil, and Australia.

References

[1] Heywood VH et al., 1995. Global Biodiversity Assessment. Cambridge University Press, 1140p.
[2] Bisby FA et al., 2008. Species 2000 & ITIS Catalogue of Life: 2008 Annual Checklist. CD-ROM; Species 2000: Reading UK.
[3] Bisby FA et al., 2008. www.catalogueoflife.org/annual-checklist/2008/
[4] Bisby FA et al., 2008. www.catalogueoflife.org/dynamic-checklist,

Who uses the present services?

Despite being only 60% complete, the Catalogue has become established as a significant global biodiversity resource with extensive electronic and human usage in Europe and all around the world. There is no other comparable source of this information.

1. On the web. Our services operating on the web [www.catalogueoflife.org; www.sp2000.org] receive ca. 40 M hits per annum, with a rising number of 30 - 40,000 unique serious users per month. But these account for only a fraction of the total usage on the web, which is by the users of other biodiversity portals that have installed a copy of the Catalogue as their taxonomic index.

2. Globally significant biodiversity portals and biodiversity programmes. Two globally significant biodiversity portals have installed the Catalogue as their principal taxonomic index, using both the taxonomic hierarchy and the synonymic species checklist to enable users to locate the organisms for which they seek information. The Global Biodiversity Information Facility (GBIF) has used it since the launch of its portal in 2003, and the Encyclopedia of Life (EoL) since its launch in 2008. Nearly every scientific use of those systems to locate data about an organism (and far outnumbering the 40 M hits on our own sites) makes use of the Catalogue one to many times. Neither organisation could have launched their portal in the way that it did without our Catalogue. Similar usage is made of the Catalogue by other portals (SeaLifeBase, SpeciesBase, NSF SEEK, the OpenModeller library used for climate change studies, uBio and the BHL taxonomically intelligent tool), and related usage as an authority file by GenBank and the Red Lists of IUCN.

The Catalogue is used by BGCI (which organises information on plants conserved in botanic gardens around the world), by the Institute of Zoology in London (which works on conservation status analyses using national Red Lists and the Sampled Red List Index), and by European partners in the BHL project. Also the European Bioinformatics Institute in Cambridge has been exploring the possibility of mapping the GenBank Taxonomy to our Catalogue so that both can be used in GenBank. This has also been promoted by iBOL, the initiative by Genome Canada, and the Consortium for the Barcode of Life (CBOL) (which has adopted our Catalogue as its 'gold standard' for taxonomic annotation of DNA Barcode records that are being deposited in EBI and GenBank).

3. National biodiversity portals and programmes. About 40 national-level portals around the world have installed the Taxonomic Hierarchy and/or the Checklist: some are GBIF nodes (eg

Netherlands), some national biodiversity portals (Norway, Thailand), some institute based (CRIA SpeciesLink, Brazil, Plants-of-India, Lucknow), and some NGO based.
4. Regional biodiversity portals and programmes. The Inter-American Biodiversity Information Network (IABIN) uses the Catalogue for its invasive species network and Species 2000 is involved with European initiatives - the European Distributed Institute of Taxonomy (EDIT), the Consortium of European Taxonomic Facilities, and the LifeWatch infrastructure programme.
5. Individual and small institute use of the Annual Checklist on CD. 3,500 copies of the Annual Checklist are distributed each year free of charge on CDs, enabling operation on PCs or laptops without internet access. During 2007 the Secretariat received written and emailed requests for the CD from 471 individuals in 79 countries. 300 are distributed by the National Institute for Environmental Sciences, Tsukuba, to members of the Japanese Society of Systematic Biologists. 300 are distributed by the Belgian Biodiversity Platform to focal points of the GTI and the Clearing House Mechanism from African and Lomé Convention countries. Distribution to the regional organisations of the BioNET International programme is about to start.
6.First commercial uses. A trial contract with Taylor & Francis plc provides taxonomic support to its electronic Dictionary of Natural Products. This has been started cautiously, as it involves permission from provider databases, and the possible return of some income to them as running costs.

What are the present services used for?

What makes the Catalogue unique is the breadth of coverage of organisms, the degree of validation in the knowledge set, and the wide European and global take-up. A wide variety of interesting and significant science is being done worldwide by a range of programmes using the Catalogue as a baseline infrastructure. So, while not in itself saving the world's biodiversity, taming the effects of climate change, enabling evolutionary study, or dealing with human food and disease crises, it does increasingly contribute as an infrastructure to research on all of these.

Basic Functions
1. Species Look-ups. Vast numbers of people, both professional and amateur, use the Web to look up species of organisms, i) to verify the name they have, ii) to place a taxon in the taxonomic hierarchy, and iii) to see other basic data, such as distribution, for that species. Ideally their search will inform them of an accepted name for the species if the one used for the search is in synonymy. Unlike general search engines or knowledge stores such as Google, Wikipedia or Wikispecies, our Catalogue provides a single clear consensus taxonomy; and taxonomic consistency in classification and naming within and between groups. Many appreciate that it is provided and maintained by taxonomic experts, so that it is likely to be more current and authoritative, for example in dealing with variant spellings and misapplied names. For instance, species look-ups are the main usage by users as diverse as our Taylor & Francis customers, users of a culture collection at the Riken Institute in Japan, and undergraduates at University of Reading.
2. Sizing higher taxa. Individuals use the Catalogue to get an estimate of the size of a taxonomic group. Typically they know one member species, but have no idea whether the group is extensive or small on a world basis.
3. Synonymic indexing and synonymic amplification. Many biologists may not appreciate either the key role played by synonymy, or the extent to which the same organisms may be referred to by different scientific names in different continents or even different states in

Europe. *Species checklists with automated synonymy have become the electronic organisers of biodiversity knowledge on the internet, and it is synonymy that enables the loose ends to be connected together.*

Synonymic indexing is automated in the Catalogue so that users in different countries who might search for Broad bean, or Fava Bean, or Faba bean, or *Vicia faba*, or *Faba vulgaris* will all arrive at the same species page. The page states clearly that the one species *Vicia faba* has a synonym *Faba vulgaris*, and common names, Broad Bean, Fava bean, Faba bean, all referring to the same species. The 'unification' or globalisation provided by accurate synonymic indexing contributes significantly to enabling international discussion and data exchange.

The reverse of this process is important – synonymic amplification. Given that one species may be referred to by many names, a person or machine searching the Internet, or a data set such as GBIF, will receive only a subset of data if they search on just one name. It is more effective to amplify the search with all synonyms of that species. There are many examples where this amplification of search strings yields sharply improved results, particularly spanning continents. The BBSRC BiodiversityWorld project introduced this as part of its data harvesting using the electronic web-service of the Catalogue, and it is now available in GBIF.

4. Taxonomic Backbone usage. GBIF has used the Catalogue as a backbone structure on which to make its own additions. It uses associative techniques to link additional species that are not presently in the Catalogue to the positions they are likely to belong. The 'taxonomically intelligent' tool provided by uBio to the BHL project using our hierarchy has broadly similar features. An email cited later, received as this proposal was in preparation, illustrates how the Institute of Zoology in London plans a related usage.

5. Download usage. The Annual Checklist edition on CD and the web is frequently used as a download facility to insert the data into a user's own database, or as the database to which a user will add further data. However the present format is not ideal for this purpose. A well-structured download facility is planned as one of the novel technical services to be added in this project.

6. Global standardisation. Organisations are starting to use the Catalogue to achieve compatibility with others that already use it, for instance GBIF nodes and some national portals. *The possibility of a quantum increase in the coherence of the world's biodiversity data and analyses is beginning to emerge, simply by the process of many organisations opting to use the same Catalogue. Given the extent of society's dependence on biodiversity, this alone is a significant goal.* Compare GBIF and EoL, that have strong data compatibility and both use the same maps, with GenBank and IUCN Red Lists where the checklists are compiled independently. The GBIF/EoL compatibility arises largely because they both use our Catalogue as a taxonomic backbone. This standardisation role favours the Catalogue continuing to be generated and maintained by an independent organisation as a scientific infrastructure embedded in the taxonomic profession, rather than by a sectoral organisation. Similar issues arise in our discussions with CITES, the European Species Directive, the Global Invasive Species Programme (GISP), the CBD Biosafety Clearing House (BCH) and other regulatory frameworks.

7. Spell-checker for users of other character sets. There is anecdotal suggestion that scientists whose working languages use other character sets find Latin scientific names quite difficult, for instance in the subtle differences that denote ranks (eg. Caesalpiniaceae, Caesalpinioideae, Caesalpinia). Spell-checker usage possibly contributes to the peaks of night-time usage of the web-site, and adds to interest in the programme from Russia, China, Thailand and Japan.

Compounding the basic functions
1. Portal usage. The importation of the Catalogue by GBIF, EoL and many of the other portals provides them with all of the basic species checklist-related functions in a single package. This re-emphasizes the extent to which this is a fundamental scientific knowledge infrastructure serving these other bodies.
2. Programme usage. It is for this same reason that the CBD places the Catalogue at the core of its Global Taxonomy Initiative, and lists usages for it in six of its programmes.
3. Community player. We believe that the Catalogue creates 'value' in the biodiversity community by being a platform to which databases provide, and from which users are supplied: it is a valued knowledge platform in the biodiversity community.

Contribution to higher level science and society
1. Impacts of Climate Change. A core issue in many climate change modelling projects is how to harvest sufficient comparable world-wide occurrence records to be able to generate climatic envelope or other models for each of the species in the study. These models are then used to interact with climate change scenarios, and to generate predictions for the future. A common taxonomic framework and synonymic amplification are key tools. Our Catalogue is used as part of this process by GBIF, the NSF SEEK project, the BBSRC BiodiversityWorld Project, CRIA SpeciesLink in Brazil, and the EC INCOFish Aquamaps system.
2. Genetic Resources for Food, Disease & Biosafety Regulation, and Invasive Species. In all these international areas of biodiversity science it is important that there is integrity across species records and scientific reports between countries. Our Catalogue is used by Bioversity International and the System-Wide Network for Genetic Resources of the international agriculture institutes, by the CBD Biosafety Clearing House (BCH) that monitors international biosafety issues, and by IABIN and GISP in international programmes relating to invasive species.
3. Identification of Biotic Materials in Foods, in Forensic Scenarios & in the Environment. Genome Canada is investing in excess of 150 M Canadian dollars in a global programme (iBOL) to establish usable DNA barcodes for identification of biotic materials in commercial, forensic and environmental programmes. To be effective this needs a common taxonomic backbone and synonymic integrity in their data systems. The International Consortium for the Barcode of Life (CBOL) has adopted our Catalogue as its 'gold standard' for annotation of Barcode records and we are presently discussing with GenBank and the Canadian iBOL programme how this is best implemented.
4. Conservation of Global Biodiversity. At present the family of organisations within IUCN, those linked to Conservation International and the Nature Conservancy in the Americas, and United Nations conventions use a wide variety of taxonomic catalogues, including some with known handicaps, such as taxonomic inflation of local endemics, and outdated taxon concepts enshrined in law. These difficulties provide a barrier to globalising efforts, and will not be solved rapidly. There are increasing signs that our Catalogue will be used in at least some of these contexts, for instance through the work of the Institute of Zoology and BGCI in London. The email excerpt of 24 June 2008 below (cited with permission) illustrates this.

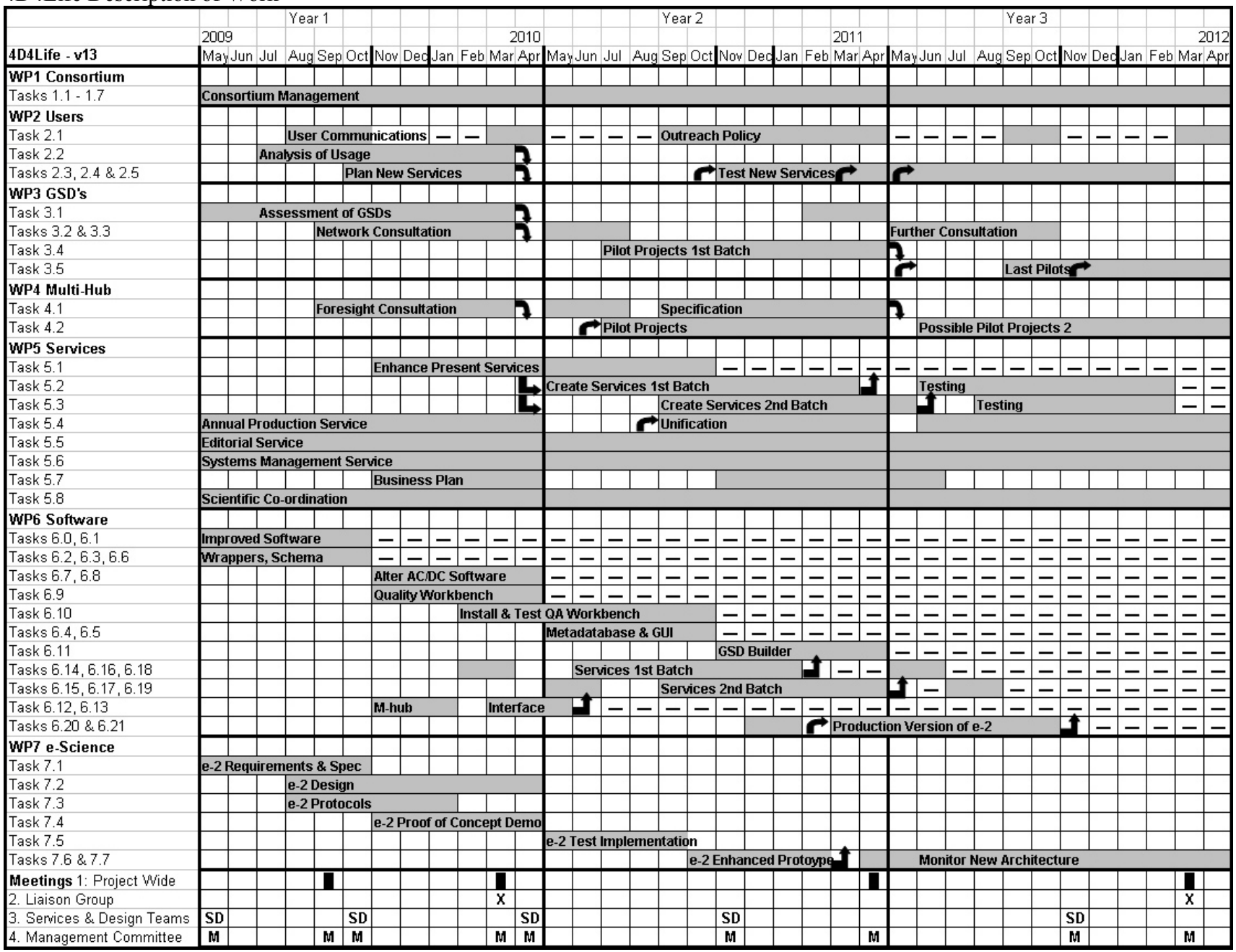
Dear Dr. Bisby,
I am writing to you to request permission to use the 2008 Annual Checklist hierarchy in my own database. At the Zoological Society of London, we are in the process of creating a master database of all National Red Lists that have been conducted worldwide. This online database will be launched at the World Conservation Congress in Barcelona this October and is intended to provide a focal point for National Red Listing. We hope that by making this data readily stored, managed and accessible, countries will be able to learn from each other in

creating and updating Red Lists, and that this will help towards effective conservation planning.

5. Uses outside Biology. Our Taylor & Francis contract is effectively for use of the Catalogue by chemists working on natural products. We are in contact with a number of wildlife image agencies, including the ARKive/WildScreen Trust in Bristol and a commercial organization in San Francisco.

Clients wish to make such uses of our Catalogue because:
i) it has a reputation for quality, because it is validated by experts and peer reviewed
ii) it has a reputation for quality, because it is internally integrated by an expert team
iii) it will cover all groups in a single system ('one-stop shopping')
iv) users can see the full extent of a group, unlike GenBank and IUCN Red Lists
v) no other organisation has aggregated the taxonomic opinions needed to provide accurate synonymy on this scale
vi) it is available electronically, and as an electronic web-service
vii) the Taxonomic Hierarchy is artificially stabilised and simplified at the higher levels as a 'management hierarchy'
viii) it has achieved a reputation as a workable standard both in Europe and worldwide.

B.1.1.1 Gantt Chart explaining how activities and workpackages are linked


Year 1
Year 2
Year 3
2009
2010
2011
2012
4D4Life - v13
May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr
WP1 Consortium
Tasks 1.1 - 1.7
Consortium Management
WP2 Users
Task 2.1
User Communications
Outreach Policy
Task 2.2
Analysis of Usage
Tasks 2.3, 2.4 & 2.5
Plan New Services
Test New Services
WP3 GSD's
Task 3.1
Assessment of GSDs
Tasks 3.2 & 3.3
Network Consultation
Further Consultation
Task 3.4
Pilot Projects 1st Batch
Task 3.5
Last Pilots
WP4 Multi-Hub
Task 4.1
Foresight Consultation
Specification
Task 4.2
Pilot Projects
Possible Pilot Projects 2
WP5 Services
Task 5.1
Enhance Present Services
Task 5.2
Create Services 1st Batch
Testing
Task 5.3
Create Services 2nd Batch
Testing
Task 5.4
Annual Production Service
Unification
Task 5.5
Editorial Service
Task 5.6
Systems Management Service
Task 5.7
Business Plan
Task 5.8
Scientific Co-ordination
WP6 Software
Tasks 6.0, 6.1
Improved Software
Tasks 6.2, 6.3, 6.6
Wrappers, Schema
Tasks 6.7, 6.8
Alter AC/DC Software
Task 6.9
Quality Workbench
Task 6.10
Install & Test QA Workbench
Tasks 6.4, 6.5
Metadatabase & GUI
Task 6.11
GSD Builder
Tasks 6.14, 6.16, 6.18
Services 1st Batch
Tasks 6.15, 6.17, 6.19
Services 2nd Batch
Task 6.12, 6.13
M-hub
Interface
Tasks 6.20 & 6.21
Production Version of e-2
WP7 e-Science
Task 7.1
e-2 Requirements & Spec
Task 7.2
e-2 Design
Task 7.3
e-2 Protocols
Task 7.4
e-2 Proof of Concept Demo
Task 7.5
e-2 Test Implementation
Tasks 7.6 & 7.7
e-2 Enhanced Protoype
Monitor New Architecture
Meetings 1: Project Wide
2. Liaison Group
X
3. Services & Design Teams
SD
4. Management Committee
M

B.1.1.2 Global View of Section B.1.3, B.1.4, B.1.5 and B.1.6

The main objectives of the 4D4Life project are to build a state-of-the-art e-infrastructure that:

- facilitates enhanced scientific information exchange within the Species 2000 networks,
- synthesises a significantly improved global resource, the *Catalogue of Life,*
- disseminates this synthesised knowledge in a new array of modern web-services and products,
- takes the e-infrastructure to a sustainable future.

Three **Networking Work packages** will enhance information exchange among the Users, GSD Network and the Global Multi-hub Network; contribute to improved data supply and content in the *Catalogue of Life*; and initiate an array of new public services.

WP 2: Engagement with scientific users and partners (led by Sara Oldfield, BGCI).
WP 3: Stengthening the GSD Partner Network: a Virtual Community (led by Thierry Bourgoin, MNHN, Paris)
WP 4: Foresight Study for geographical extension: the Global Multi-hub Network (led by Jiri Kvacek, Narodni, Prague)
(plus WP 1: Management of the Consortium)

A further two **Services Work packages** will incorporate these improvements and new services into the enhanced production and service provision process of the Catalogue of Life:

WP 5: Scientific Services of the *Catalogue of Life* (led by Frank Bisby, Univ. of Reading).
WP 6: Software Support Services (led by Peter Schalk, ETI, Amsterdam).

Finally one **Joint Research Work package** will create a new service-based Architecture that will provide the multi-layered ecosystem of data supply services to take the programme into a sustainable future.

WP 7: New Service-based Architecture (led by Richard White , Cardiff University).

B.1.1.3 Overall Deliverables List

Del. no.	Deliverable name	WP no.	Lead beneficiary	*Estimated indicative person-months*	Nature	Dissemi-nation level	Delivery date (proj. month)
1.1	Launch the 4D4Life project web-site at www.4D4Life.eu and update regularly	1	UR	1	D	PU	5
2.1	(Rolling deliverable) Six-monthly e-bulletin.	2	BGCI	8	O	PU	6, 12, 18, 24, 30, 36
2.2	User-driven proposals for new services for each class of users	2	BGCI	10	R	PP	11
3.1	Report on agreed target set of improvements to the data supply by the Global Species Databases (GSDs) (includes Revised Data Standard & Best Practice Documents)	3	MNHN	55	R	PU	11
4.1	Report on preliminary concept, requirements and specifications for the Global Multi-Hub Network: the *Preliminary Requirements Document*	4	NARODNI	20	R	PP	12
5.1	Initial software improvements and streamlining incorporated in production of the enhanced Annual Checklist 2010, published on the web, and on CD/DVD	5	UR	25	D	PU	12
6.1	Improved management interfaces and streamlined production of the Catalogue of Life delivered for existing software, (includes tasks 6.0, 6.1, 6.2, 6.3, 6.6, 6.7, 6.8, 6.9, 6.12 & 6.13)	6	ETI	39	D	PP	12
7.1	Report on Requirements, Specification and Design of new e-2 architecture completed	7	CU	14	R	PP	12

7.2	'Proof of Concept' demonstrated for the e-2 architecture	7	CU	14	D	RE	12
1.2	First Periodic Management Report of Networking, Services & Joint Research for year 1	1	UR	6	R	PP	14
7.3	Report on the implementation and testing of operative prototypes of the new architecture	7	CU	30	R	RE	24
3.2	Report on pilot projects for data provider improvements (1st batch)	3	MNHN	56	R	RE	24
4.2	Global Multi-Hub Network Design Document	4	NARODNI	20	R	PP	24
5.2	Unification of the production process and admission gate achieved for the Annual (2011) and Dynamic Checklists, published on the web, and (Annual Checklist) on CD/DVD.	5	UR	41	D	PU	24
6.2	Altered software for both unified AC & DC processes, and improved synthesis of the Catalogue of Life operational and installed, (includes tasks 6.4, 6.5, 6.10 & 6.11)	6	ETI	13	D	RE	24
6.3	Software for new custom services prototyped and ready for test, (includes tasks 6.14 – 6.19)	6	ETI	25	D	RE	24
5.3	Business Plan for sustainable development of the services	5	UR	25	R	RE	26
1.3	Report on exploitation of the foreground IPR created across the project	1	UR	5	R	RE	26
6.4	Production software for the new e-2 architecture fine-tuned and delivered, underpinning the virtual community, (includes tasks 6.20 & 6.21)	6	ETI	6	D	PU	30

2.3	Report on new services tested with users and ready for launch	2	BGCI	15	R	RE	34
5.4	New array of services available and launched with pack of publicity materials and electronic products	5	UR	55	D	PU	34
3.3	Report on pilot projects (2nd batch) including some that roll out the new architecture for data providers	3	MNHN	56	R	RE	36
1.4	Final Report on the 4D4Life enhanced scientific data e-infrastructure	1	UR	14	R	PU	36
			TOTAL	553			

B.1.1.4 Planning of Reviews

It is proposed that the Workpackage leaders hold the project reviews in Brussels at dates agreed with the Commission in months 14, 26 and 38, that is after each year, allowing for the 60 day report submission time. With a start date of 1 May 2009, these will occur in June 2010, 2011 and 2012.

B.1.2 Progress beyond the state-of-the-art

The present Species 2000 Catalogue of Life programme is a unique, significant and leading component in the landscape of international biodiversity programmes. There are a number of different levels at which it can be variously described as unique (in a class of its own), in which it is at the state of the art (along with other leading programmes), and in which it is joining with the front runners pushing 'beyond the state of the art'.

i) The programme has been unique from its start in 1996 in developing the concept of Global Species Databases (GSDs) and the principle that only by multiplying this process can a comprehensive and validated global catalogue be developed. This principle was subsequently accepted by the more recent GBIF and EoL programmes that have come to use our Catalogue as their taxonomic backbones, and by the very recent WoRMS and FADA programmes that we are working with to complete coverage of primarily marine and primarily freshwater taxonomic groups.

ii) The programme has been unique from its start in adopting a distributed model, and pushing to embed the responsibility for creating and maintaining the taxonomy of the many different taxonomic groups into the institutions and groups of experts within the taxonomic profession. It recognises that virtually the whole taxonomic profession is needed to complete and maintain the task, and that no single taxonomic institution has the capacity to do so. GSDs provide the vehicle by which expertise is focussed from the experts into the single aggregated view. The distributed model is tolerant of, and responsive to, a wide heterogeneity of purpose and style of activity across the taxonomic community. For instance it uses databases from

large institutions, medium-sized networks, and individual expert amateurs; it uses databases across a spectrum of purposes from primarily nomenclatural to full species bank, and the databases come from natural resources, veterinary, microbiological and health organisations as well as the purely taxonomic.

iii) The programme is possibly unique and certainly distinctive in its community participation model, which differs markedly from the models used by the other major global biodiversity infrastructures. Each participating centre or individual, and particularly each supplier database custodian, is invited to become a member of the Species 2000 residual legal body that owns and governs the Species 2000 programme, and the Species 2000 component of the Catalogue of Life. Each supplier database custodian is invited to sign an Access Agreement with Species 2000 that specifies and regulates how Species 2000 may publish, credit and handle expert knowledge from the supplier databases. These two features have combined to create a solid but flexible community organisation in which the partners and contributors feel a sense of ownership. It is for this reason that the European GSD Network custodians were invited to become full members of the 4D4Life Consortium and also that Species 2000, the residual legal body, is a member of the consortium, involved both in the annual production cycle, and in the moves towards sustainability.

iv) The SPICE distributed system was innovative and ahead of its time when it was first released in 2003, and was used to implement the current version of the Dynamic Checklist launched in 2005. Some of its features have since been superseded by a new generation of community standards launched by TDWG, such as TCS, but others, such as the SPICE protocol for harvesting connected checklists and taxonomic hierarchy branches have yet to be bettered in fully available protocols. It is against this background that we propose to replace the SPICE system with a fully service-based and open architecture as a main part of this proposal, and to include both significant adjustments to altered requirements, and the use of appropriate current community standards wherever possible.

v) In 2007/8 the programme introduced Globally Unique Identifiers (using the LSID system) for the Taxa recognised in the Catalogue, putting it very much at the leading edge, and preparing the ground for the range of GUID-based tracking services proposed in this project.

vi) The programme is working with major community partners GBIF and EoL towards the establishment of a new Global Names Index that will be part of a higher level Global Names Architecture. The Global Names Index (a giant list of all known names and orthographies) and the Catalogue of Life (giving a quality species list with synonymy) will be two significant items in the plans for a seamless biodiversity informatics landscape under development through this project, through EOL, GBIF, TDWG and EDIT.

B.1.3 Methodology to achieve the objectives of the project, in particular the provision of integrated services

The programme already provides a simple integrated service based on the 60% complete Catalogue. This proposal is
i) to provide a substantial enhancement and diversification of the integrated services,
ii) to re-engineer and install an enhanced state of the art service-based architecture,
iii) to strengthen the supplier base, the system infrastructure, and the sustainability,
iii) to extend community participation and taxonomic coverage to other continents.

B.1.3.1 The present process of integration, and provision of integrated services.

The Editorial and Production Team at Reading, Los Baños and ETI operates the steady state process of integrating the Catalogue using the taxonomic sectors supplied by the array of GSDs in Architecture 1. In the Annual Checklist process, there is a 12-month cycle with the Annual Checklist brought to publication and release for 1 April each year, and this is the present product of first choice to users. The Dynamic Checklist process is operated continuously, but in the proposed project we should like to increase the frequency of caching, and complete the wrapper connections for those GSDs not yet connected.

There are four processes in the integration.
i) Existing supplier GSD databases provide enhanced data sets. These contain additional species, and edited species records in the species checklist, and they may also contain alterations to the taxonomic hierarchy joining the species up the hierarchy to the top point (highest taxon) in the sector. They may also contain what are considered by the Editorial Team as additional or separated taxonomic sectors. In the Annual Checklist process the enhanced data set is downloaded, lightly scrutinised by the Editorial Team, and inserted to replace the prior data set. In the Dynamic Checklist the enhanced data set is cached automatically to overwrite that part of the cache, and only the part below the registered top point is utilised. (New sectors need to be notified and taken through the new sector process.) Note that a proportion of the yearly increase in the size of the catalogue comes from enhanced data sets generally including more species than in previous occasions – especially as some of the GSDs are incomplete, and their staff are adding more species.

ii) During each year new GSD databases are enlisted to the programme, and also some GSDs offer new sectors that they have added, extending the taxonomic coverage. These new GSDs and sectors go through an extensive admission gate procedure that may take several months. Most important is that the metadata questionnaire, extensive discussion with the custodian about cross-mapping the schemas, technical review of the database, and peer review of the taxonomic quality, all precede the incorporation of their material into the Catalogue. The initial work is done using the Annual Checklist download process, followed both by wrapper writing and connection to the Dynamic Checklist, and by administrative work of signing Access Agreements and initiating membership of the Species 2000 legal body.

iii) Underlying the seemingly straightforward admission work for a new GSD sector is an extremely skilled task for the head of the Editorial and Production Team – where and in what taxonomic rank to connect the incoming GSD sectors, and how to place them in relation to adjacent sectors, and how to check for overlaps. The Executive editor has two consultation groups available to assist with this task – the Species 2000 Taxonomy Group (that advises the global Species 2000 Team), and an appointed Editorial Panel, plus the structure provided by the upper levels of the Catalogue of Life taxonomic hierarchy adopted from time to time by the global Team. This is a contentious and potentially troublesome area, but the policy that Species 2000 should adopt a preferred position and structure for the higher level taxonomy has been a source of approval by user organisations, even though some mistakes have inevitably been made and then rectified.

iv) The annual cycle of the Annual Checklist production means that there is a phase in which the whole product is reviewed by the editorial panel, and double-checked by the provider database custodians. This is currently not available in the Dynamic Checklist process, but it is hoped to introduce this with the unification of the processes planned in this project.

B.1.3.2 Methodology to re-engineer and install the enhanced service-based architecture

Significant changes are needed because the simple Architecture 1 needs to be superseded by an architecture permitting a multi-layered ecosystem of providers and harvesters. The current distributed system needs to be enhanced with a fully service-based architecture suitable for incorporation into the programmes of other organisations such as GBIF, EoL, EBI and LifeWatch.

The methodology to be used involves the full cycle of requirements analysis, specification and design from an expert team based at the Welsh e-Science Centre at Cardiff University. A prototype system will be tested with three GSDs, and a second release produced that will be handed to WP 6 for the creation of a further robust production system with management interfaces and user documentation. Further detail is described under WPs 6 and 7 that will carry out the development.

B.1.3.3 Methodology to strengthen the supplier base

A renewed assessment and certification will be made for each GSD in the supplier array in the work plan of WP 3. This will include both the databases that joined the programme in the previous EuroCat project, and for which full certification was carried out, and the more recent recruits for which not all of the certification was completed because of shortage of resources and staff.

After the assessment each GSD will be assigned two 2-month pilot projects. In most cases at least one of these is likely to be to make and test additions/alterations to their database to increase their compliance either with the Data standard, or with the best practice guidelines developed in the EuroCat project. Although it will take time for the GSDs to start filling these additional areas, the result is expected to be a significant improvement in the uniformity and degree of fill provided by the GSD databases, and these in turn will improve the fill and quality of the integrated service content.

B.1.3.4 Methodology to strengthen the system infrastructure

Some key alterations to the currently used Annual Checklist and SPICE/Dynamic Checklist middleware are expected to make it possible to unify the Annual Checklist and Dynamic production routes. This will provide two improvements – that all tools in both processes can be used in a single production route, and that the difficult taxonomic positioning and insertion process can be partly automated by software that links between the metadatabase and the taxonomic hierarchy.

B.1.3.5 Methodology to strengthen the sustainability

Substantial work towards achieving sustainability has already been carried out. Although the programme is currently not in a sustainable position without grant support, there has been progress on several fronts, and the programme has been able to progress even through the 2006 – 2008 period without major grants.

The work needed includes discussion and financial negotiation with three of the classes of users and stakeholders – the large global programmes that depend quite heavily on the Catalogue for their own infrastructure, the many smaller portals that are in fact quite well funded at national levels, and other commercial companies that may wish to enter into small contracts such as the experimental one with Taylor & Francis. This needs to be accompanied

by legal and discussion work with the supplier databases, both to gain their support for making this activity sustainable, but also for modest contributions to their running costs to be included in the financial model.

B.1.3.6 Methodology to extend the community participation and taxonomic coverage to other continents.

The proposed development of Regional Hubs in a Multi-Hub Network under Architecture 2, and also the devolution of active work in the programme to Centres in each continent, have the technical and institutional potential to enlarge this currently Europe-based programme to become a truly global knowledge network and scientific infrastructure.
In WP 4 it is proposed to network with the leaders of these developments in the first batch of agreed regional centres – in China, New Zealand, Australia, Brazil and N. America. At the technical level the group will both evaluate the test-bed implementation available with three regional hubs (European, N. America and China) and draw up a proposed concept and specification for the working of the regional Hubs. At the institutional level it is proposed that they also draw up a model for how the regional centres may interact with the present global programme, with its residual legal body, its elected Directors, and its global Team.

B.1.4 Networking Activities

B.1.4.1 Overall Strategy

The Networking Activities are designed to foster a culture of co-operation both between the partners in the programme (as is already developing), and between them and the different user communities – something that is new and has not been undertaken before. The intention is to generate from these activities both a stronger more professional infrastructure organisation and an enriched and more valuable set of scientific services. These activities are grouped in three work packages (WP 2, 3, and 4) each of which has the goal of both raising levels of co-operation and bringing new or enhanced service capability to the programme. These work packages are led by Sara Oldfield, Secretary General of Botanic Gardens International, an international organisation based at Richmond, UK (WP 2), Thierry Bourgoin, Deputy Director of the Collections at the Museum National d'Histoire Naturelle, Paris (WP 3), and Jiri Kvacek, Leader of Science programmes at the National Museum in Prague (WP 4). The whole project is co-ordinated by Frank Bisby at the Species 2000 Secretariat at the University of Reading, including the consortium management (WP 1).

WP 1: Management of the Consortium
Work Package 1 provides the consortium/financial management required by the Commission. The proposal is submitted and co-ordinated by the Species 2000 Secretariat, with support by the University of Reading Research Enterprise Services, and will be handled in the same way as the FP5 EuroCat project.

WP 2: Engagement with scientific users and partners
In Work Package 2 it is proposed initially to survey and identify more accurately which are the classes of users, and how they make use of the present services. Two of these classes are reasonably well-known to us – the major global biodiversity portals, and the national portals. However, the third and fourth classes, the individual scientists and citizens in Europe and worldwide who use the web-site and the distributed CD systems, and the commercial organisations that may be interested, all need to be better characterised, and processes put in place to engage with at least some of them. The needs of these four classes of users are quite

likely to be rather different, so the intention is to engage separately with each of them. Surveys, bulletins and small workshops will be used to elicit needs and suggestions that can then be turned into pilot proposals for new services. In the case of the major global portals we are already talking in terms of User/Partnerships with further joint enterprises as well as our contribution to the approaching 'seamless biodiversity informatics' architecture, for instance with the Global Names Architecture (GNA) under discussion with EoL and GBIF. A further task associated closely with user communications is to establish a communications strategy for the programme, and establish the necessary outreach and communication programme.

Without prior knowledge as to exactly what they will be, we have broken these new services into 'educational and popular' services likely to address students and citizen scientists, and 'technical services', such as a variety of new web services built on the service-based architecture of the programme as a whole and designed to serve the technical informatics communities in biodiversity, climate change, food security and conservation. These new services will be designed and prototyped by staff working under WP 5 (Services) and supported by software innovations under WP 6 (Software Support), before being returned to the originating user groups for testing in the later part of the project. Additionally some of the suggestions may be best handled as alterations or improvements to the present services and interfaces.

In practice we already have a substantial set of suggestions in hand to which the new ones can be added, but the interactions with user groups and the service work packages are much needed – to determine which suggestions should be given high priority, and to balance these priorities against their practical and cost implications.

Initial suggestions for the two groups of new services include:

i) Educational & Popular

- Common Names in the Taxonomic Hierarchy
- Thumbnail images of representative or iconic species in the Hierarchy
- Species total tally counter on the web-site
- Recording current species totals in the hierarchy
- Catalogue of Life 'snippets' for use in browsers and installation on web-sites.
- Services to hand-held devices
- Downloadable subsets and displays for use in hand-held devices
- Downloadable subsets for countries and continents
- Fuzzy spelling and phonetic variant searching

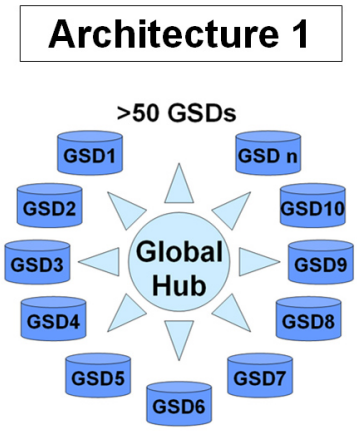
ii) Technical

- GUID-based services highlighting name changes, and circumscription changes
- GUID-based services for incremental cross-maps between annual editions
- GUID-based alert services and dynamic real time provision to certain users
- A range of download services using standard protocols and data models
- A range of web services, e.g. synonymy server, geographical subset server.
- Resilient and load-sharing services able to serve certain users at high load.

WP 3: Strengthening the GSD Network: a Virtual Community

In Work Package 3 we plan to build on the participatory network of supplier Global Species Databases (GSDs) already established in the previous FP5 project ("Architecture 1"), but with a number of issues to be discussed and agreed, each of which has the potential to enhance the existing services. This is the European part of the distributed array of databases on whom the project depends totally – they supply the expert taxonomic sectors that are harvested and

integrated to make up the content of the Catalogue. This Network pulled together in a very co-operative way during the FP 5 project, agreeing data standards, best practice, access agreements, and linking up to enlarge the Catalogue and bring the whole integrative process to production scale.



A network of 29 organisations in a project such as this would normally be seen as a management liability. We stress that the network is already well established with an administration in place and its scientific contribution under control. Not only has it functioned continuously since the FP 5 EuroCat project, both in administrative terms (email list, access agreements and membership of the legal body) and in scientific terms (supply and integration of their data into the Catalogue yearly or continuously), but it has also continued to hold annual meetings, and another 10 databases have joined.

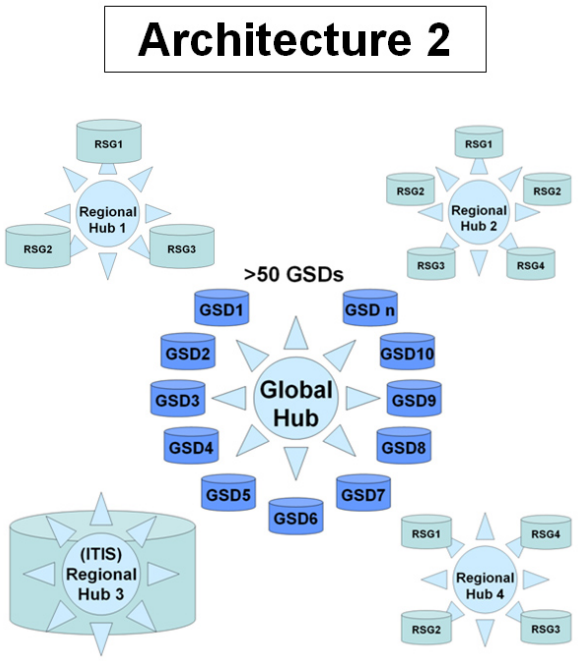
The issues for debate and decision among the Network are:
i) Compliance issues: moving some of the databases towards better compliance with the agreed data standards and best practice: this will impact directly on improved quality and better uniformity in the Catalogue services.
ii) Proposals to agree a small increase in the data set, to cover life forms etc., and to standardise the distribution records - both features that will enhance the services in the eyes of many users. For instance if the distribution data can be structured better, it will eventually become possible to download the complete all-taxon species list as a selected subset for each country in the world, a major step forward.
iii) Wrapper programs and caching frequency: now that many of the GSDs are connected on-line to the SPICE cache, we should like to experiment with increased frequency of caching, so that the Dynamic Checklist can be refreshed at a higher frequency. Some new databases are still to be connected, and we will want to test the wrapper system for the new architecture when it comes on line at the end of the project. These steps will increase the currency of the Catalogue services for users.
iv) Certification, Peer Review and Access agreements. Many of the GSDs in the network became connected in 2005/6, and will therefore need renewed Certification, Peer review, and renewed Access agreements – all of which were planned with a 3 – 5 year cycle time. In the long term it is these that secure the supply and quality of supply to the Catalogue and its services.
v) GSD alerts and services from Nomenclators. This is to discuss possibilities for a new layer in the 'ecosystem' of services creating the Catalogue. Traditionally the GSD custodians have gleaned and entered newly published species to their checklist sectors by hand – but now there is the possibility of gaining them electronically as alerts or feeds from the Nomenclator databases that assemble lists of names for certain sectors – Fungi, Plants etc.
vi) In addition to these material tasks, it is vitally important for the sociology of the whole programme that the GSD custodians in this Network continue to meet, communicate, and feel ownership of the programme. Most are also legal 'members' of the Species 2000 residual body, thus playing a formal part in the long-term governance.

Each GSD partner is funded for 4-months of work on two 2-month pilot tasks. The exact tasks undertaken will be negotiated with each GSD after their assessment in year 1. Four classes of pilot project will be possible:

i) work on the current SPICE-based wrapper and installation
ii) alterations to the GSD database to increase compliance with the Sp2000 data standard and best practice for management
iii) alteration to the GSD database and trialling of the proposed new fields
iv) trials of wrapper development for the new e-2 system in year 3.

The outcome from these discussions and the array of pilot projects will be:
a) improved harvesting performance of each GSD in the Network
b) improved fill and uniformity of fill across the taxonomic sectors of the Catalogue – and hence an improved resource and service capability.

WP 4: Foresight Study for Geographical Extension: the Global Multi-Hub Network
Workpackage 4 establishes a completely new network that represents the extension of the programme to become truly global, and to establish the Architecture 2 foreseen for future operation. It networks the leaders of the Regional Centres that will establish the ring of Regional Taxonomic Database Hubs. This development will extend both the data coverage and the organisational coverage of Species 2000 eventually to every region of the world.



A key issue in the programme is data for the missing species – at 1.2 million species in the current system, there are an estimated 700,000 species still missing. For those taxonomic sectors with GSD databases, these already provide to the global hub nearly complete world coverage of the species in those sectors. But for those taxonomic sectors where no GSD exists, where in the world may we find electronic records for the remaining species? To a large measure we expect to find those records in Regional Species Databases (RSDs), particularly in the so-called mega-diverse countries: China, Brazil, New Zealand, Colombia etc. So this is one of the reasons for embarking on Architecture 2, and the installation of regional hubs. However the notion of forming a regional database hub in each of these areas has other reciprocal attractions for biodiversity managers in those regions, including benefits from linking their own taxonomic databases with the global hub of the Catalogue.

So this Workpackage is to bring together the leaders of the first batch of agreed Regional Hubs in what is now being called the 'Multi-Hub Network'. The object is to discuss a range of concepts and proposals, and to agree an overall concept and implementation strategy under Architecture 2 so that the Multi-Hub Network can be prototyped, and later implemented. The Workpackage does depend on modest funds being made available for the Multi-Hub Network partners from outside Europe to take part. In each case the major funding of the Regional Hub has already, or will shortly be made available from other sources in their own region.

For each of these five partners, we make the special case that a senior or technical staff member from that organisation will be needed for these consultations and for several meetings

in Europe. The *Species 2000 China Node,* already established and funded by the Chinese Academy of Sciences, and the *Catalogo da Vida Brasil,* in development by CRIA in Campinas, working with the Brazilian Ministry of Science & Technology, are both nationally designated centres within third countries listed as International Co-operation Partner Countries. The *New Zealand Organism Register (NZOR)* nationally funded at LandCare Research, a Crown Institute; the *Integrated Taxonomic Information System (ITIS)* a US Federal Government project at the Smithsonian Institution, National Museum of Natural History; and the *Atlas of Living Australia (ALA)* a federally funded programme at CSIRO (Entomology) are nationally designated centres within third countries. We make the case that only nationally designated centres within these five countries can provide the technical and institutional expertise to operate these regional hubs successfully within those countries, and for the biota of those regions. In each case this expertise is not available within Europe, and is essential for the global extension of the programme.

The three Networking work packages largely engage different subsets of partners for rather different purposes, although all related to extending and enhancing the public services. Much of the time they will work separately, but at the second project workshop we plan a day-long session at which each group will be present to discuss their proposals. This is not only for information and interaction, but also against the possibility that synergies across the groups will generate yet further ideas for enhancing the services. For instance, the regional databases to be linked in WP 4 may be able to provide a richer set of common names than can be provided by the global databases in WP 3. Also the Regional centres may be able to set up mirror sites with services provided in their different languages, for example, making the Catalogue available in Chinese and Portuguese interfaces.

B.1.4.3 Networking Activities: Work package list

Work package No[11]	Work package title	Type of activity[12]	Lead partic no.[13]	Lead partic. short name	Person-months [14]	Start month[15]	End month[5]
1	**Management of the Consortium**	**MGT**	**1**	**UR**	**26**	**m1**	**m36**
2	**Engagement with scientific users and partners**	**COORD**	**2**	**BGCI**	**33**	**m1**	**m36**
3	**Strengthening the GSD Partner Network: a Virtual Community**	**COORD**	**3**	**MNHN**	**167**	**m1**	**m36**
4	**Foresight Study for geographical Extension: the Global Multi-Hub Network**	**COORD**	**4**	**NARODNI**	**40**	**m1**	**m36**
	TOTAL			**4**	**266**		

B.1.4.4 Work Package Descriptions

Work package 1: Management of the Consortium

Work package number	WP 1		Start date or starting event:		Month 1		
Work package title	Management of the Consortium						
Activity type[20]	MGT						
Participant number	1						
Participant short name	UR						
Person-months per participant	26						

Objectives
i) To establish and manage the consortium
ii) To manage delivery and reporting of the project tasks and deliverables for the consortium.
iii) To administer the community financial contribution for the consortium.
iv) To manage the audit for the consortium.
v) To participate in the Commission's Concertation and Dissemination programmes.

Description of work (possibly broken down into tasks) and role of partners
Task 1.1 Establish the 4D4Life project web-site and its links to the Catalogue of Life.
Task 1.2 Establish and administer a Management Committee to manage, monitor and report the progression of the project tasks, the completion of milestones and deliverables, and compliance with the grant agreement.
Task 1.3 Review the reports on tasks, and the achievement Project Milestones and Deliverables prior to transmitting them to the Commission for the Consortium.
Task 1.4 Manage the keeping of records and financial accounts, and the receipt, claiming and distribution of funds to beneficiaries at the start, each year, and at the close of the project.
Task 1.5 Prepare and submit the periodic Management Reports, and the Final Report.
Task 1.6 Concertation: The project will actively participate in joint activities and meetings related to the Commission's e-Infrastructures Programme and other related areas, including the participation and contribution to relevant working groups established under the above initiative. The objective of the concertation activity is to optimise synergies between projects and the collective impact and value of the Programme.
Task 1.7 Dissemination: The project will provide input for relevant European Commission initiated dissemination activities (e.g. press releases, news bulletins, brochures, success stories, posters, web-based publications, multimedia material etc.). In this context the project's dissemination messages will also reflect its broader societal and economic impact. The project's dissemination material in relation to the above goal will be regularly updated in deliverable D1.1.

Deliverables (brief description) and month of delivery
D1.1 mth 5 Launch the 4D4Life project web-site at www.4D4Life.eu and update regularly. [1 PeM] (D/PU)
D1.2 mth 14 First Periodic Management Report of Networking, Services & Joint Research for year 1. [6 PeM](R/PP)
D1.3 mth 26 Report on exploitation of the foreground IPR created across the project. [5 PeM] (R/RE)
D1.4 mth 36 Final Report on the 4D4Life enhanced scientific data e-infrastructure. [14 PeM] (R,PU)

Milestones (brief description) and month of delivery
M1.1 mth 6 Management Committee and full task monitoring in place.

Work package 2: Engagement with scientific users and partners

Work package number	WP 2	Start date or starting event:				Month 1	
Work package title	Engagement with scientific users and partners						
Activity type[20]	COORD						
Participant number	2						
Participant short name	BGCI						
Person-months per participant	33						

Objectives

i) To establish a communication and promotional programme that will provide ongoing communication with the classes of users and disseminate the infrastructure services even more widely than has occurred so far.
ii) To locate and characterise the main classes of external and internal users of Catalogue of Life services, and to engage with each of them, including:
- the major global biodiversity portals and biodiversity organisations that utilise the Species 2000 Catalogue of Life services, to review current and potential uses and enhancements.
- the substantial number of national portals and programmes, NGOs, and higher education institutes that utilise the services, to review current and potential uses and enhancements.
- particular subsets of the 40,000 individuals who make serious use of the services each month to review current and potential uses and enhancements.
- the array of commercial organisations that either do use the services, or may be considered likely to do so – for instance in natural product chemistry, in the seed and horticultural services, and the biodiversity image agency trade.
iii) To work with each class of users and partner organisations, to specify a list of possible new or altered services.
iv) To work with the Services Team to select a subset of priority services that are feasible to provide.
v) To work with the Services Team to test and then launch the selected new services.

Description of work

Task 2.1 Test, establish and operate user and partner communication platforms designed to inform on progress, stimulate engagement and promote the Species 2000 Catalogue of Life, through
- a quarterly e-bulletin (that will include information on new services),
- workshops and events (to which major institutional and potential users are invited).

Task 2.2 A renewed analysis of the main classes of external and internal users, including:
- *current and potential users*, to analyse changes over the last 5 years, the potential for retaining them, and for engaging additional users.
- *geographical distribution of users*, to analyse changes over 5 years, and likely change in the future,
- *uses to which the CoL is put*, how this changed over five years, and is likely to change in the future.

Task 2.3 Engagement with the different classes of users to generate a prioritised list of:
- proposals as to how the current user-interface and web-services might be enhanced,
- proposals for completely new educational and popular services, including mobile device services,
- proposals for completely new technical services, including further web-services.

Task 2.4 To work with the Services Team to conduct a triage for which priority services are feasible, taking into account data supply (WP 3), software alterations cost (WP 6) and sustainability (WP 5).
Task 2.5 To assist in beta testing and public launch of two batches of new services once they are implemented.

Deliverables

D2.1 mth 6,12,18,24,30.36 Rolling deliverable. Six-monthly e-bulletin. [8 PeM] (O, PU)
D2.2 mth 11 User-driven proposals for new services for each class of users. [10 PeM] (R,Re)
(to include educational, popular, technical and electronic, as well as internal services)
D2.3 mth 34 Report on new services tested with users and ready for launch [15 PeM](R, PP)

Milestones	
M2.1 mth 11	Analysis of Use and User classes for the Catalogue of Life.
M2.2 mth 24	Ongoing outreach policy implemented

Work package 3: Strengthening the GSD Partner Network: a Virtual Community

Work package number	WP 3		Start date or starting event:			Month 1			
Work package title	Strengthening the GSD Partner Network: a Virtual Community								
Activity type[20]	COORD								
Participant number	1	3	7	9	11	10,12-22, 24, 26-32	23	25	33
Participant short name	UR	MNHN	NNM	VLIZ	NHM	(GSD Network)	CSIC	SNSB	ITZN
Person-months per participant	8	46	4	4	6	4 each (total 80)	8	8	3

Objectives

i) To engage the custodians of the GSD databases in discussions and actions in how best to develop and sustain their participation in and ownership of the Species 2000 Catalogue of Life programme.

ii) To initiate active steps among the GSDs to move towards improved compliance with the current Species 2000 Data Standard and the current Sp2000 GSD Best Practice document, so as to raise the quality of the Catalogue of Life by making the product better filled and more uniform.

iii) To explore and pilot possible small extensions to the bio-data knowledge set provided by the GSDs, so as to enhance the Catalogue of Life services for users, including responses to service requests prioritised by WP 2 and the Services Team.

iv) To explore and pilot possible extensions of internal services within the programme, such as name-alerts from nomenclator databases, feedback comments from users, taxa harvested from regional hubs, cross-checks with other lists, and dealing with unassigned name batches supplied from a variety of sources, so as to increase the quality, currency and responsiveness of the Catalogue of Life services, and to report on these to the Software Design Team and the Services Team.

v) To roll out service-based delivery to the new e-2 Architecture being developed in WP 7 and WP 6.

Description of work (possibly broken down into tasks) and role of partners

Task 3.1 A survey to assess, peer review and certificate each provider GSD, its wrapper & delivery, and its formal membership of the programme, including data capture for the improved metadatabase.

Task 3.2 Undertake prior consultation with appropriate partners (e.g. GSD custodians, nomenclator custodians, editorial services staff) and prepare a Network Discussion Document, including proposals for i) Changes to the Data Standard for existing fields (if any), ii) Changes to the Best Practice Document (if any), iii) Possible extensions to the bio-data fields, including data supply requests (if any) from the Services Team, and iv) possible internal service supply issues (e.g. name-alerts, comments from users, taxa from regional hubs, cross-checks with other lists, and unassigned name batches).

Task 3.3 Network consultation and workshop with GSD custodians, Nomenclator custodians and Editorial Services Staff (WP 5), to explore each of the items highlighted in the Network Discussion Document; to explain and establish priorities for the pilot project for each database; to respond to requests from the Services Team, and to provide additional training days.

Task 3.4 Two 2-month Pilot Projects undertaken per GSD database:

- for partially compliant GSDs to take a step up the compliance ladder (eg add fields or best practice)
- or for new GSDs (since EuroCat project), to install and test wrapper and delivery process.
- or to install and test fill the new bio-data fields from WP 8.
- or to pilot participation in one of the new internal services of the programme (name-alerts, etc)
- or a fourth class, not available until later (to test install new e-2 system wrapper)

Task 3.5 Network consultation and workshop reporting on progress with the four previous classes of Pilot Project and additional training days for the new e-2 system wrapper, with opportunity for the last set of GSDs to:

- undertake the fourth class of Pilot Project – to test install the new e-2 system wrapper.

Deliverables (brief description) and month of delivery
D3.1 mth 11 Report on agreed target set of improvements to the data supply by the Global Species Databases (GSDs), (includes Revised Data Standard & Best Practice Documents). [55 PeM] (R/PU)
D3.2 mth 24 Report on pilot projects for data provider improvements (1st batch). [56 PeM] (R/RE)
D3.3 mth 36 Report on pilot projects (2nd batch) including some that roll out the new architecture for data providers.
[56 PeM] (R/RE)

Milestones (brief description) and month of delivery
M3.1 mth 11 GSD Survey complete, brief listing available.
M3.2 mth 11 Network Discussion Document, including data supply requests from Services Team.
M3.3 mth 12 Response available on data supply requests to Services Team.
M3.4 mth 34 GSD Survey updated, brief listing available, and data in metadatabase

Work package 4: Foresight study for geographical extension: the Global Multi-Hub Network

Work package number	WP 4		Start date or starting event:		Month 1	
Work package title	Foresight study for geographical extension: the Global Multi-Hub Network					
Activity type[20]	COORD					
Participant number	4	34	35	36	37	38
Participant short name	NARODNI	CAS	CRIA	National Museum of Natural History	CSIRO	Landcare Research
Person-months per participant	32	4	4	n/a	n/a	n/a

Objectives

i) To bring together for the first time leaders of the various international centres from around the world that have committed to create Regional Hubs in the Sp2000 Multi-Hub Network, so that they can work together in Networking, Discussion and Design.
ii) Foresight Study to report the Needs, Conceptualisation, Specification and subsequently the Design both for individual Regional Hubs, and for the Multi-Hub network as a whole.
iii) To carry out a limited amount of prototyping and testing to explore the concepts and design features agreed.
iv) Foresight Study to point the way for other potential regional hub centres, not yet part of the Multi-Hub Network, to form Regional Hubs and join the Network in future, with a view of eventually completing coverage of all continents and oceans, and of including centres from less-advantaged parts of the world in possible future projects.

Description of work (possibly broken down into tasks) and role of partners

Task 4.1 Network consultation and workshop discussions among regional hub organisers:

- to debate the Needs, Concepts and Specification for the Multi-Hub Network at both the taxonomic and the technical level, leading to a Preliminary Specification with outstanding issues highlighted.
- to discuss building on the Preliminary Specification and the Pilot Projects leading to a firm Concept and Specification of the Multi-Hub Network.
- to propose an institutional structure frot eh Multi-hub Network within Sp2000, and what structural changes may be needed.
- to report on the Pilot Projects.
- to prepare an outline plan for the extension of the Multi-Hub Network to cover the whole world, and to include new centres, some from less advantaged parts of the world.

Task 4.2 A small number of possible Pilot Projects, depending on the partners resources:

- for prototyping and testing based on the limited Multi-Hub Linkage tools available with the present SPICE system within Sp2000.

Deliverables (brief description) and month of delivery

D4.1 mth 12 Report on preliminary concept, requirements and specifications for the Global Multi-Hub Network: the *Preliminary Requirements Document*. [20 PeM](R/PP).
D4.2 mth 24 Global Multi-Hub Network Design Document, including both a technical design component and an institutional design component. [20 PeM](R/PP)

Milestones (brief description) and month of delivery

None

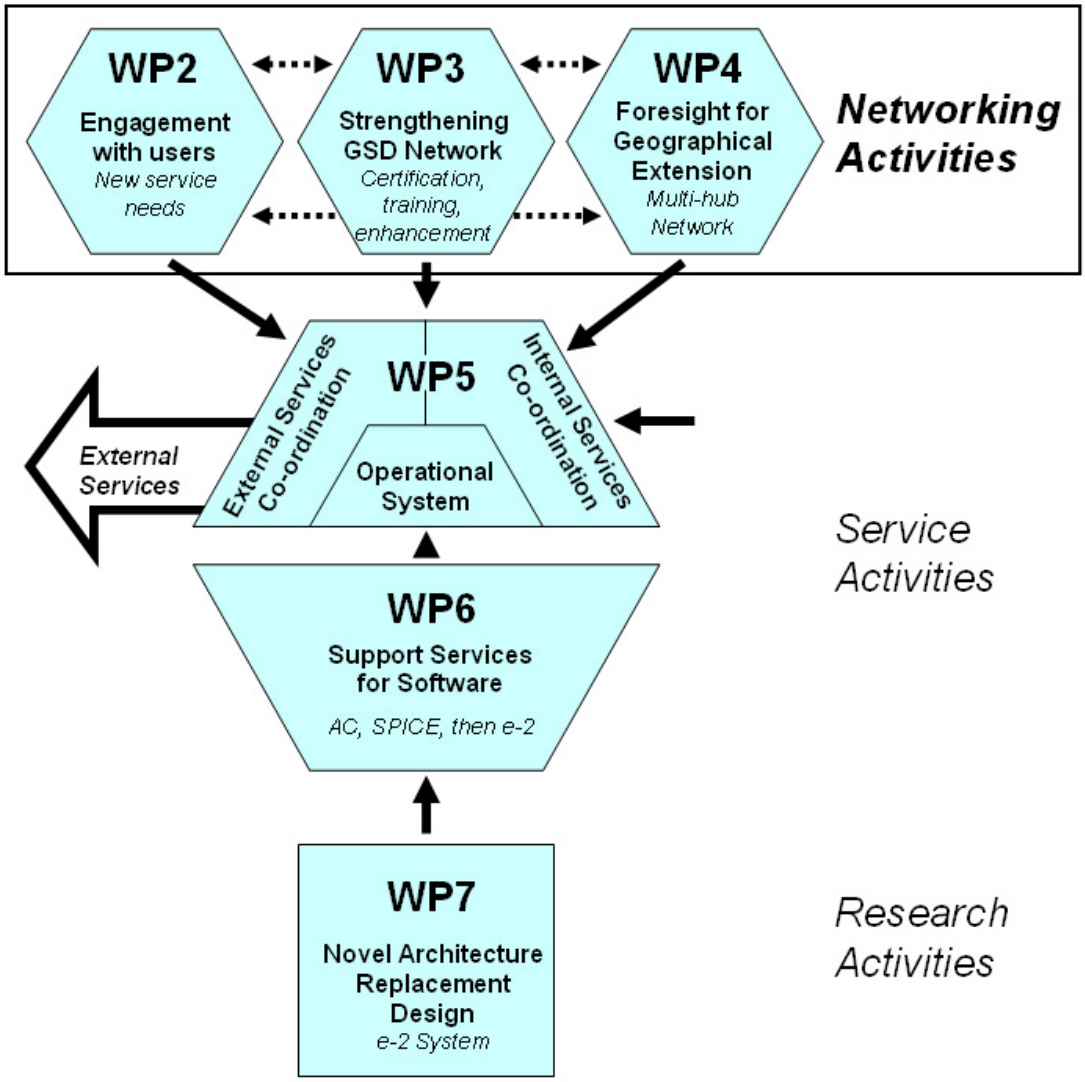
B.1.4.5 Networking Activities: Efforts for the full duration of the project

Partic. no.	Partic. short name	WP 1	WP 2	WP 3	WP 4	Total person months
1	UR	**26**		8		34
2	BGCI		**33**			33
3	MNHN			**46**		46
4	NARODNI				**32**	32
5	ETI					0
6	CU					0
7	NNM			4		4
8	SP2000					0
9	VLIZ			4		4
10	MfN, Berlin			4		4
11	NHM			6		6
12	CABI			4		4
13	RBGK			4		4
14	UNIVIE			4		4
15	UU			4		4
16	NUI Galway			4		4
17	IFM-GEOMAR			4		4
18	DKFZ			4		4
19	MIZ-PAS			4		4
20	Land OOE			4		4
21	NMGW			4		4
22	IRD			4		4
23	CSIC			8		8
24	TSJ BVBA			4		4
25	SNSB, Munich			8		8
26	UNIPD			4		4
27	UVA			4		4
28	RBINS			4		4
29	University of Oxford			4		4
30	WU			4		4
31	RBGE			4		4
32	UCPH			4		4
33	ITZN			3		3
34	CAS				4	4
35	CRIA				4	4
36	National Museum of Natural History				0	0
37	CSIRO				0	0
38	Landcare Research				0	0
Total		26	33	167	40	266

B.1.4.6 Networking Activities: Milestones

number	Milestone name	Work package(s) involved	Expected date	Means of verification
M1.1	Management Committee and full task monitoring in place.	WP 1	mth 6	PPR&PO
M3.1	GSD Survey complete, brief listing available.	WP 3	mth 11	PPR&PO
M3.2	Network Discussion Document, including data supply requests from Services Team.	WP 3	mth 11	PPR&PO
M2.1	Analysis of Use and User classes for the Catalogue of Life.	WP 2	mth 11	PPR&PO
M3.3	Response available on data supply requests to Services Team	WP 3	mth 12	PPR&PO
M2.2	Ongoing outreach policy implemented	WP 2	mth 24	PPR&PO
M3.4	GSD Survey updated, brief listing available, and data in metadatabase	WP 3	mth 34	PPR&PO

PPR&PO – Achievement of each Milestone will be reported in the 12-monthly Project Periodic Report, and notified to the Programme Officer.



Relationships between Workpackages

B.1.5 Service Activities

B.1.5.1 Overall Strategy

The whole purpose of the Species 2000 Catalogue of Life programme is to synthesize the Catalogue of Life as an ongoing integrated knowledge resource, and to use this as the basis for scientific infrastructure services to science and society. This section of the project operates those two processes – the production of the integrated resource, and the generation of services using that resource: it is at the very core of the programme. Also it will be the beneficiary of ideas for new services and enhancements generated by the Networking Activities and the new e-infrastructure developed in the Joint Research Activity, so that by the end of the project these new services or enhancements will be incorporated in the main production stream leading to the public services. But between the start and finish, these two service work packages have also a very substantial role in conceptualising, creating, testing and launching those new services and features.

The work is stratified into two service layers. The taxonomic integration of the knowledge and its synthesis into the Catalogue, plus the operation of the public services are done from Reading. As seen by the public, and as seen by the participants in the project it is the Reading centre that undertakes these tasks. But the integration process and the operation of the public services both depend on having specialised software in place with which to run these internal and external services. ETI will underpin software support services for these two processes. It will take responsibility for generating and maintaining the production software, both as an internal service to participants in the programme, and for the outward facing public services themselves. The integration and synthesis of the Catalogue, and the operation of public services are led by Frank Bisby at University of Reading (WP 5), and the software support services for both the internal and external services are led by Peter Schalk at ETI in Amsterdam (WP 6).

WP 5 Scientific Services of the Catalogue of Life
Work Package 5 provides the scientific co-ordination of the whole project. This includes the steady state synthesis and production of the Annual and Dynamic Checklists, the steady state Systems operation that operates both the internal and external services, and the co-ordination of the new features to be tested and introduced.

Given that the services are already widely used, it is important that they should be maintained and strengthened continuously throughout the three year project. One task that must be continued in steady state operation is the integration of the Catalogue and production of the Annual Checklist for publication each April. This involves both the introduction of new taxonomic sectors from new supplier GSD databases, and the integration of updated sectors from existing suppliers. Each year a number of new suppliers are added, and it is these that require the greatest work for the ‘admission gate’ process: recording metadata, technical scrutiny, peer review, administrative agreements, as well as test downloading data and writing and testing the wrapper program. Receiving or re-caching from the existing suppliers involves less time. Integrating the new sectors into the taxonomic Hierarchy and Checklist is a skilled taxonomic editorial task, part of the work of the Editorial and Production Team, divided between Reading and Los Baños in the Philippines.

A second task needed to maintain the services is the systems operation. Not only is there the need to operate components 24/7, but there are ongoing complexities to be managed, particularly in the area of assisting suppliers with writing and testing wrapper programs, operating the test and production hubs, and managing the availability and caching from the various databases – a complex task with 52 suppliers, expected to rise to 66 and more, divided between Europe and world-wide.

Against this steady state background, a substantial set of changes have to be managed and intercalated during the three years of this project. These are:

i) improvements planned for the production process in year 1, working with WP 6. In yesr 2 there will be the unification of the production process for the Annual and Dynamic Checklists, rather separate up till now, and a system giving more support from the metadatabase when new databases are connected. Implementation of these alterations will be managed by the head of the Editorial and production team. It will need to be closely co-ordinated with the annual production cycle.
ii) Work with WP 2 and WP6 to select, design and create the new services, to test them, and then to introduce them into the operational system. The Services Team and a new services manager will assess the possible implementation of each suggested new service, taking into account both the supply side issues with GSD databases in WP 3, and the middleware/software enhancement issues with WP 6. Decisions as to which new services to go ahead and implement will be taken after a 'triage' between priorities set by the users and the practicalities of putting it in place.
iii) Work with WP 4 to test the present Multi-Hub Network features and to agree a long-term specification.
iv) Work with WP 7 and WP 6, to test the new distributed system, and then to roll it out in the third year.

This work package will also undertake the task of creating a second generation business plan, and of initiating the financial and legal engineering needed for the programme to approach sustainability. It will be undertaken with advice from the user community contacted in WP 2, with discussion with the GSDs partners in WP 3, as their endorsement is a prerequisite, and in partnership with the directors of Species 2000 (partner 8), the residual legal body through and for which financial sustainability will be achieved.

WP 6 Software Support Services
Work Package 6 will provide a newly co-ordinated middleware and software support service within the Species 2000 community. In the past the programme has either adopted software products resulting from informatics research pilot projects, or contracted individual tasks to particular IT laboratories. Despite some extremely successful outcomes, this has left the programme dependant on more than one system, and in some cases on working prototypes that lacked the robustness, management features and documentation that are needed for what is becoming a mature public service. The middleware and software involved provide both the internal processes by which taxonomic sectors are harvested from the many databases, but also the user-facing processes by which the public services are served from the servers.

Responsibility for co-ordinating this service is taken in this project by ETI in Amsterdam, although they will carry out some of the work, and spread other tasks among other specialist members of the work package.

The work package undertakes quite a long list of tasks that are needed for four main themes:

Theme 1: Unification and improvement of the existing Checklist creation processes.
The secure editorial process of the present synthesis and production of the Annual Checklist product needs to be combined with the automated caching process of the Dynamic Checklist product, and linked more closely with the metadatabase. This will be done by securing an altered base schema that is suitable for both systems, and making adjustment so that both systems can operate on this one schema and linked to the metadatabase. Once this change is made, all the supporting tools of both systems can be brought to bear on a single production process. Other associated improvements to the SPICE 5.1 distributed system concern a better wrapper kit for the GSD suppliers, better automation of the attachment process for new GSDs, and better management interface and functionality for caching and re-caching the sectors supplied by the GSDs. Many of these features are improvements in the present middleware as seen and used by the partners within the programme.

Theme 2: New services
WP 5 will co-ordinate the response to suggested new services, but this work package will assist them, first to assess how much software alteration may be needed for each, and once it is decided which to go ahead with, to make those alterations or new features prior to testing, and eventual implementation. The interface and software alterations made will primarily impact on users, but to provide content for these, it may be necessary to make middleware changes back to the origin of the information.

Theme 3: Multi-Hub connections
Features needed for the Dynamic Checklist to link regional hubs to the Global Hub were foreseen in the EuroCat project and first steps to implement these were made at the end in 2006. They were incorporated into SPICE version 5.2, the last version of the Dynamic Checklist software in that project, and used to link the three European regional databases into the Euro-Hub, and to link the Euro-Hub with the Global Hub. It is now proposed to extend that system so that the N. American Hub, and the China Hub can be linked in year 1 on an experimental basis. That structure, with three regional hubs connected in this simple way, will then provide a background test and prototype for the discussions and long-term specifications being formulated in WP 4.

Theme 4: Introducing the new e-2 infrastructure.
Scoping, designing, and preliminary testing of the new e-2 distributed system are the separate responsibility of WP 7, although many of the altered requirements will be coordinated with opinions from this work package. But once the new system prototypes have been tested and a second prototype made available, this Work package 6 will take over the process of preparing a second, production version of the system to be rolled out across the programme.

The separation of tasks between this Work Package 6, the specification of future needs for the Multi-Hub Network in WP 4, and the prototyping of the next generation e-2 distributed system in WP 7 means that for the first time the Species 2000 Catalogue of Life programme will have proper attention to the maintenance and functionality of its current production middleware and software, at the same time as others undertake the design and prototyping for the next generation of production systems.

B.1.5.3 Service Activities: Work package list

Work package No	Work package title	Type of activity	Lead partic no.	Lead partic. short name	Person-months	Start month	End month
5	**Scientific Services of the Catalogue of Life**	**SVC**	**1**	**UR**	**146**	**m1**	**m36**
6	**Software Support Services**	**SVC**	**5**	**ETI**	**83**	**m1`**	**m36**
	TOTAL			**2**	**229**		

B.1.5.4 service Activities: Work Package Descriptions

Work package 5: Scientific Services of the Catalogue of Life

Work package number	WP 5		Start date or starting event:		Month 1		
Work package title	Scientific services of the Catalogue of Life						
Activity type[20]	SVC						
Participant number	1	8					
Participant short name	UR	Sp2000					
Person-months per participant	110	36					

Objectives
i) To strengthen and substantially enhance the public scientific infrastructure services provided by the Species 2000 Catalogue of Life programme, including an array of novel services both at the popular and educational level, and at the electronic services level. This will involve engaging closely with the User and Partner communities through the networking tasks of WP 2, as well as working with the knowledge providers in the GSD Network of WP 3, and of the new Multi-Hub Network in WP 4.
ii) To strengthen and make secure the internal community services that take the raw taxonomic materials from an ecosystem of providers and synthesise it into the authoritative Species 2000 Catalogue of Life and its public infrastructure services. These internal services include the year on year production of the Annual and Dynamic Checklists; the work of the Editorial Team in taxonomic integration, editorial scrutiny and quality control; and the operational systems management needed to operate the internal and public services continuously.
iii) To undertake the financial and legal engineering needed to make the Species 2000 Catalogue of Life programme sustainable as a scientific infrastructure service on a European and global scale.
iv) To undertake the scientific co-ordination of the whole project whose central objective is to enhance and secure the public services for the Species 2000 Catalogue of Life.

Description of work (possibly broken down into tasks) and role of partners
Task 5.1 Design, test and implement a preliminary batch of interface enhancements to the present public services, working with the Services Team.
Task 5.2 Advise the Services Team on the sustainability of each of the new services in the prioritised list of proposals, and contribute to the triage by which the most feasible are selected for implementation.
Task 5.3 Work with the Services Team to design the novel services that have been selected, and to co-ordinate their implementation, testing and launch.
Task 5.4 Work with the Systems Design Team and WP 6 to enhance and operate the annual production service for the Annual and Dynamic Checklists, including unification of the admission gate for the two products, and the roll-out of the new e-2 system.
Task 5.5 Enhance, secure and operate the editorial, scrutiny and quality control services for the unified production route, working with the GSD custodians in WP 3 and others worldwide.
Task 5.6 Secure and operate the ongoing systems management for the internal and public services of the whole programme working closely with WP 6.
Task 5.7 Work with the partner 8 (Sp2000) to establish an enhanced business plan for sustainability, and undertake the associated financial and legal engineering needed to implement this plan over the life of this project for continued operation by partner 8.
Task 5.8 Co-ordination of scientific, IT, networking and business model components of the entire project, including the interaction of this central work package with all the other work packages, and supporting the co-ordination work of the Services Team and the Systems Design Team.

Deliverables (brief description) and month of delivery

D5.1 mth 12 Initial software improvements and streamlining incorporated into production of the enhanced Annual Checklist 2010, published on the web, and on CD/DVD. [25 PeM](D/PU)

D5.2 mth 24 Unification of the production process and admission gate achieved for the Annual (2011) and Dynamic Checklists, published on the web and (Annual Checklist) on CD/DVD. [41 PeM](D/PU)

D5.3 mth 26 Business Plan for sustainable development of the services. [25 PeM](R/RE)

D5.4 mth 34 New array of services available and launched with pack of publicity materials and electronic products. [55 PeM](D/PU)

Milestones

M5.1 mth 12 Base schema and associated tools accepted for editing and attaching databases.

M5.2 mth 24 First batch of new services implemented and under test

M5.3 mth 30 Second batch of new services implemented and under test.

Work package 6: Software Support Services

Work package number	WP 6		Start date or starting event:		Month 1		
Work package title	Software Support Services						
Activity type[20]	SVC						
Participant number	5	3	9	35			
Participant short name	ETI	MNHN	VLIZ	CRIA			
Person-months per participant	68	5	2	8			

Objectives
To operate a software development service that provides and supports production-quality software for deployment across the Catalogue of Life networks. This service will include middleware and software for internal services among partners, as well as the software with which the external services are provided to the user communities.
i) Improvement of the existing software, such as that used in the production workflows for the Annual and Dynamic editions of the Catalogue of Life, to create a robust and sustainable infrastructure.
ii) Unification of the two workflows. Unification of the two workflows into a single workflow, an important internal improvement alongside other improvements for the project.
iii) Preliminary linkage of Multi-Hub connections. Test preliminary linkages to enable the N. American and Chinese regional hubs to be linked alongside the European regional hub, to provide a background test and prototype for the discussions and long-term specification being formulated in WP 4.
iv) Implementation of software enhancements needed for the array of new services being created and launched by the project, in two batches.
v) Introduction of the new e-2 infrastructure, by preparing a production version of the e-2 system that will have been researched and prototyped in WP 7.

Description of work n.b. "(+doc)" = plus documentation
i) Improvement of existing software (Specifications and test routines agreed with Systems Design Team)
Task 6.0: Establish development and documentation infrastructure, and test installation, plus preparation and production subcontract for the AC on CD, in co-operation with WP 5 task 5.6.
Task 6.1: Simple SPICE management interface: control caching, attachments, and re-starts (+doc).
Task 6.2: Improved wrapper development kit for SPICE (+doc)
Task 6.3: Improved user-interface for use with both Annual and Dynamic Checklist (+doc).
Task 6.4: Improved CoL metadatabase (+doc)
Task 6.5: GUI for graphical control of attachments (+doc)
ii) Unification of existing workflows (Specifications & test routines agreed with Systems Design Team)
Task 6.6: Research, propose new base schema with Design Team and adopt it at a workshop (+doc).
Task 6.7: Alter AC software to use new base schema (including Philippines routines)
Task 6.8: Alter DC software to use new base schema
Task 6.9: Create & alpha test Quality Assurance Workbench (with editor, filter profiler and cache up-loader) (+doc)
Task 6.10: Install QA Workbench (for beta test and then production unification)
Task 6.11: Version of QA Workbench for use as freestanding GSD Builder and/or GSD Editor (+doc)
iii) Preliminary linkage of Multi-Hub connections (Specs. & test routines agreed with Syst. Design Team)
Task 6.12: Improve and test multi-sector wrapper and cache controls for regional hub attachment (+doc)
Task 6.13: Alter DC Interface to permit toggling and improved labelling when multiple hubs in use (+doc).
iv) Software for new services (Specifications & test routines agreed with Services Team)
Task 6.14: Preliminary scoping and costing for all new services listed, prior to selection.
Task 6.15: Full scoping and costing for just those new services selected for implementation.
Task 6.16: Implement 1st batch of services for test (+doc).
Task 6.17: Implement 2nd batch of services for test (+doc).
Task 6.18: Respond to testing 1st batch, ready for launch

Task 6.19: Respond to testing 2nd batch, ready for launch
v) Production version of e-2 infrastructure (Specs. & test routines agreed with SystemsDesign Team)
Task 6.20: Design and planning for production version of e-2 prototype provided by WP 7.
Task 6.21: Implement production version (+doc), including prioritised enhancements and responses to testing.

Deliverables (brief description) and month of delivery
D6.1 mth 12 Improved management interfaces and streamlined production of the Catalogue of Life delivered for existing software, (Release 2, includes tasks 6.0, 6.1, 6.2, 6.3, 6.6, 6.7, 6.8, 6.9, 6.12 & 6.13). [39 PeM](D/PP)
D6.2 mth 24 Altered software for both unified AC & DC processes, and improved synthesis of the Catalogue of Life operational and installed, (Release 4, includes tasks 6.4, 6.5, 6.10 & 6.11). [13 PeM](D/PU)
D6.3.mth 24 Software for new custom services prototyped and ready for test, (includes tasks 6.14 – 6.19). [25 PeM](D/RE)
D6.4 mth 30 Production software for the new e-2 architecture fine-tuned and delivered, underpinning the virtual community, (includes tasks 6.20 & 6.21). [6 PeM](D/PU)

Milestones
M6.1 Month 6 Base schema proposals ready for review and adoption by Systems Design Team (includes task 6.6).
M6.2 Month 6 Release 1 of enhanced AC & DC Software (includes tasks 6.0, 6.1, 6.2, 6.3, & 6.12).
M6.3 Month 11 Scoping and costing of services software available for Services Team (task 6.14).
M6.4 Month 12 Unified software and Quality Assurance Workbench available for test at Secretariat (task 6.10 beta test part).
M6.5 Month 18 Release 3 of enhanced AC & DC Software (includes tasks 6.4, 6.5, 6.10 production part).
M6.6 Month 21 Software for 1st batch of services available for test by Services Team (tasks 6.15 & 6.16).
M6.7 Month 24 Software for 2nd batch of services available for test by Services Team (task 6.17).
M6.8 Month 34 Copies of all software components and documentation placed in the public software repository operated by the Secretariat.

B.1.5.5 Service Activities: Efforts for the full duration of the project

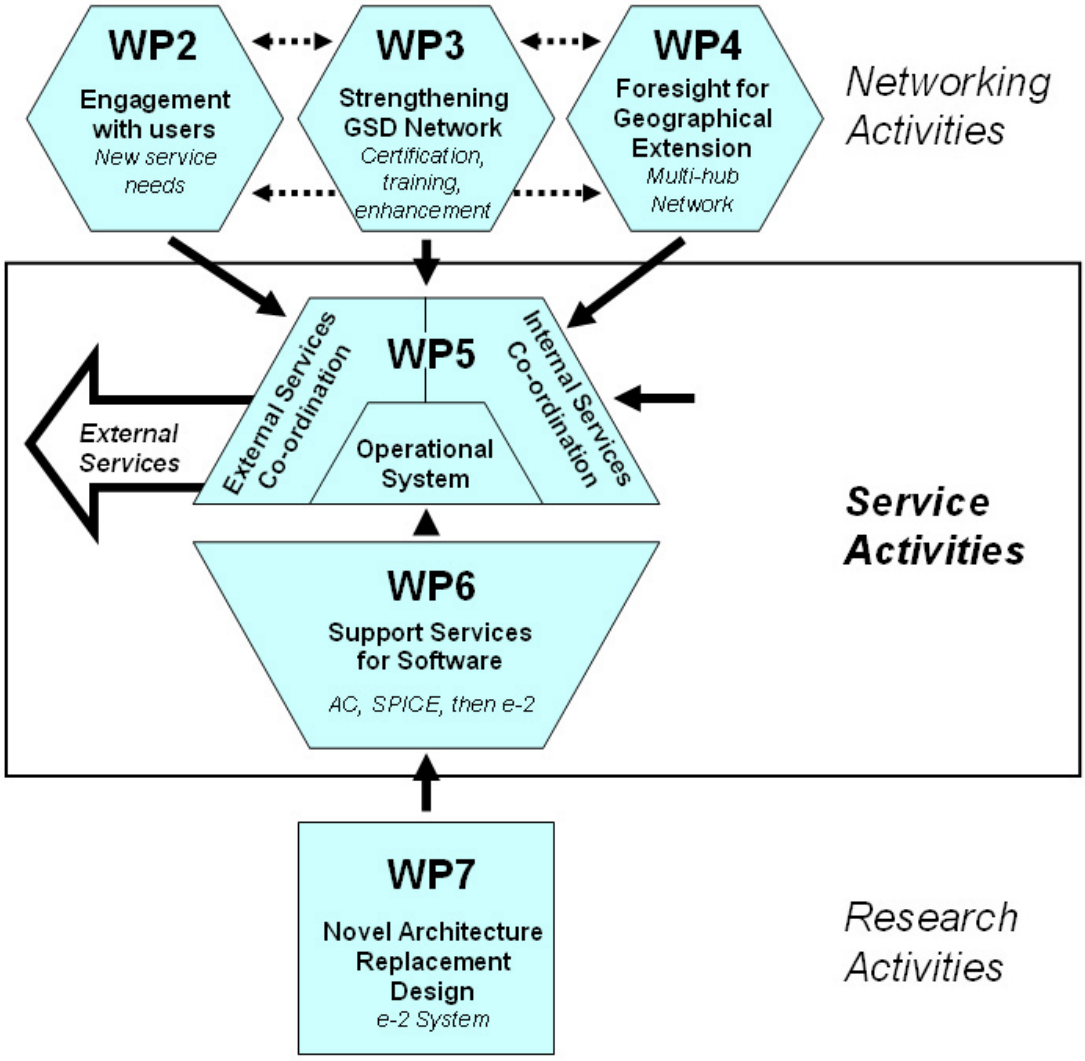
Partic. no.	Partic. short name	WP 5	WP 6	Total person months
1	**UR**	**110**		**110**
3	**MNHN**		5	**5**
5	**ETI**		**68**	**68**
8	**Sp2000**	36		**36**
9	**VLIZ**		2	**2**
35	**CRIA**		8	**8**
Total		**146**	**83**	**229**

B.1.5.6 Service Activities: Milestones

.

number	Milestone name	Work package(s) involved	Expected date	Means of verification
M6.1	Base schema proposals ready for review and adoption by Systems Design Team.	WP 6 & 7	mth 6	PPR&PO
M6.2	Release 1 of enhanced AC & DC Software.	WP 6 & 5	mth 6	PPR&PO
M6.3	Scoping and costing of services software available for Services Team.	WP 6 & 2 & 5	mth11	PPR&PO
M6.4	Unified software and Quality Assurance Workbench available for test at Secretariat.	WP 6 & 5	mth 12	PPR&PO
M5.1	Base schema and associated tools accepted for editing and attaching databases	WP 5 & 7	mth 12	PPR&PO
M6.5	Release 3 of enhanced AC & DC Software	WP 6& 5	mth 18	PPR&PO
M6.6	Software for 1st batch of services available for test by Services Team.	WP 6 & 2	mth 21	PPR&PO
M6.7	Software for 2nd batch of services available for test by Services Team.	WP 6 & 2	mth 24	PPR&PO
M5.2	First batch of new services implemented and under test	WP 5 & 2	mth 24	PPR&PO
M5.3	Second batch of new services implemented and under test	WP 5 & 2	mth 30	PPR&PO
M6.8	Copies of all software components and documentation placed in the public software repository operated by the Secretariat.	WP 6 & 5	mth 34	PPR&PO

PPR&PO – Achievement of each Milestone will be reported in the 12-monthly Project Periodic Report, and notified to the Programme Officer.



Relationships between Workpackages

B.1.6 Joint Research Activities

B.1.6.1 Overall Strategy

The Joint Research Activity is a single work package to replace the current distributed system with a modern service-based architecture using the robust e-Science protocols that the community needs to ensure effective interoperability. It is led by Richard White, assisted by Alex Hardisty and Andrew Jones at the Welsh e-Science Centre and the School of Computer Science at Cardiff University (WP 7).

WP 7: Novel e-Science Service-based Architecture
Work Package 7 is to research, design, prototype and test a completely new service-based architecture and modern implementation for the Species 2000 Catalogue of Life community of contributors. There are several reasons for doing this.

i) Although Species 2000 is proud to have been one of the first European scientific infrastructure services to be based on a distributed system, the fact is that the SPICE system currently in use is distinctly outdated in a number of ways. First, it is based on technologies that were at the cutting edge in 1998: CORBA is no longer an appropriate technology for our purposes, and our use of XML needs to be revised to comply with current standards.

ii) It was about three years in advance of other biodiversity programmes that subsequently used similar technologies, and so the system made innovations in Common Data Model, and communications formats that have in recent years been overtaken by more extensive models and protocols that have now become generic standards, particularly of the TDWG organisation. We now need to make use of these more widely used generic standards.

iii) The most important reason is simply that the requirements for the architecture have changed.
- One issue is that the 'ecosystem' of providers and harvesters has changed. (Architecture 1 referred to in the introduction will here be referred to as 'knowledge architecture 1' to distinguish the knowledge architecture from the computational architecture used in the implementation – they need not be congruent.) Knowledge Architecture 1 required a single level array of GSD databases to be harvested by a single global hub. Discussions in WP 3 of this project will finalise ideas on inserting a layer of providers (the Nomenclators – databases of names) 'beneath' the layer of GSDs. These nomenclators can assist the GSDs by sending them alerts to the publication of new names, even though they cannot provide the expertise that the GSD provides, of positioning these newly discovered organisms within the classification. Similarly the discussion and specifications for the Multi-Hub Network in WP 4 will consolidate the higher level array of multiple hubs. So what started as a simple one-layer distributed array, does now resemble the idea of an ecosystem with a three level 'food chain' of providers and harvesters.
- Another issue is that the original federated model allowed for real time harvesting, as there were severe reservations about permissions to cache our partners' data. In fact the community has come to respect and trust the responsible handling of other people's data by Species 2000 both by seeing how it has behaved, and by virtue of the Access Agreements agreed with the suppliers, and their becoming members and part owners of the system and its governance. It also became clear that real-time harvesting required unrealistic performance over the network,

leading in practice to poor response times, something much improved by caching. As a result the system requirement now is to operate from a cache and indeed with a buffer with potential editorial intervention, and certainly not in real time harvesting.
- A further issue is of the openness of the ecosystem. In the present system the individual GSD SPICE wrappers would be open to harvesting by any organisation using the SPICE protocols, but in reality other organisations have not done so. There is the prospect that new standard providers coming from TDWG or GBIF may make it possible for a database to serve specimen and taxon data using a common provider protocol, thus avoiding the need for different wrapper/provider middleware for different projects, and making individual service providers available for direct harvesting by others.

Several other features may be needed in the new architecture.
i) integration of LITCHI tools for testing the integrity of checklists covering the same species in different hubs
ii) possible 'whole biota arithmetic' and other features from the Multi-Hub Network, some requiring advanced network capacities
iii) structures for mirroring between the regional hub sites, both for resilience against down times, and for load-sharing under heavy load.

The proposal is to make a fresh requirements analysis in year 1, and to enhance this with possibly additional requirements coming from the new services proposed in WP2, and the needs of the Multi-Hub network coming from WP 4. The new system is arbitrarily referred to as the 'e-2' system until a name or acronym can be found.

B.1.6.3 Joint Research Activities: Work package list

Work package No	Work package title	Type of activity	Lead partic no.	Lead partic. short name	Person-months	Start month	End month
7	New Service-based Architecture	RTD	6	CU	58	M1	M36
	TOTAL			1	58		

B.1.6.4 Joint Research Activities: Work package descriptions

Work package 7: New Service-based Architecture

Work package number	WP 7	Start date or starting event:		Month 1			
Work package title	New Service-based Architecture						
Activity type[20]	RTD						
Participant number	6						
Participant short name	CU						
Person-months perparticipant	58						

Objectives
i) To replace the existing distributed model with a purpose-built service-based Architecture.
ii) To build the new Architecture around the 'open ecosystem' model in which all distributed components are implemented as open services, so that user or partner communities can integrate both low- and high-level services into their own e-Science communities.
iii) To build the new Architecture to reflect the recent enlargement and raised objectives of the Catalogue of Life community. This will include replication from one distributed hub to multiple distributed hubs, for the need for the enlarged concept to be managed by an overall management system, and for the management to become a 24/7 operation involving several sites around the world.

Description of work (the system to be implemented under this new Architecture is here referred to as 'e-2'.
Task 7.1 To assemble the requirements and specification for the new Architecture and the new implementation of the 'e-2' system.
Task 7.2 To design the e-2 implementation with an open service-based architecture.
Task 7.3 To establish common communication protocols, data model, and management framework suitable for e-2.
Task 7.4 To implement as 'proof-of-concept' demonstration of the new e-2 architecture, with selected test and management tools.
Task 7.5 To complete and install a test implementation for e-2 with three GSDs as test-beds.
Task 7.6 To respond to testing with an enhanced prototype for e-2
Task 7.7 Continue to monitor the new architecture, and to assist the Systems Design Team.

Deliverables
D7.1 mth 12 Report on Requirements, Specification and Design of new e-2 architecture. [14 PeM](R/PP)
D7.2 mth 12 'Proof of Concept' demonstrated for the e-2 architecture.[14 PeM](D/RE)
D7.3 mth 24 Report on the implementation and testing of operative prototypes of the new architecture. [30 PeM](R/RE)

Milestones
M7.1 mth 22 e-2 Prototype handed to Software Services (WP 6) for development of production version.

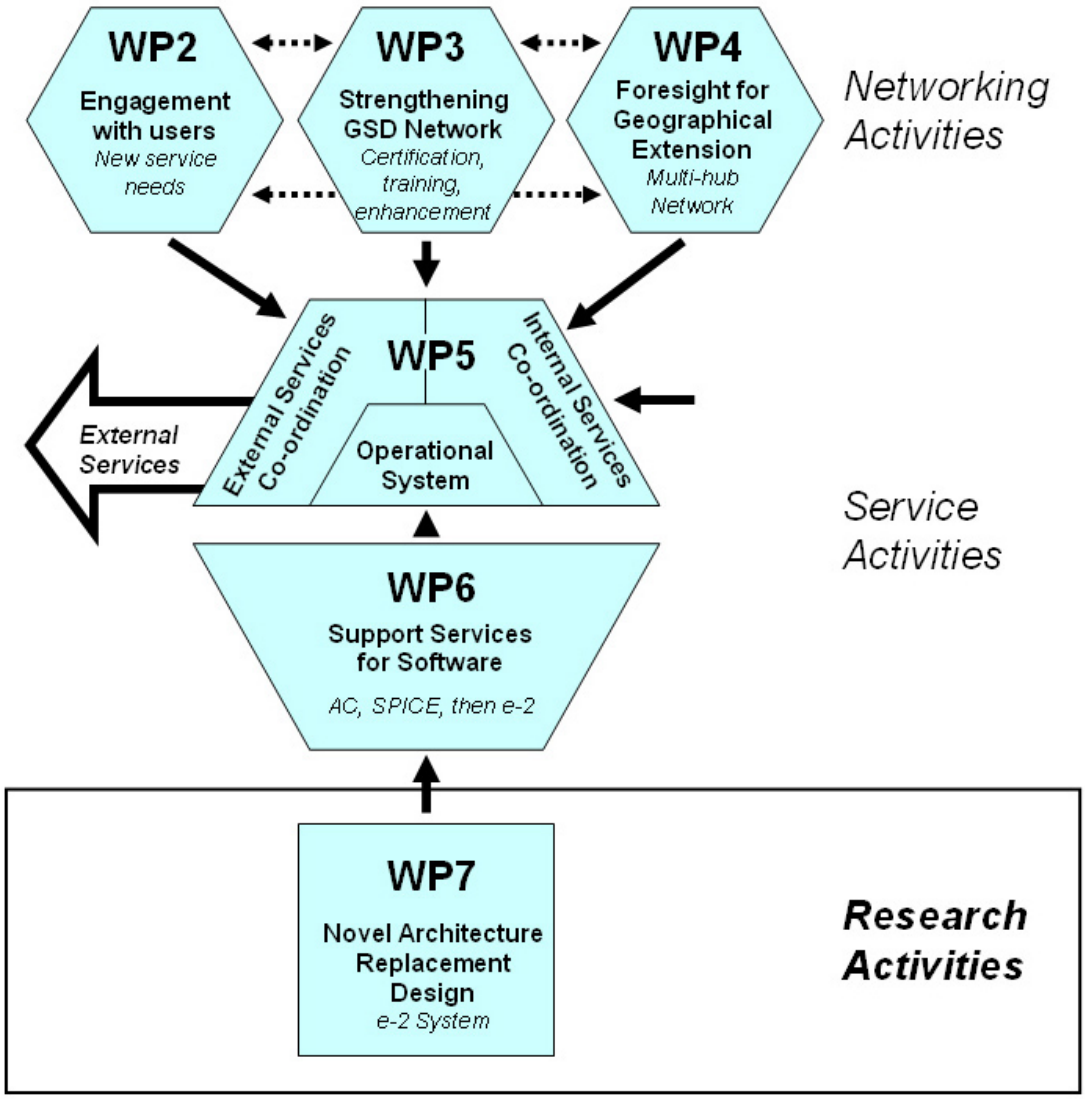
B.1.6.5 Joint Research Activities: Efforts for the full duration of the project

Partic. no.	Partic. short name	WP 7	Total person months
6	**CU**	**58**	**58**
Total		**58**	**58**

B.1.6.6 Joint Research Activities: Milestones

number	Milestone name	Work package(s) involved	Expected date	Means of verification
M7.1	e-2 Prototype handed to Software Services (WP 6) for development of production version	WP 7 & WP 6	mth 12	PPR&PO

PPR&PO – Achievement of each Milestone will be reported in the 12-monthly Project Periodic Report, and notified to the Programme Officer.



Relationships between Workpackages

B.1.7 Project Effort Form

Project Effort Form 2 - indicative efforts per activity type per beneficiary

Project number (acronym):

Activity Type	UR	BGCI	MNHN	NARODNI	ETI	CU	NNM	SP2000	VLIZ	MfN, Berlin	NHM	TOTAL ACTIVITIES
Joint Research (RTD) activities												
Total 'JRAs'						58						58
Networking (Coordination) activities												
Total 'NAs'	8	33	46	32			4		4	4	6	137
Support activities												
Total 'Support'												
Consortium Management activities												
Total 'Management'	26											26
Service (Other) activities												
Total 'SAs'	110		5		68			36	2			221
TOTAL BENEFICIARIES	144	33	51	32	68	58	4	36	6	4	6	442

Project Effort Form 2 - indicative efforts per activity type per beneficiary

Project number (acronym):

Activity Type	CABI	RBGK	UNVIE	UU	NUI-GALWAY	IFM-GEOMAR	DKFZ	MIZ-PAS	LAND OOE	NMGW	TOTAL ACTIVITIES
Joint Research (RTD) activities											
Total 'JRAs'											
Networking (Coordination) activities											
Total 'NAs'	4	4	4	4	4	4	4	4	4	4	40
Support activities											
Total 'Support'											
Consortium Management activities											
Total 'Management'											
Service (Other) activities											
Total 'SAs'											
TOTAL BENEFICIARIES	4	4	4	4	4	4	4	4	4	4	40

Project Effort Form 2 - indicative efforts per activity type per beneficiary

Project number (acronym):

Activity Type	IRD	CSIC	TSJ BVBA	SNSB, Munich	UNIPD	UVA	RBINS	University of Oxford	WU	RBGE	UCPH	TOTAL ACTIVITIES
Joint Research (RTD) activities												
Total 'JRAs'												
Networking (Coordination) activities												
Total 'NAs'	4	8	4	8	4	4	4	4	4	4	4	52
Support activities												
Total 'Support'												
Consortium Management activities												
Total 'Management'												
Service (Other) activities												
Total 'SAs'												
TOTAL BENEFICIARIES	4	8	4	8	4	4	4	4	4	4	4	52

Project Effort Form 2 - indicative efforts per activity type per beneficiary

Project number (acronym):

Activity Type	ITZN	CAS	CRIA	National Museum of Natural History	CSIRO	Landcare Research	TOTAL ACTIVITIES
Joint Research (RTD) activities							
Total 'JRAs'							
Networking (Coordination) activities							
Total 'NAs'	3	4	4	0	0	0	11
Support activities							
Total 'Support'							
Consortium Management activities							
Total 'Management'							
Service (Other) activities							
Total 'SAs'			8				8
TOTAL BENEFICIARIES	3	4	12	0	0	0	19

Project Effort Form 2 - indicative efforts per activity type per beneficiary

Project number (acronym):

Activity Type	Total Page 1	Total Page 2	Total Page 3	Total Page 4	GRAND TOTAL
Joint Research (RTD) activities					
Total 'JRAs'	58				58
Networking (Coordination) activities					
Total 'NAs'	137	40	52	11	240
Support activities					
Total 'Support'					
Consortium Management activities					
Total 'Management'	26				26
Service (Other) activities					
Total 'SAs'	221			8	229
GRAND TOTAL BENEFICIARIES	442	40	52	19	553

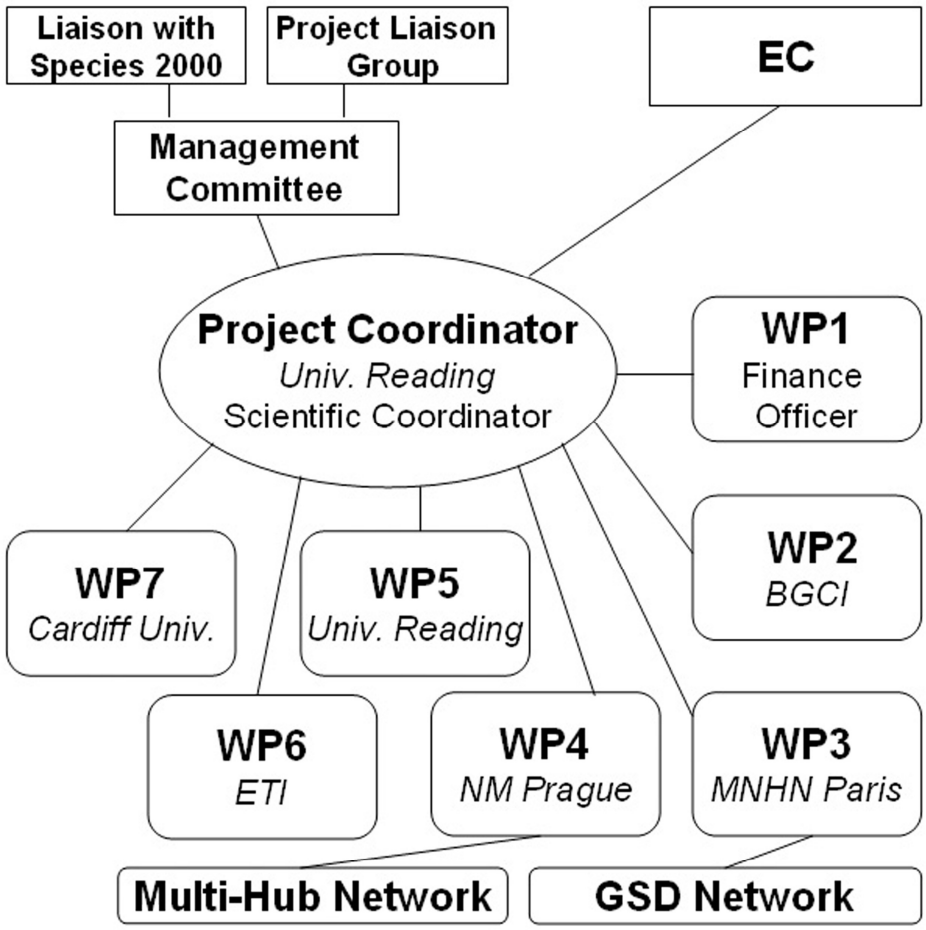
Section B2: Implementation

B.2.1 Management Structure and Procedures

Management Structure
4D4Life will be organised from the Project Coordinator's institute, the University of Reading (UREAD). The organisation of the consortium will consist of:

i) Project Coordinator
ii) Scientific Coordinator (full time employee of WP 5)
iii) Finance Officer (part time employee of WP 1)
iv) Work Package Leaders (of WP2, WP 3, WP 4, WP 5, WP 6 and WP 7)
v) Management Committee
vi) Project Liaison Group

The management structure of 4D4Life is illustrated below. The structure is simple and straightforward, and well suited to the size and scope of the project. In addition to the main members of the work packages, the consortium contains the GSD Network members and the Regional Multi-Hub Network members.



4D4Life Management Structure

In addition to the management structures shown, there will be four project meetings at which all members of the consortium will meet to review the project and its progress:

a) Month 5 Project Start-up Meeting with start-up meetings of work packages
b) Month 11 Project Meeting and associated workshops of work packages
c) Month 24 Project Meeting and associated workshops of work packages
d) About Month 35/36 Showcase Meeting with a wider public audience

The responsibilities of the members in the organisation are as follows:

The Project Coordinator (F. Bisby) has the overall responsibility for the project. He is in charge of organising and running the 4D4Life networking and coordination. The Project Coordinator oversees the partner's progress according to the work package descriptions, evaluates that progress, the meeting of milestone and deliverable dates, creation of final reports, motivates and monitors collaboration across institutions and disciplines, encourages joint publications, is responsible for the project and Liaison Group Meetings and their reporting, and organises resolution of any administrative, scientific or technical conflict that may occur. The Project Coordinator is also responsible for WP1 to manage the consortium and the financial contribution, and is responsible for project communications with users and with programmes across the biodiversity landscape through the project web-site, through the Project Liaison Group, through representing the project at European and international meetings and through the Showcase Meeting at the end of the project. The Project Coordinator has responsibility for all forms of reporting to the European Commission, for promoting gender equality in the project, and for conflict resolution should it occur.

The Scientific Coordinator (located at University of Reading, employed under WP5) will assist the project Coordinator with internal and external communication (including the use of one or more wikis, and realisation of the www.4D4Life.eu web-site), the organisation of the Project Meetings, and the recording of scientific reporting and communication with the Commission. This will occupy half that person's time, the other half being work under WP5 on designing and realising services.

The Financial Officer (located at University of Reading, employed part-time under WP1) is responsible for the claiming, distribution, accounting and reporting of funds distributed from the community financial contribution from the Commission.

The Management Committee consists of the work package leaders and is chaired by the Project Coordinator. The Management Committee follows progress of the project on a monthly basis via telephone conferencing, and normally two meetings per year, including one at each Project Meeting. Any deviations from the work plan, as well as methods for dealing with these deviations, will be discussed, and action agreed at the project level and/or in communication with the Commission. This will also be the place to resolve technical conflicts that may arise between work packages.

Agenda items will be circulated at least 7 days ahead of the meeting, decisions taken by majority vote, and minutes recorded and circulated within one month.

Within the Management Committee two 'teams' will work with the project Co-ordinator and the Scientific Co-ordinator to ensure cross-workpackage co-ordination in two particularly important areas:
i) Services Team
Sara Oldfield (leader of WP 2) will chair the Services Team that will co-ordinate the selection and delivery of the array of new services – involving WP 2, WP 3, WP 5 and WP 6.
ii) System Design Team
Alex Hardisty (co-leader of WP 7) will chair the Systems Design Team that will co-ordinate the design of systems across the project – involving WP 4, WP 5, WP 6, WP 7, as well as work with the Services Team.

The Project Liaison Group is a liaison group invited to meet with the Management Committee on just a few occasions to maintain awareness and liaison between 4D4Life and the ring of related but separate projects within the taxonomic and biodiversity informatics landscape, particularly in Europe. The following have initially been invited to join this group:

Professor Simon Tillier, EC EDIT
Dr David Remsen, GBIF-ECAT
Professor David Patterson, EoL Informatics
Dr Yde de Jong, EC PESI
Professor Stefan Claesson, CETAF
Professor Walter Berendsohn, EC BioCase

(The Group may meet in association with certain of the 4D4Life Project Meetings such as the Showcase meeting, but also possibly at other venues including those organised, for example, by EDIT, or GBIF.)

Liaison with Species 2000. 4D4Life will become a significant component of the wider Species 2000 Programme, and as such will need to work closely both with the Species 2000 Global Team that coordinates scientific policy for the programme world-wide, and with other funded projects within the wider programme. This is needed because there are projects and contributors to the wider programme that are not based in Europe, and in which the partners are not part of 4D4Life. Species 2000 has indicated that it will invite the Project Coordinator (who is already a member) and one other member of the Management Committee to join the Global Team and act as regular conduit for coordination and cooperation between the wider programme and 4D4Life. Communication with the Global Team is important for this project to reach its full potential, and it is possible that two of the Global Team meetings can be arranged to coincide with two of the 4D4Life Project Meetings. An outline of the 4D4Life proposal was endorsed by the Global Team at its April 2008 meeting.

A similar arrangement worked well between the Global Team and the FP 5 Eurocat project in 2003 - 2006, with back-to-back and joint meetings held at the Malta and Stockholm project meetings of that project.

B.2.2 Beneficiaries

1. The University of Reading, Reading, (UR)

The Centre for Plant Diversity and Systematics, part of the School of Biological Sciences at the University of Reading, is known world-wide for its leadership of flagship research programmes in taxonomy - Flora Europaea, the UNEP Global Biodiversity Assessment, the Families of Flowering Plants, ILDIS and most recently the Species 2000 and ITIS Catalogue of Life, BiodiversityWorld (BDWorld). The Centre has a proven track record in coordinating EC FP5 project Species 2000 europa (EuroCat) & Euro+Med, and BBSRC projects (SPICE, LITCHI, BDWorld). Participating in others EC projects: ERMS, Fauna Europaea, BioCASE, ENBI, LifeWatch.

The Centre will provide management infrastructure for operation of four GSDs in this project including ILDIS and the Conifer database.

ILDIS World Database of Legumes by YR Roskov, FA Bisby, JL Zarucchi, BD Schrire & RJ White (eds). Taxonomic database of world legumes: Plantae – Magnoliophyta – Magnoliopsida – Fabales – Fabaceae/Leguminosae; 19,937 spp.

Conifer Database by A Farjon. Taxonomic database of world conifers: Plantae – Pinophyta; 631 spp.

Tasks: Consortium management and finances in WP1.

Coordination of WP5 "Scientific services of the Catalogue of Life". Enhance the public scientific infrastructure of the Catalogue of Life, including an array of novel public and electronic services. (Design, test and implement (i) enhancements to the present public services, (ii) novel services at the technical level, (iii) novel educational and popular services). Enhance and operate the annual production service for the Annual Checklist, including unification of the admission gate for the Annual and Dynamic Checklist products. Enhance, secure and operate the editorial, scrutiny and quality control services for the unified production route. Secure and operate the ongoing systems management for the internal and public services of the whole programme. Establish an enhanced business plan for sustainability, and undertake associated financial and legal engineering. To undertake the scientific co-ordination of the whole project (co-ordination of scientific and IT components of the entire project, including the interaction of this central workpackage with all the other workpackages).

ILDIS and Conifer databases will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Prof. Dr. Frank Bisby (Project Co-ordinator, WP5 leader), played a significant role in establishing biodiversity informatics, as a lecturer and then reader at the University of Southampton, and now as Professor of Botany and Director of the Centre for Plant Diversity & Systematics and member of the Informatics Research Centre at the University of Reading. He is a founding member of the Biodiversity Information Standards organisation (TDWG), instigator of one of the first global species databases, ILDIS, and leader of the Species 2000 organisation since its launch in 1996, as well as co-ordinator of EC FP5 EuroCat project, member/workpackage leader in other EC projects such as ERMS, Euro+Med, ENBI, and EDIT. As Executive Director of Species 2000 he works with what is now an extensive network of taxonomic database hubs around the world from China and New Zealand to Brazil

and N. America, in projects to approach completion of known organism coverage in the Catalogue of Life, and to work with partner organisations on the Global Names Architecture.

Dr Yury Roskov, plant taxonomist, Catalogue of Life content manager. Executive editor of the Annual Checklist editions 2006, 2007 & 2008. Co-director and responsible manager of ILDIS database.

Dr Aljos Farjon, world leading taxonomist of conifers, author and custodian of the Conifer database.

2. Botanic Gardens Conservation International (BGCI)

Botanic Gardens Conservation International (BGCI) is a UK Registered Company limited by guarantee (Company No. 04673175) and a Registered Charity (No. 1098834). It operates as a membership organisation that represents and unites the skills and expertise of botanic gardens globally, and works with partners, to conserve the world's imperilled flora. Linking more than 800 botanic gardens in 118 countries BGCI forms the world's largest plant conservation network. The HQ of BGCI is based in the UK at Royal Botanic Gardens, Kew. It has offices in South China Botanical Garden, Chicago Botanic Garden and Singapore Botanic Garden. A staff member is based in the IUCN East Africa Office, and also leads a regional bioinformatics project for East Africa. BGCI also has national representatives in Colombia, Germany, the Netherlands, Russia and Spain.

BGCI provides the Secretariat for the Global Partnership for Plant Conservation (GPPC) mandated by the CBD to facilitate the implementation of the Global Strategy for Plant Conservation. The GPPC currently has over 30 member in a lose affiliation including IUCN, FAO, BioNet International, Bioversity International and other major providers and users of biodiversity information.

BGCI is a member of GBIF, a member of the Encyclopedia of Life Institutional Council and a member of the EoL Species Sites Group for plants. BGCI has good contacts with the corporate sector and is currently in discussion with or developing partnerships with companies in the global banking, pharmaceutical, agricultural, energy and transport industries.

Tasks:
Coordination of WP2 "Engagement with scientific users and partners". To engage with (i) the major global biodiversity portals and biodiversity organisations, (ii) national portals and programmes, NGOs, and higher education institutes, (iii) individuals who make serious use of the Catalogue of Life Services, and review current and potential uses and enhancements in these groups. To locate and engage with the array of commercial organisations that either do use the Catalogue of Life Services, or may be considered likely to do so. To establish an outreach and promotional programme that will disseminate the infrastructure services. To establish the Catalogue of Life as one of the major international biodiversity programmes that can take its place as a partner and provider at the global scale alongside GBIF, EoL, BHL, IUCN, GenBank and CBOL. Working with the users and partner organisations, to specify, and subsequently test a range of enhanced and novel scientific services.

Sara Oldfield, Secretary General of BGCI with responsibility for overall management and development of the organization. She has 30 years experience of working in the biodiversity sector, including data analysis, policy development and management of a wide range of conservation initiatives..

Suzanne Sharrock, Global Programmes Director for BGCI. She manages BGCI's climate change and human well-being programmes. Suzanne also works on policy issues connected with the Global Strategy for Plant Conservation.

Meirion Jones, BGCI's Head of Information Management. He maintains and develops the web-based PlantSearch and GardenSearch Databases as part of BGCI's overall information strategy.

3. Muséum National d'Histoire Naturelle, Paris (MNHN)

MNHN is a scientific institute dealing with biodiversity conservation. It is the scientific advisor of the French ministry for environment. Its areas of activities are research, education and training managements and enrichment of its collection (almost 80 million specimens), expertise and diffusion of scientific knowledge. The MNHN plays a key role in the organisation of the French information system on biodiversity and landscape, being the scientific coordinator of this national project. It also hosts the European thematic centre for biodiversity of the European Environment Agency. In addition to international research collaboration at the individual level, MNHN plays also several important international roles. It has been and is heavily involved in various European programmes related to biodiversity:

- The European Topic Centre on Nature Protection and Biodiversity is hosted by MNHN.
- The French national GBIF, GTI and CBD nodes, the French Focal Point for the EU-funded BioCASE and PESI projects are located in the MNHN.
- Leading partner / work package leader in EC ParSyst, ColParSyst, Fauna Europaea, ENBI and EuroCAT, Synthesys 1, EDIT, MarBEF.

MNHN is a home organization for four GSDs:

FLOW: Fulgoromorpha Lists On the Web by T Bourgoin. Taxonomic database of world planthoppers: Animalia – Arthropoda – Insecta – Hemiptera – infraorder Fulgoromorpha; 9,209 spp.

COOL: Cercopoidea Organised On Line by A Soulier-Perkins. Taxonomic database of world froghoppers: Animalia – Arthropoda – Insecta – Hemiptera – superfamily Cercopoidea; 2,086 spp.

Psyl'list: Psylloidea database by D Ouvrard. Taxonomic database of world jumping plant-lice: Animalia – Arthropoda – Insecta – Hemiptera – superfamily Psylloidea; 1,756 spp.

CLEMAM by P Bouchet. Taxonomic database of molluscs. Animalia – Mollusca.

Tasks:
Coordination of WP3 "Strengthening the GSD Partner Network: a Virtual Community". Assessement of participating GSDs. Network consultation and workshop with GSD custodians towards Catalogue of Life standardisation. To accelerate and enhance GSD completeness and data quality. To coordinate set of Pilot Projects for enhancement of the Catalogue of Life Services for Users.

FLOW, COOL and Psyl'list databases will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Prof. Thierry Bourgoin, Deputy Director of the MNHN Collections, coordinator of Species 2000 europa database network (EuroCat), member of the Global Species 2000 Team and Taxonomy group, leading taxonomist of planthoppers, author and manager of FLOW database.

Dr Philippe Bouchet, taxonomist of molluscs, editor of CLEMAM database.

Dr Adeline Soulier-Perkins, taxonomist of froghoppers, author of COOL database.
Dr David Ouvrard, taxonomist of jumping plant-lice, author of Psyl'list database.

4. Narodni Museum, Prague (NARODNI)

The Museum is a public scientific institution which systematically enriches its collections consisting of objects of natural and historical sciences all over the world, with particular interest to the Czech Republic. It conducts research in various fields of natural and historical sciences and has a large exhibit activity. The National Museum is the most distinguished and the largest museum in the Czech Republic. It consists of five professional institutions: Natural History Museum, Historical Museum, The Náprstek Museum of Asian, African and American Cultures, Czech Museum of Music, National Museum Library. These are collections of mineralogy, palaeontology, mycology, botany, entomology, zoology, anthropology, archaeology. From the very beginning the collections have been regarded also as a treasury of the most important monuments of the Czech national history. They are located in more than 15 public museums and exhibition halls all over the country. The museum publishes 12 scientific journals and numerous other non periodical publications.

Tasks:
Coordination of WP4 "Foresight study for geographical extension: the Global Multi-Hub Network". Network consultations and workshop discussions among regional hub organisers (to bring together for the first time leaders of the various international centres from around the world that have committed to create Regional Hubs in the Sp2000 Multi-Hub Network, so that they can work together in Networking, Discussion and Design). Pilot Projects for prototyping and testing Multi-Hub Linkage tools, installing and testing Multi-Hub linkages using tools from the new e-2 Architecture of WP 7.

Dr Jiří Kvaček, Head of the Department of Palaeontology, leader of Science programmes. His scientific work is focused on palaeobotany of gymnosperms and angiosperms. Since 1987 he is working for the National Museum in Prague in various functions. From 2005 to 2008 he was science director of the National Museum. He was involved in management of several scientific projects in palaeontology, a project of Scientists night and three exhibitions. In 2007 he took part in the management team responsible for the development of the new permanent exhibition of the National Museum, Prague.

5. Stichting Expertisecentrum voor Taxonomische Identificatie, Amsterdam (ETI)

ETI BioInformatics is a project-driven organization with ample experience in biodiversity informatics and service development. It promotes access to biodiversity information and specializes in the development and implementation of professional, sustainable information infrastructures that facilitate standardization, integration, interoperability, accessibility and efficient dissemination of biodiversity information. ETI expertise concerns:

- database implementation, information management tools, user interfaces
- interoperability solutions, data merging, database wrapping and webservices
- web portals, dynamic websites, mobile applications
- computer aided identification, diagnostic systems, e-learning applications
- e-content management, product development, exploitation planning and marketing

With and for a global network of taxonomists and biodiversity specialists ETI developed tools and a mechanism for the World Biodiversity Database: a series of over 100 e-monographs published on CD-ROMs and the web. ETI is one of the founders of Species 2000 and closely involved with the technical aspects of producing the CoL Annual Checklist. ETI participates in a range of EC and global projects and has an excellent track record with the timely delivery of professional ICT applications, e.g.: in ENBI, EuroCat, LifeWatch, MARBEF, and KeyToNature. Other recent partners include: Netherlands GBIF Node, Tanzanian GBIF Node, US Dept of Agriculture, NL Ministry of Agriculture Nature Conservation and Food, National Herbarium NL, Zoological Museum Amsterdam. ETI is a member of GBIF, EMBNet and participates in TDWG.

Tasks:
Coordination of WP6 "Support Services for Software Deployment".

(i) Public applications Services. Provide stability and functional enhancements to the present user interfaces, search facilities and web-services, using input from WP 5 & 4. Specification testing, blackbox usability and stability testing involving all partners. Test and provide additional or variant software for the novel technical, education and popular services, following selection and prioritization by WP5. Improve documentation and version control.

(ii) Middleware for Internal Services. Hold design team meeting with WP7 and WP5 to determine the middleware strategy and changes. Define a new central base schema to assist unification of all assembly and publication tools, based on current schemas in AC en DC; and produce a Global Species Database (GSD) builder tool (batch functions) and GSD editor tool (atom by atom editing) based on the new schema pre-fitted for uploading to the SPICE cache (SPICE is a distributed-computing engine that runs the Dynamic Checklist). (with Paris & VLIZ provides prototype, ETI develops production version). Prepare robust components for the SPICE 5.1 interoperability system without changing the fundamental communication protocol, enabling steps towards unification of the AC and DC production streams; produce control components for assembled data. Produce a universal GSD wrapper [kit] for data providers to link GSDs to the SPICE system (with CRIA). Alter the SPICE DC cache to adopt the new central base schema, Alter the GSD editing and inspection tools that currently work with the AC to work with the new base schema and make these available for editorial staff to edit and inspect the cache an intermediate state datasets in the base schema. Produce a stable production tool for linking GSDs to the SPICE system using experiences from task 6.8 , including a new version of the Metadatabase software and a tool for AC export & conversion from the DC cache for publication on CD. Prepare a version of the SPICE 5.1 system for supporting linkages, cross-maps and enhanced interface for use in the Multi-Hub Network. Prepare documentation and version control system for all SPICE components. (with CARDIFF). Prepare a production version of the components in the new e-2 service-based system from WP7

Dr Peter H. Schalk, ETI Managing Director. Specialization: HR and project management, use-user analysis, data quality, information management, knowledge product development. Chairman of the Board of Directors of Species 2000, Chairman of the Board of Directors of ETI Information Services Ltd, Chairman of the Netherlands Foundation for Nature Research, Vice-chair of the GBIF Budget Committee.

Wouter Addink, ICT project-leader, system developer.

Dr Maarten Schermer, system developer.

Dr Ruud Altenburg, **Dr Edwin van Spronsen, Gideon Gijswijt** application programmers.

6. Cardiff University, Cardiff (CU)

The Cardiff School of Computer Science (COMSC) undertakes internationally ranked research in areas related to the focus of the present proposal through its Knowledge and Information Systems (KIS) research group. The Welsh e-Science Centre (WeSC) within COMSC complements and works with KIS. It is one of the UK's established e-Science centres, set up in 2001 as part of the national programme for e-Science/Grid/e-Infrastructure. The School brings relevant knowledge and experience in the areas indicated below through its participation in international and national funded projects. Funders include: EC, UK Research Councils, and JISC.

- ***e-Science for Biodiversity Informatics:*** Building an e-Science platform for biodiversity research, addressing relevant research problems in workflow system integration and usability, security, interoperability and integration of heterogeneous, distributed and legacy systems and databases, in projects such as BiodiversityWorld, GRAB, BioDA, ASMIMA and Ark 2010. TDWG-sponsored SPICE TIP project researched and prototyped the deployment of LSIDs. Executive partner for technical planning in the EC funded project for preparing the LifeWatch infrastructure.
- ***Biodiversity Informatics:*** Building distributed information systems, including the Catalogue of Life, addressing research problems in architectures for interoperability, quality control and standards for data and software and how to work with different quality levels, in projects such as ENBI, EuroCat, SPICE and LITCHI.
- ***e-Science:*** Partner operating node in the UK National Grid Service; Configuration and management of Grid middleware, build and operation of Service-oriented (SOA) Grid test-beds, in projects such as RTGrid, GECEM and Grid4Wales. International Grid and e-Science networking activities; links to Coregrid, Gridlab and EGEE. Addressing research problems in Virtual Organisations, Problem Solving Environments (Portals), use of varying quality data, workflows and mechanisms for provenance, in projects such as CONOISE, COVITE, GECEM; PROVENANCE, PASOA, Triana, WoSE, Gridlab and GridOneD.

Tasks:
Coordination of WP7 "Novel e-Science Service-based Architecture". To assemble the requirements and specification for the new Architecture and the new implementation of the 'e-2' system. To design the e-2 implementation with an open service-based architecture. To establish common communication protocols, data model, and management framework suitable for e-2. To implement e-2 test tools, and e-2 management tools. To create and install a test implementation for e-2. To respond to testing with an enhanced prototype for e-2

Alex Hardisty, Grid Centre Manager of the Welsh e-Science Centre and a technical planner for EC FP7 LifeWatch. Chair of the Steering Committee for the Community Engagement and Support strand of the JISC e-Infrastructure programme.

Richard White, Lecturer in Bioinformatics, the convenor of the Species 2000 Information Systems Group.

Andrew Jones, Senior Lecturer, a member the Species 2000 IS group.

7. Stichting National Natuurhistorisch Museum Naturalis, Leiden (NNM)

Naturalis is a home organisation for Odonata database. Member of Species 2000. Participated in EC FP5 EuroCat project.

Odonata: Catalogue of the Odonata of the World by J van Tol. Taxonomic database of world dragonflies and damselflies: Animalia – Arthropoda – Insecta – Odonata; 5,321 spp.

Tasks: Odonata database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality. Prof. van Tol will also lead discussion and prototyping of nomenclator services to GSDs in WP 3.

Prof. Jan van Tol, leading taxonomist of dragonflies and author of Odonata database, also nomenclatural specialist and Commissioner of the International Commission on Zoological Nomenclature.

8. Species 2000 (Sp2000)

Species 2000 is a Network organisation that creates an index of the world's known organisms. It is a not-for-profit company limited by guarantee registered in England (Company No. 3479405). The Species 2000 distributed model synthesises the index from sectors supplied by taxonomic databases across Europe and around the world, many from the major European institutions of CETAF. The programme reached production scale as an EC scientific infrastructure under the FP5 EuroCat project, and as a member of EDIT celebrated coverage of one million species in 2007. Its Catalogue of Life is a global service (www.catalogueoflife.org) recognised by the UN Convention on Biological Diversity, and presently comprising a synonymic species checklist of 1.1 million plants, animals, fungi and micro-organisms, about 2.5 million names, and a comprehensive taxonomic hierarchy. It provides a taxonomic backbone for global biodiversity portals, such as EoL and GBIF, for about 40 national portals worldwide, and is used by scientists in 79 countries. It contributes content and the taxonomic hierarchy used by the uBio taxonomically intelligent tool in the BHL programme.

Sp2000 operates a Global Secretariat at University of Reading (partner 1) and an office within the WorldFish Centre at Los Banos , Philippines.

Sp2000 provides the residual legal body for the Species 2000 Global Programme, holding its IPR, copyright, domain names, access licences, MOU's, taking responsibility for continuity between major projects and providing the ongoing governance of the global programme. It is structurally a federation, owned and governed by the participants that become its formal members, such as those other members of the consortium listed as 'Member of Species 2000'.

Tasks: Participation in WP2, WP3, WP4 and WP5 activities, plus international liaison through the Global Species 2000 Team, and ongoing attention to the business model and IPR issues.

Prof. Dr. Frank Bisby, Executive Director of Species 2000, works with what is now an extensive network of taxonomic database hubs around the world from China and New Zealand to Brazil and N. America, in projects to approach completion of known organism coverage in the Catalogue of Life, and to work with partner organisations on the Global Names Architecture.

Dr Yuri Roskov, Dr Yuri Roskov, Species 2000 Content Manager.
Ms Rebeca Mann, Species 2000 Secretariat.
Dr Nicolas Bailly, Species 2000 Philippines Office Manager.
Ms Luvie Paglinawan, Species 2000 database compiler, Philippines Office.

9. Vlaams Instituut voor de Zee, Ostend (VLIZ)

VLIZ is a home organisation for WoRMS database.

WoRMS: World Register of Marine Species, compile by the consortium of marine biologists. Taxonomic database of marine organisms of the world, 122,500 spp.

Tasks: WoRMS database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality. IT specialists of VLIZ will take active part in software development in WP 6 (GSD builder and editor based on the new schema).

Ward Appeltans, WoRMS data manager.
Bart Vanhoorne, Aphia system manager and programmer.

The Global Species Database (GSD) Network

10. Museum für Naturkunde - Leibniz Institut für Evolutions – und Biodiversitätsforschung an der Humboldt-Universität zu Berlin (MfN, Berlin)

Museum für Naturkunde is a home organisation for GloBIS database. Participated in EC FP5 EuroCat project.

GloBIS/GART: Global Butterfly Information System by C Häuser, J Holstein & A Steiner (eds). Taxonomic database of world swallowtail butterflies: Animalia – Arthropoda – Insecta – Lepidoptera – Papilionidae; 553 spp.

Tasks: GloBIS database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Christoph Häuser & **A Steiner**, specialists in taxonomy of Lepidoptera and authors of GloBIS database.

11. Natural History Museum, London (NHM)

NHM is a home organisation for 5 databases: LepIndex, UCD, Tineidea NHM, BSF and PSF. Member of Species 2000. Participated in EC FP5 EuroCat project.

BSF: Blattodea Species File Online by GW Beccaloni. Taxonomic database of world cockroaches: Animalia – Arthropoda – Insecta – Blattodea; 4,430 spp.

LepIndex: The Global Lepidoptera Names Index by GW Beccaloni, MJ Scoble, GS Robinson & B Pitkin. Digitised version of the index card archive to the scientific names of the living and fossil butterflies and moths of the world at the Natural History Museum: Animalia – Arthropoda – Insecta – Lepidoptera; 250,203 spp.

PSF: Phasmida Species File by PD Brock. Taxonomic database of world stick and leaf insects: Animalia – Arthropoda – Insecta – Phasmida; 2,821 spp.

Tineidae NHM: Global Taxonomic Database of Tineidae by GS Robinson. Taxonomic database of world clothes-moths: Animalia – Arthropoda – Insecta – Lepidoptera – Tineoidea – Tineidae; 2,340 spp.

UCD: Universal Chalcidoidea Database by J Noyes & A Polaszek. Taxonomic database of world chalcidoid wasps: Animalia – Arthropoda – Insecta – Hymenoptera – superfamilies Chalcidoidea & Serphitoidea; 19,847 spp.

Tasks: LepIndex, UCD, Tineidea NHM, BSF and PSF databases will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Christopher Lyal, Department of Entomology, leading taxonomist of weevils, co-author of WTaxa database.

Dr George Beccaloni, taxonomist of cockroaches, content manager of LepIndex and author of BSF database.

Dr Andrew Polaszek, taxonomist of chalcidoid wasps and manager of UCD database.

12. CAB International (CABI)

CABI is a home organisation for Species Fungorum database. Member of Species 2000. Participated in EC FP5 EuroCat project.

Species Fungorum by PM Kirk. Taxonomic database of various fungal taxa: Fingi (diverse taxa), Chromista – Hyphochytriomycota, Labyrinthulomycota, Oomycota; Protozoa – Acrasiomycota, Dictyosteliomycota, Plasmodiophoromycota; 30,085 spp.

Tasks: Species Fungorum database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Paul Kirk, editor of Species Fungorum database.

13. Royal Botanic Gardens Kew (RBGK)

RBG Kew is a home organisation for World Checklist of Selected Plant Families database. Member of Species 2000. Participated in EC FP5 EuroCat project.

World Checklist of Selected Plant Families by R Govaerts (ed). Taxonomic database of 94 plant families: Plantae – Ginkgophyta, Gnetophyta, Magnoliophyta – Liliopsida (except Iridaceae, Juncaceae, Poaceae & Potamogetonaceae) & Magnoliopsida – Araliaceae, Betulaceae, Euphorbiaceae, Fagaceae, Garryaceae, Magnoliaceae, Nothofagaceae, Pandaceae, Phyllanthaceae, Picrodendraceae, Putranjivaceae, Sapotaceae & Ticodendraceae; 83,298 spp.

Tasks: Kew database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Alan Paton, Assistant Keeper of the Herbarium, official representative of RBG Kew in Species 2000.

Dr Rafael Govaerts, editor of Kew database.

Dr Robert Allkin, Information Projects Manager.

14. Universitaet Wien (UNIVIE)

The University is one of the home organisations for AnnonBase.

AnnonBase: World Species List of Annonaceae by H Rainer & LW Chatrou (eds). Taxonomic database of world custard apples: Plantae – Magnoliophyta – Magnoliopsida – Magnoliales – Annonaceae; 2,235 spp.

Tasks: AnnonBase will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Heimo Rainer, editor and IT manager of AnnonBase.

15. Universiteit Utrecht (UU)

The University is a home organisation for TicksBase. Participated in EC FP5 EuroCat project.

TicksBase by AM Nijhof, AA Guglielmone & IG Horak. Taxonomic database of world ticks: Animalia – Arthropoda – Arachnida – Parasitiformes (Acari – suborder Ixodida); 868 spp.

Tasks: TicksBase will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Frans Jongejan & **Dr Ard Nijhof**, authors of TicksBase.

16. National University of Ireland, Galway (NUI Galway)

The University is a home organisation for AlgaeBase. Member of Species 2000. Participated in EC FP5 EuroCat project.

AlgaeBase by M Guiry . Biodiversity database of world algae: Bacteria – Cyanobacteria; Chromista – Cryptophyta, Haptophyta, Sagenista; Protozoa – Acritarcha, Cercozoa, Dinophyta, Euglenozoa, Flagellata, Parabasalia; Plantae – Bacillariophyta, Chlorophyta, Cyanidiophyta, Glaucophyta, Prasinophyta, Rhodophyta; 30,487 spp.

Tasks: AlgaeBase will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Prof. Michael Guiry, taxonomists of seaweeds, author of AlgaeBase.

17. Leibniz-Institut Fuer Meereswissenschaften an der Universitaet Kiel (IFM-GEOMAR)

IFM-GEOMAR is a home organisation for FishBase. Participated in EC FP5 EuroCat project.

FishBase by R Froese & D Pauly (eds). Biodiversity database of world fishes: Animalia – Chordata – Actinopterygii, Cephalaspidomorphi, Elasmobranchii, Holocephali, Myxini, Sarcopterygii; 29,879 spp.

Tasks: FishBase will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Prof. Rainer Froese, editor and leading manager of FishBase, founding Member of Species 2000.

18. Deutsches Krebsforschungszentrum (DKFZ)

FZK will proviode administrative and technical facilities for Reptile Database.

Reptile Database by P Uetz. Taxonomic database of world reptiles: Animalia – Chordata – Reptilia; 8,624 spp.

Tasks: Reptile Database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Peter Uetz, taxonomist of reptiles, author and custodian of Reptile Database. Member of Species 2000. Participated in EC FP5 EuroCat project.

19. Museum and Institute of Zoology – Polish Academy of Sciences, Warsaw (MIZ-PAS)

The Institute is a home organisation for Salticidae Database. Participated in EC FP5 EuroCat project.

Salticidae Database: Global Species Database of Salticidae (Araneae) by Jerzy Proszynski. Taxonomic database of world jumping spiders: Animalia – Arthropoda – Arachnida – Araneae – Salticidae; 5,000 spp.

Tasks: Salticidae Database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Wieslaw Bogdanowicz, Director of the Institute.
Dr Jerzy Proszynski, taxonomists of spiders and author of Salticidae Database.

20. Land Oberosterreich, Linz (Land OOE)

The Linz Museum (part of Land Oberosterreich) is a home organisation for ZOBODAT database. Member of Species 2000. Participated in EC FP5 EuroCat project.

ZOBODAT: Zoological-Botanical Database (Vespoidea) by J Gusenleitner. Taxonomic database of world vespid wasps: Animalia – Arthropoda – Insecta – Hymenoptera – superfamily Vespoidea; 5,925 spp.

Tasks: Zobodat Database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Gerhard Aubrecht, zoology of vertebrates.
Michael Malicky, ZOBODAT database system manager.

21. National Museum & Galleries of Wales, Cardiff (NMGW)

The Museum is a home organisation for MolluscaFW database. Participated in EC FP5 EuroCat project.

MolluscaFW: Checklist of Freshwater Mollusca by M Seddon & al. Taxonomic database of world freshwater molluscs

Tasks: MolluscaFW database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Mary Seddon, MolluscaFW database editor.
Ben Rowson, MolluscaFW database manager.

22. Institut de Recherche pour le Développement, Marseille (IRD)

IRD is a home organisation for TITAN database. Participated in EC FP5 EuroCat project.

TITAN: Cerambycidae database by G Tavakilian. Taxonomic database of world longhorn and timber beetles: Animalia – Arthropoda – Insecta – Coleoptera – Cerambycidae; 27,221 spp.

Tasks: TITAN database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Gerard Tavakilian, leading taxonomist of large family of timber beetles, author of TITAN database.

23. Agencia Estatal, Consejo Superior de Investigaciones Cientificas, Madrid (CSIC)

The Museo Nacional de Ciencias Naturales CSIC & Real Jardín Botánico de Madrid CSIC are home organisations for 3 databases. Member of Species 2000.

WTaxa: Electronic Catalogue of Weevil Names by M Alonso-Zarazaga & C Lyal. Taxonomic database of world weevils: Animalia – Arthropoda – Insecta – Coleoptera – superfamily Curculionoidea; 80,728 spp.

Eumycetozoa.com: Nomenclatural Database of Eumycetozoa (Myxomycota) by C Lado. Taxonomic database of world slime moulds: Protozoa – Mycetozoa; 1,074 spp.

RJB Geranium: Geranium Taxonomic Information System by C Aedo. Taxonomic database of world cranesbills: Plantae – Magnoliophyta – Magnoliopsida – Geraniales – Geraniaceae – genus Geranium; 409 spp.

Tasks: all 3 database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Marian Ramos, Museo Nacional de Ciencias Naturales.

Dr Miguel Alonso-Zarazaga, taxonomist of weevils and co-author of WTaxa database.

Dr Carlos Lado, taxonomists of slime moulds and author of Eumycetozoa.com data portal.

Dr Carlos Aedo, taxonomists of flowering plants and author of RJB Geranium database/

24. TSJ BVBA

TSJ BVBA is a home organisation for Scarabs database.

Scarabs: World Scarabaeidae Database by P Schoolmeesters. Taxonomic database of world scarab beetles: Animalia – Arthropoda – Insecta – Coleoptera – superfamily Scarabaeoidea; 17,444 spp.

Tasks: Scarabs database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Paul Schoolmeesters, author and custodian of Scarabs database. Member of Species 2000. Participated in EC FP5 EuroCat project.

25. Bayerische Staatsministerium fur Wissenschaft, Forschung und Kunst (SNSB, Munich)

Staatliche Naturwissenschaftliche Sammlungen Bayerns (SNSB, Munich) is a home organisation for SysMyr and LIAS databases.

SysMyr: Systematic Myriapod Database by J Spelda. Taxonomic database of world diplopods, pauropods and symphylans: Animalia – Arthropoda – Diplopoda, Pauropoda & Symphyla; 10,344 spp.

LIAS: A Global Information System for Lichenized and Non-Lichenized Ascomycetes by D Triebel & al. Biodiversity database of world lichens: Fungi – Ascomycota – Ascomycetes.

Tasks: SysMyr and LIAS databases will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Joerg Spelda, taxonomist, author of SysMyr database.

Dr Dagmar Triebel, editor and manager of LIAS database.

26. Universita Degli Studi Di Padova (UNIPD)

UoP Department of Biology is a home organisation for ChiloBase.

ChiloBase: A World Catalogue of Centipedes (Chilopoda) for the Web by A Minelli (ed). Taxonomic database of world centipedes: Animalia – Arthropoda – Chilopoda; 3,139 spp.

Tasks: ChiloBase database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Prof. Alessandro Minelli, leading taxonomist of centipedes and editor of ChiloBase.

27. Universiteit Van Amsterdam (UVA)

The Museum is a home organisation for Porifera database.

Porifera: World Porifera Database by R van Soest (ed). Taxonomic database of world sponges: Animalia – Porifera; 8,052 spp.

Tasks: Porifera database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Rob van Soest, taxonomist of sponges and editor of Porifera database.

28. Institut Royal Des Sciences Naturelles De Belgique, Brussels (RBINS)

The Institute is a home organisation for the Rotifera database, and operates a capacity building programme for the focal points of the CBD Global Taxonomy Intiative, in which copies of the Annual Checklist on CD are distributed to delegates from developing countries.

Rotifera database by H Segers. Taxonomic database of world rotifers: Animalia – Rotifera; 1,983 spp.

Tasks: Rotifera database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Hendrik Segers, taxonomist of rotifers and author of Rotifera database.

Yves Samyn, one of the team organising of the Belgian capacity-building programme for GTI and CBD focal points in developing countries.

29. The Chancellor, Masters and Scholars of the University of Oxford (University of Oxford)

Department of Zoology is a home organisation for WCS database.

WCS: World Catalogue of Strepsiptera by J Kathirithamby. Taxonomic database of world Strepsiptera (in progress): Animalia - Arthropoda - Insecta – Strepsiptera; 600 spp.

Tasks: Rotifera database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Jeyaraney Kathirithamby, taxonomist of Strepsiptera and author of WCS database.

30. Wageningen Universiteit (WU)

WU Department of Plant Sciences is a home organisation for Compositae Checklist database. Participated in EC FP5 EuroCat project.

Global Compositae Checklist by C Flann (ed). Taxonomic database of world Compositae (in progress): Plantae – Magnoliophyta – Magnoliopsida – Asterales – Asteraceae.

Tasks: Compositae Checklist database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Christina Flann, plant taxonomist and editor of Global Compositae Checklist.

31. Royal Botanic Garden, Edinburgh (RBGE)

RBGE is a home organisation for World Umbellifer Database.

World Umbellifer Database by M Watson (ed). Taxonomic database of world umbelifers (in progress): Plantae – Magnoliophyta – Magnoliopsida – Apiales – Apiaceae.

Tasks: World Umbellifer Database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot projects to enhance GSD quality.

Dr Mary Gibby, plant taxonomist.

Dr Mark Watson, plant taxonomist and editor of World Umbellifer Database.

32. Københavns Universitet, Copenhagen, Denmark (UCPH)

University of Copenhagen is a new home organisation for BioSystematic Database of World Diptera.

BioSystematic Database of World Diptera by NL Evenhuis, T Pape, AC Pont & FC Thompson (as editors); and T Pape as a custodian. Taxonomic database of world dipterans (flies, mosquitoes, bots, midges and gnats): Animalia – Arthropoda – Insecta – Diptera; 135,505 spp.

Tasks: Diptera database will participate in WP3 GSD Network and respond to GSD assessment, network consultations and workshops, acceleration of completeness and data quality, pilot project to enhance GSD quality.

Dr Thomas Pape, taxonomist of dipterans, co-author and custodianof Diptera database.

33. International Trust for Zoological Nomenclature (ITZN)

International Trust for Zoological Nomenclature is a legal body for International Commission on Zoological Nomenclature (ICZN).

ITZN plans to create world nomenclator database for animalian names - ZooBank.

Tasks: ITZN will participate in WP3 GSD network and respond to GSD network

consultations; designing two-way relations between taxonomic databases and nomenclators; participation in project workshops.

Dr Ellinor Michel, Executive Secretary of the ICZN.

Multi-Hub Network:

34. Chinese Academy of Sciences, Beijing, China (CAS)

Institute of Botany and Institute of Zoology CAS are non-for-profit organisations, which belongs to and powered by the Chinese Academy of Sciences. It is national leading institution for research and education in biodiversity. Since October 2006 IBCAS become a home organisation for "Species 2000 China Node", a regional Catalogue of Life hub in China. In present, two key organisations contribute to the Node activity: Institute of Botany and Institute of Zoology CAS. The Biodiversity Committee of the Chinese Academy of Sciences supports and supervises the establishment of the Node. The Node issued two editions of the Catalogue of Life China database (2007, 2008). Member of Species 2000.

Tasks: CAS will represent Species 2000 China Node and collaborate with WP 3 "Foresight study for geographical extension: the Global Multi-Hub Network". It will take significant role in Multi-Hub Network discussion, design and prototype project, which will link Catalogue of Life China to the Second Architecture of the Catalogue of Life.

Prof. Keping Ma, Director of the Institute of Botany, Secretary General of the Biodiversity Committee of the CAS, and Co-ordinator of Species 2000 China Node.

Dr Haining Qin, (at the Institute of Botany) plant taxonomist, content manager of China Checklist of Higher Plants.

Prof. Liqiang Ji; (at the Zoological Institute) co-ordinator of Chinese Animal Information System.

Huijie Qiao, (at the Zoological Institute) system manager of Chinese Animal Information System.

Xuehong Xu, & **Shao Qing**, (at the Biodiversity Committee office) Secretaries for Species 2000 China Node.

35. The Centro de Referência em Informação Ambiental, Campinas, Brazil (CRIA)

CRIA (Reference Center on Environmental Information) is a not-for-profit, non-government organization. Its aim is to contribute towards a more sustainable use of Brazilian biodiversity through the dissemination of high quality scientific information to support policy and decision making and to promote education for sustainable development. CRIA´s activities are focused on the development of information systems for the dynamic integration and visualization of biodiversity data from distributed information resources.

Tasks: CRIA will collaborate with WP 3 "Foresight study for geographical extension: the Global Multi-Hub Network". It will take significant role in Multi-Hub Network discussion, design and prototype project, which will link Brazilian data to the Second Architecture of the Catalogue of Life.

Dr Vanderlei Perez Canhos, President Director of CRIA, member of the Species 2000 Board of Directors.

Sidnei de Souza, Director of Information Systems at CRIA.

36. Smithsonian Institution, National Museum of Natural History, Washington, USA (National Museum of Natural History)

The Smithsonian Institution is an educational and research institute and associated museum complex, administered and funded by the US government . Smithsonian Institution is a home for the Integrated Taxonomic Information System (ITIS). This is a partnership designed to provide consistent and reliable information on the taxonomy of biological species. It is an international body, with Canadian and Mexican government agencies participating. The primary focus of ITIS is North American species, but many groups are worldwide. ITIS is an MOU partner of Species 2000 with the common goal to create the Catalogue of Life.

Tasks: ITIS will collaborate with WP 3 "Foresight study for geographical extension: the Global Multi-Hub Network". It will take significant role in Multi-Hub Network discussion, design and prototype project, which will link ITIS North America to the Second Architecture of the Catalogue of Life.

Dr Thomas Orrell, Acting Director of ITIS.
David Nicolson, ITIS Data Development Coordinator.
Dr Michael Ruggiero, ITIS Consultant

37. Commonwealth Scientific and Industrial Research Organisation, Canberra, Australia (CSIRO)

CSIRO will work in 4D4Life with its partner the Australian Biological Resources Study (ABRS) through their joint programme The Atlas of Living Australia (ALA). ABRS is part of the Parks Australia Division of the Commonwealth Department of the Environment, Water, Heritage and the Arts. It was established to coordinate research in taxonomy and document the flora and fauna of Australia and is Australia's national focal point for taxonomy. ALA is a partnership of Australian scientific institutions and organisations, funded under the Australian Government. The ALA is developing a data management system to catalogue and organise information relevant to the study of Australia's biodiversity. The project will bring together information from a wide variety of sources, including many of the country's most significant natural history collections and herbaria, ecological and observational data sets, images, online literature, diagnostic tools and molecular data.

Tasks: ALA will collaborate with WP 3 "Foresight study for geographical extension: the Global Multi-Hub Network". It will take significant role in Multi-Hub Network discussion, design and prototype project, which will link Atlas of Living Australia to the Second Architecture of the Catalogue of Life.

Cameron Slatyer, (Director, ABRS) with responsibility for the direction and function of ABRS and the ABRS National Taxonomy Research Grant Program.
Donald Hobern, ALA Director, coordinating and planning the development of the ALA as a set of coordinated data services.
Pam Beesley and **Alice Wells** manage the *Australian Faunal Directory* database.

38. Landcare Research New Zealand Limited, New Zealand (Landcare Research)

Landcare Research is one of nine independent Crown Research Institutes founded in 1992 from a reorganisation of Government funded research in New Zealand. New Zealand's leading provider of solutions and advice for sustainable development and the management of land-based natural resources. Research focuses on three key outcome areas: Sustaining & restoring biodiversity, Sustaining land environments, Sustainable business & living. About 380 staff, in nine locations: at Lincoln, Palmerston North, Hamilton, Auckland, Gisborne, Nelson, Dunedin and Alexandra. It is home for Species 2000 Asia-Oceania initiative which aims to create database of New Zealand and adjusted marine areas biota.

Landcare leads the national New Zealand Organisms Register (NZOR) Project which is aggregating data on all New Zealand species.

Tasks: Landcare Research will collaborate with WP 3 "Foresight study for geographical extension: the Global Multi-Hub Network". It will take significant role in Multi-Hub Network discussion, design and prototype project, which will link NZOR to the Second Architecture of the Catalogue of Life.

Jerry Cooper, member of Landcare Research staff in Lincoln, system manager and coordinator of NZOR biodiversity database.

B.2.3 Consortium as a whole

A proportion of the partners in the Consortium have worked together before, primarily in the FP5 EuroCat project that established the Species 2000 Catalogue of Life as a european scientific infrastructure, but also in other related activities such as EC ENBI and the EC EDIT NoE. This is important because the large GSD Network is already constituted and its members play a vital continuing role in providing the content that is refreshed and integrated to make the Catalogue.

The Consortium is well-balanced in that it builds where needed on parts of the existing programme, and brings in new partners with particular skills related to the new developments. Two of the new partners have joined the Consortium to lead its new developments. BGCI joins because of its skills in networking and outreach in the wider biodiversity community, and Narodny Museum Prague joins because of its experience with long-distance international partnerships.

Each of the GSD Network Partners is self-selecting on the basis of its prior taxonomic expertise and long-term database coverage – all that we know of in Europe have been included. In several cases the partner hosts more than one GSD, and five of the main partners also have GSDs at their institutions.

Partner 8, the Species 2000 legal body, brings a rather different role to the Consortium. It was established by the Species 2000 Catalogue of Life Programme to provide both a holding company that could hold IPR and Access Agreements on its behalf, and as a residual body that could persist between funded projects. It is registered in England as a not-for-profit scientific organisation (technically a company limited by guarantee), and is owned and governed by the participating organisations and individuals in the Species 2000 Catalogue of Life Programme. In 4D4Life it provides access to the prior achievements of the programme; it is the vehicle for any legal or financial sustainability; and it will again provide the residual body carrying the programme forward when the project is complete. Consortium partners in 4D4Life that are not already members will be invited to become members, so that they too have an interest in the residual body.

B.2.3.1 Subcontracting

i) Subcontracts for CD production in WP 6.
In addition to the main 4D4Life services propagated from the web, we intend to continue the distribution on CD, a service requested by the UN Convention on Biological Diversity, because many folk in developing countries still do not have sustained access to the web. The CD is presently requested by users from 79 countries and distributed to developing countries by two programmes, one run by RBINS in Brussels (partner 29), and one by BioNET International.

We request subcontracts up to a total of 18,000 euro from ETI to the manufacturer selected each year, as part of WP 6 (total approx. 14,000 CDs).

ii) Subcontract for freelance programmer in WP 3.
The IOPI (International Organisation for Plant Information) 'Global Plant Checklist' database delivers several plant families (Rosaceae, Cyperaceae) to the Catalogue of Life. The database

is 'hosted' at BGBM Berlin. This global database, and several other regional databases linked to the European Hub, are all linked to the Catalogue of Life Dynamic Checklist through a re-usable wrapper programme called the Pywrapper that originated in the EC BioCASE project. However the programmer who maintains and tailors the Pywrapper program and these Pywrapper linkages (Mr Markus Doering) has now left BGBM Berlin and works freelance (that is as a Self-employed person), largely for GBIF.

We request a subcontract for Mr Doering work for up to 18,800 euro on upgrading the configuration of Pywrapper linkages to IPT linkages needed to maintain or upgrade connections to the Catalogue of Life Dynamic Checklist.

B.2.4 Resources to be committed

B.2.4.1 Resources committed by the Consortium

4D4Life is the development of an e-infrastructure with an array of new services and improvements based on an existing programme with existing public services. The project thus inherits not only a very major public network of prior investment and development, but also the ongoing development and investment of its existing partners, both those that are participants in 4D4Life, largely based in Europe, and those outside Europe participating in the wider Species 2000 Catalogue of Life Programme but not part of this 4D4Life project. We here below attempt to characterise these resources for the three years of the 4D4Life project. In total they represent a very major ongoing commitment across this community.

1. The Global Species Database Network
The main Catalogue of Life is synthesised from knowledge provided by an array of global species databases (GSDs). At the start of 4D4Life this array is composed of 66 GSDs enhanced taxonomically and maintained in IT availability by the 66 custodians in their parent institutes, as academic networks or as individuals.

44 European GSDs that participate in the GSD Network of this project*
Presently 44 of the GSDs are based in Europe and are committed to 4D4Life by participants in the consortium. Five are committed by Partner 11 (NHM), five by partner 12 (CABI), five by partner 9 (VLIZ), four by Partner 4 (MNHN), four by partner 1 (UR), and 3 by partner 24 (CSIC). The remaining 18 are individual participants in the project (participants 7, 10, 13 – 23, 25 – 33). Together this array of GSDs provided by these participants is collectively referred to as ‘the GSD Network’ thoughout this Description of Work. This European array of GSDs currently supplies 809,520 species, and a reasonable estimate of the resource committed is 2 euro per species per year. Using this estimate, the total commitment is of 1.6 million euro per year, or 4.8 million euro over the three years.

22 Non-european GSDs*
Additionally partner 36 (CRIA) commits one GSD from Brazil, and partner 37 (National Museum of Natural History) commits two GSDs (one very large, with multiple sectors) from the USA, although their participation in 4D4Life is primarily for their work in WP 4. The remaining 19 GSDs are committed by organisations outside this 4D4Life project but inside the Species 2000 Catalogue of Life Programme, and based outside Europe. Together these non-european GSDs currently provide 351,191 species, estimated at a total of 700k euro per year, or 2.1 million euro over the three years.

*(Figures given here are for the 2009 Annual Checklist in preparation for publication in April/May 2009. They are dynamic and will change substantially during the course of the project.)

2. The non-ICPC country and other members of WP 4
Three of the six regional participants (partners 37, 38 & 39, National Museum of Natural History (USA), CSIRO (Australia) & Landcare (New Zealand)) are participating in Work package 4 at their own expense. Over the three years this is likely to involve them in costs directly related to this project of the order of 16,000 euro each, a total of 48,000 euro.

In addition to these costs of direct participation, all six regional hub participants bring the commitment of very substantial resources in existing and ongoing databases to the project, (e.g. the databases on Plants, Animals and Microbes of China being connected by the Chinese Academy of Sciences, the land (Landcare) and marine (NIWA) databases being connected to the New Zealand Organism Register by Landcare, etc, etc.). It is unclear exactly which expenditure should be tallied, but in each regional hub case it is clear that several million euro will be committed to these databases by the regional centre during the three years of the project. The purpose of WP4 is to design the structure that will bring all this knowledge into availability across the Multi-Hub network.

3. The existing infrastructure
As explained elsewhere the ownership of the pre-existing infrastructure, its management, its permissions and IPR agreements, are under the governance and management of the holding legal body Species 2000 (partner 8), which itself is controlled by all the partners through their membership of that legal body.

Partner 8 makes available both the combined IPR, permissions, copyright, domain names, user MOUs, achievements, product identity, user community, prior software and goodwill of the prior projects, but also the continuing global organisation (e.g. liaison with the Global Species 2000 Team), the benefits of having a residual legal body and the commitments of other non-european partners to the global programme, such as the non-european GSDs listed under 1) above. It is difficult to put a value on these extensive commitments, even though it amounts to a very substantial contribution to the project.

B.2.4.2 Resources requested from the Commission

4D4Life requests 3.3 M € contribution from the Commission for expenditure as follows:

(NB All partners, with the exception of Natural Musuem of Natural History, N.America, CSIRO/ALA Australia, and LandCare New Zealand receive minimum travel funds to attend all four Project Meetings)

WP 1 Management of the Consortium
The University of Reading requests one part-time clerical assistant to deal with the large consortium, including the GSD Network, and the associated financial administration, reporting and audit, plus 5% of Project Coordinator.

WP 2 Engagement with scientific users and partners
Botanic Gardens Conservation International requests one full time equivalent position to organise the networking events, organise the e-bulletin and communications strategy and platform, and both to elicit from user groups the prioritised proposals for new technical and

new educational and popular services, and to get the new services trialled and tested when they become available..
Staff travel(including Project Meetings & Management Committee), workshops, and promotional materials costs are requested, including permission to travel outside Europe when appropriate.

WP 3 Strengthening the GSD Network: a Virtual Community
MNHN Paris requests one full time position to lead and organise networking and funded pilot projects with all the custodians of databases in the GSD Network.
Staff travel (including Project Meetings & Management Committee), minor equipment and minor consumables are requested, including permission to travel outside Europe when appropriate.
Funds allocated to members of the GSD network are a standardised allocation of 16,000€ for up to 4 months work on the pilot projects, and 2800€ travel to project meetings. (If the personnel cost for the particular pilot projects agreed exceed 4,000 euro per PeM, then either the PeM will be reduced (e.g. to 3 PeM or to a cost of 16,000 euro), or the personnel effort will not be covered in full by the EC contribution of 16,000 euro).

WP 4 Foresight Study for geographical extension: Global Multi-Hub Network
Narodni requests one full time position to lead and organise networking with all partners in the Multi-Hub Network, plus pilot projects by CRIA Brazil and CAS China, plus create, maintain, and publish significant concept, design and prototype test documents.
Travel and subsistence for attendance at meetings in Europe for CRIA Brazil and CAS China. (National Museum of Natural History, N. America, CSIRO/ALA Australia and NZOR self-funded)
Staff travel (including Project Meetings & Management Committee), minor equipment and minor consumables are requested for Narodni, including permission to travel outside Europe when appropriate.

WP 5 Scientific Services of the Catalogue of Life
UR requests three full time staff plus 5% of the Project Coordinator:
1) Scientific Coordinator for the entire project, also leading on specifying, creating and testing each of the new or enhanced services, working with WP2, WP3 and WP6, and the production & systems team at Reading.
2) Taxonomic Editorial Assistant to lead on the continued and developing production cycle of the Annual Checklist and Dynamic Checklist production of an integrated Catalogue of Life.
3) Systems Manager to install, upgrade and maintain the internal and external service systems 24/7, to manage caching and refreshing processes, to manage the Annual Checklist test cycles for suppliers to view, and to lead on installation, development, and management of wrapper programmes across the entire project, as well as project web-sites and web-services.
Travel requested for Project Coordinator & staff to Project Meetings, staff visits to partners, Coordinator visits to partners, staff and coordinator presentations to international meetings, including outside the EU.
Equipment, materials and legal requested for Servers & rack, server services, promotion materials, legal fees for Access Agreements and sustainability issues, and consumables.
Sp2000 requests full time time assistant for Annual Checklist and CD-ROM compilation and production cycle, at low rate in Sp2000 Philippines Office.
Staff travel (including Project Meetings & Management Committee), minor equipment and minor consumables are requested for UR, including permission to travel outside Europe when appropriate.

WP 6 Software Support Services
ETI requests 2 full-time equivalent systems and programmer staff for the multiple programming and design tasks in WP6.
Staff travel (including Project Meetings & Management Committee), minor equipment including test server for ETI, including permission to travel outside Europe when appropriate.
VLIZ requests programmer time at 4 months.
CRIA Brazil requests programmer time at 6 months.
MNHN Paris requests programmer time at 3 months.

WP 7 Novel e-Science Service-based Architecture
CU requests 2 full time staff, 24 months of system designer, 18 months of programmer, plus 10% of e-science designer/manager.
Staff travel (including Project Meetings & Management Committee), minor equipment including test server are requested for CU, including permission to travel outside Europe when appropriate.

Section B.3: Potential Impact

B.3.1 Strategic Impact

4D4Life will increase the scale and stratification of federation within the Catalogue of Life community, and promote yet further interoperation of digital repositories, including across the world. The project has a strategy for enhancing the content and curation of the contributing repositories, and this, tied with middleware enhancements will increase the robustness of the data infrastructure. Also, by adopting a common management strategy to unify the presently separate integration of our two products (the Annual Checklist and the Dynamic Checklist) we anticipate reduced costs and more robust outcomes.

4D4Life does require a European level approach, that is it is neither national/local, nor yet truly global in activity and focus of expertise. The size of the task, as well as the technical architecture needed to complete it, does require major large scale cooperation between taxonomists and taxonomic institutions on a continental and global scale. However it is also a fact that even at the global and continental scale, the main concentration of taxonomists, expertise and activities is in Europe. This has been established in the EDIT NoE for the profession as a whole, and those statistics are echoed by the distribution of GSD activity seen in the Species 2000 Catalogue of Life programme. 44 of the 66 GSDs currently supplying data are european based, compared with 18 in the USA, and 4 from elsewhere.

Leveraged value of existing databases connected
Lower bound mean (extremes excluded) £ 78 per species
Upper bound mean (including extremes) £ 131 per species

	Databases	Lower bound	Upper bound
USA	18	£ 39,600k	£ 59,680k
Europe	44	£ 68,870k	£194,084k
Japan	1	£ 520k	£ 860k
Russia	1	£ 160k	£ 270k
Australia	1	£ 74k	£ 124k
Brazil	1	£ 45k	£ 75k
TOTAL	**66**	**£ 109M**	**£ 255M**

Invested cost of GSDs leveraged by central expenditure in GBP, Y. Roskov, Secretariat

An important issue within the landscape of international scale biodiversity and taxonomic programmes is a perception possibly raised by the exceptional publicity attending the start-up of GBIF, and the recent launch of EoL. Both of those programmes are deeply involved in the delivery of taxonomic information, and in the use of taxonomic structures to link and index other information. But it is now absolutely clear that neither GBIF nor EoL is attempting the same core task as we are carrying out in 4D4Life and the wider Species 2000 Catalogue of

Life programme. In recent months both those organisations have confirmed, in public and in private, that they continue to use the Catalogue of Life as taxonomic backbone for their activities. Enthusiastic letters of support were attached to the original proposal from both organisations – indeed two from EoL.

4D4Life is projected to make significant impact in five areas:

i) Opening up a range of electronic e-taxonomy services as infrastructure that can be built into the main international biodiversity information portals: GBIF, EoL, CBOL, iBOL, EBI/GenBank, IUCN.

ii) Opening up the same range of electronic e-taxonomy services as infrastructure that can become part of the seamless architecture of biodiversity informatics, and thus available to be built into the much wider range national local and NGO portals in Europe and around the world, and into novel experimental systems that are the subject of a growing research and development community.

iii) Opening up the same range, plus additional specially tailored e-taxonomy services to certain scientific communities, such as the biodiversity and climate change modelling community, the oceanographic community, or the genomics community.

iv) Launching a completely new range of taxonomic products based on the Catalogue of Life, and to be tailored and downloaded onto laptops and hand-held devices. These will provide educational and popular products that will impact students, fieldworkers, and amateurs.

v) and finally, simply by enhancing our present services we anticipate some additional impact on the wider user community.

B.3.1.1 International Biodiversity Information Portals

The new range of electronic e-taxonomy services will have an immediate and beneficial effect on our technical relations with the major international biodiversity portals that use our Taxonomic Hierarchy and Species Checklist as their taxonomic backbone. These include EoL, GBIF. uBio, IUCN, iBOL, CBOL. Indeed, an eagerness for this project to move to enhanced electronic services can be clearly seen in their letters of support appended to the proposal Part B document.

The services outlined in the description of Work Package 2 in section 1.4 Networking Activities fall into two classes: i) those that will be used for uploading to the users portal once only the entire or incremental content of the Hierarchy or Checklist by those who hold a full copy; and ii) those who plan to use our service repeatedly in real time, and not to hold a full copy. In the first class are several services to notify users of incremental changes to the Catalogue including species name changes, species circumscription changes, and higher taxon classification changes. These will be based on the taxon Globally Unique Identifiers (GUIDS) recently introduced using the Life Science Identifiers (LSID) system. These will be of interest especially to GBIF, EoL and uBio. They will allow GBIF, and uBio to track taxonomic concepts between annual editions, and to link external data with high accuracy. In the second class are “Find this name”, and “Synonymy server” services, and these will be of interest to iBOL, EoL and IUCN.

B.3.1.2 Services to the general Biodiversity Informatics and Biodiversity Portals community

In the longer term, the same types of electronic service as already mentioned in 3.1.1 are likely to have a more significant impact on the biodiversity community as they become available as standard components of the emerging seamless architecture for biodiversity informatics. This will mean that an even wider range of national, local and NGO biodiversity portals will incorporate our Catalogue or make use of our services in real time. It also means that our Catalogue and services become available to the growing community of innovative researchers in biodiversity informatics. This relatively new community is starting to realise the exciting potential and to experiment with novel ideas, structures and connections. Many of these may soon lead to completely new products and portals in what is a comparatively open field for future development. 4D4Life expects to be part of these new developments.

B.3.1.3 E-taxonomy electronic services to particular scientific communities.

The Catalogue of Life is already used in a number of scientific communities that have particular and different requirement. Examples include:

i) Climate Change Studies.
Using the catalogue as an index for harvesting specimen and observation records – for instance over the Internet or from GBIF or SpeciesLink. Such harvested datasets are frequently used in creating range maps to show distribution, or for use in species envelope modelling to predict distribution changes in climate change studies. To be accurate, and to yield the maximum harvest of records, the process needs to use both full synonymic expansion (to locate records of a species under all its names), but countered with alerts or blanking for ambiguities caused by homonyms, misapplied names and pro-parte synonyms (cases where the name alone cannot give accurate searches).

ii) Collections data cleansing
Data cleansing processes for herbarium, botanic garden and museum specimen record databases is another area where special issues are involved. The tools being developed are particularly good at alerting the users to errors caused either by small spelling variants, by regular spelling flips related to latin (phytos vs phyllos etc), and by confusions with phonetically close words. But at present we do not provide any facility to sort the species checklist and synonyms by these error-associations.

iii) Accessing DNA Barcode and Nucleotide data
The Programme has been approached by EBI and GenBank to discuss the idea of maintaining a cross-map between the Catalogue and the NCBI taxonomy used by the nucleotide databases and GenBank. Although the NCBI taxonomy is used in the established database, there is a need to provide indexing using the Catalogue of Life, because of its increasing use with DNA BarCode records being placed in Genbank by the CBOL and iBOL initiatives that cite the Catalogue of Life as a gold standard for annotation.

These examples illustrate the way in which the Catalogue of Life will provide special infrastructure services to particular disciplines, and how user communities wish to build our services into their own systems for daily use. These are examples of true infrastructure applications that we believe will play an increasing part in the impact of our public services – to professional communities rather than directly to the public.

B.3.1.4 Educational and popular products

This is a new area for the Catalogue of Life. A new generation of field biologists, students on field courses, expert amateurs and others expect to make heavy use of portable and hand-held electronic devices in the field and at home, no longer thinking of books as the principal medium of expert knowledge delivery.
Our new products may include components of the hierarchy and species list, attractive presentation with icons or illustrations of iconic species, and be both tailorable by taxonomic group and geographical area, and downloadable to laptops or hand held devices. Students attending fieldcourses, field study centres, field survey workers, amateurs travelling abroad, trainee field staff may all be interested in better tools to overcome the "taxonomic impediment", in a world where basic taxonomic knowledge is no longer taught to every biology undergraduate.

B.3.1.5 Simple improvements to our existing products

The Species 2000 Catalogue of Life programme has developed up to now as a primarily academic programme in taxonomy to complete as much as possible, responding to the issue of reaching critical mass, the need to establish the resource in the taxonomic community, and the need to convince critics that it can be created and can be maintained. The result is that development of a range of services has not been a high priority. It is clear from what users tell us that we can make a larger impact, and increase the regular public usage simply *by enhancing our present services, interfaces, web-services and web-sites.* The focus on services that this project will provide, will enable us to add to our impact by this relatively simple change of emphasis, in addition to the novelties discussed above.

B.3.2 Plan for the use and dissemination of foreground

B.3.2.1 Dissemination

4D4Life will become the dissemination arm of the wider Species 2000 Programme. The programme is already tuned to the concept of developing as a scientific infrastructure, and has a clear and unique role in the landscape of biodiversity platforms and services. What it has lacked up to now is a project outside academic taxonomy to develop its service role. 4D4Life provides the opportunity to realise and give momentum to the new services that we know are wanted by the user communities.

The present services are branded as "Species 2000 & ITIS Catalogue of Life", and are made available in four editions:

Annual Checklist 2008 (and preceding years to 2000) on the web:
http://www.catalogueoflife.org/annual-checklist/2008/search.php

Annual Checklist 2008 (and preceding years back to 2000) on CD-ROM (ISSN 1473-009X).

Annual Checklist 2008 (and preceding years to 2005) web-service:
http://webservice.catalogueoflife.org/annual-checklist

Dynamic Checklist on the web, launched October 2005:
http://www.catalogueoflife.org/dynamic-checklist/search.php

Dynamic Checklist web-service

http://webservice.catalogueoflife.org/dynamic-checklist/CASWebService

LSID resolver service (requires LSID plug-in for users browser)

The present public services are operated from servers at University of Reading, accessed via the web-sites of:

Species 2000 & ITIS Catalogue of Life - www.catalogueoflife.org
Species 2000 – www.sp2000.org
ITIS – www.itis.gov

The CD-ROM is distributed at conferences and mailed out, or it can be either ordered or downloaded as a CD-ROM image from the Species 2000 web-site

The Annual and Dynamic Checklists attract more than 40 million hits per year, with usage by serious unique visitors rising in 2007 from 20,000 per month to 40,000 per month. Other much larger usage is achieved when GBIF and EoL users make use of the catalogue or traverse between species on the GBIF and EoL portals.

The 3,500 CD-ROMS are distributed each year, including 600 to GTI and BioNET International focal points in developing countries, and including mail requests from 79 countries.

In addition to the dissemination of the Catalogue products, 4D4Life represents a scientifically interesting development that includes novelties, and it will be the policy of the project both to encourage partners to publish research articles about the project, but also to encourage all partners to present he developments of the project or its outcomes at relevant scientific and professional organisational events.

B.3.2.2 Management of Intellectual Property

It is proposed that 4D4Life will adopt the existing IPR framework set in place by the EC EuroCat project in FP5, and work in liaison with Species 2000 (partner 8) if it appears that changes are needed.

The key components in the existing framework are these:

i) 4D4Life and previous components of the wider Species 2000 programme operate an integrated service based on an integrated resource composed with contributions harvested from supplier databases. Standard Access Agreements are made with the supplier databases, and held by Species 2000 (partner 8) the residual legal body. These Access Agreements make clear that ownership, copyright and responsibility for taxonomic enhancement and maintenance of each database remain with the original custodian.
They also place a responsibility on Species 2000 to provide the three level credits in the published catalogue, and to make maintaining visibility of those credits a condition of third party licenses.

The entire Catalogue of Life is copyrighted as a taxonomic work by Species 2000 on behalf of its suppliers, and on behalf of ITIS, the North American partner. This enables Species 2000 and ITIS:

i) to enforce that third party licencees maintain the three level credit system that is important both to supplier databases, and to individual taxonomists named in the "latest taxonomic scrutiny" field.
ii) and to attempt to maintain control of the Catalogue as a single product in the community, that is without 'speciation' of variant forms.

The joint policy of the Species 2000 programme and of ITIS has been that the Catalogue is available to all individuals free of charge at point of use.

During 2007/2008 Species 2000 and ITIS have embarked on a trial license to a commercial company, where the service offered by the commercial company is quite outside the area of our own services. The model employed establishes an annual financial contribution, and includes the possible return of contributions to running costs from Species 2000 to the supplier databases. A letter of support from Taylor & Francis Ltd. for this 4D4Life proposal was appended to the original proposal.

B.3.2.3 Risk Assessment

The Project has identified the following risks, and put in place strategies for mitigating them or, in the worst case, dealing with a failure.

1. Conflicting priorities for new services, and the risk of disappointment.

Risk
Since the end of the EC EuroCat Project in 2006 the Catalogue of Life has, despite minimal finances, continued to build its knowledge resource quite strongly. But it has been unable to keep pace with rapid innovation in the area of e-services, such as multiple web-services and RSS feeds. This 4D4Life Project is intended to address this and to put in place a new array of services. The leaders have already accumulated an extensive list of suggestions and many more will be added to this. However, the resources of this project are limited, some suggestions will anyway prove impracticable, and it is likely that not all the services can be put in place.

There is a risk that users may have their expectations raised, only to discover that new services put in place are restricted to a modest improvement only, leading to some disappointment in the user community.

Mitigation
In addition to just listing and describing the new services wanted by users, Work package 2 will provide clarification of the different classes of user, and prioritisation for the suggested services. By providing separate prioritisations of the suggested new services in each class of users, it will then be possible for the Services Team to make decisions that relate to these user classes and to optimise user satisfaction with the outcome of the project. For instance it may be possible to please each major class with at least one or two top priority services, or alternatively to decide to meet the priorities of one particularly important class of users before others.

Failure
Risk of major failure in this is small. In the extreme case – of making little headway with services that are widely used – the Catalogue of Life would fall back to its present position, where it has successfully achieved both a reputation and substantial usership purely on the basis of its quality content, and despite the minimal basic services offered.

It can be argued that this is not a risk to the successful completion of this 4D4Life project, but a risk to the ongoing infrastructure after the project is over. Successful completion of this project is therefore not at risk.

2. Conflicts between multiple alterations to the annual production cycle at the same time, and the risk of interruption to the cycle.

Risk
The Annual Checklist, presently the flagship product of the Programme, is produced on a tight well-organised annual schedule. The schedule is tight because of the growing number of databases to be dealt with each year – itself one of the reasons for switching to the planned unified process, in which existing supplier databases would follow a more automated caching process. One of the strengths of the present programme has been the on-time regular publication of the new edition each year.

There is a risk that delays or faults with new and altered software will disturb this pattern, especially in year 2, when the switchover from separate production streams (for Annual and Dynamic Checklists) to the unified production stream.

There is a risk that unforeseen complications will result from making multiple changes to the process in the same period, for instance in year 3 with introduction of some data suppliers using the new e-2 architecture: because of the change over, the two systems will have to work in parallel.

Mitigation
We have built in a full scale test implementation for the unification, for instance of the Quality Assurance Workbench, to be tested before the decision is made to switch to the new routing. That decision is expected to be taken during year 2, so that the 2011 Annual Checklist and the Dynamic Checklist operated from the unified system by the end of year 2. Because of the ability to 'wrap' a partly made Annual Checklist, and then bring it in to the unified process through the wrapper, there is some room for minor delays in the switch-over during year 2, but with the Annual Checklist still to be created through the new system by the end of the year.

Failure
i) In all cases it is possible to revert to the previously used software, for instance creating the Annual Checklist and the Dynamic Checklists by the original processes.

ii) By switching to the unified process in year 2 we have one yearly cycle in hand. In the worst case the switch-over could be abandoned in year 2, and then completed in year 3.

iii) On-time publication of the yearly edition is not a contractual requirement, either of this project, or for other users, so a minor slippage of up to a month and a half could be tolerated without major consequences.

3. Risks in the voluntary participation model for the array of supplier databases, and the regional hub participants around the world on which this infrastructure depends, and the risk of their withdrawal or cessation of participation.

Risk
Global Species Database custodians provide the knowledge from which we synthesize the Catalogue of Life. The continental Regional Hub organisations will provide the knowledge that will bring the Catalogue of Life to full global coverage and ever nearer the goal of complete coverage. These databases and their custodians, these continental centres take part in the Species 2000 Catalogue of Life programme on a voluntary basis because of the 'positive' participation model established by Species 2000 over a number of years.

There is a risk that for any reason, one, some or even all of these organisations would leave the Programme, leading to a major collapse in the product and services. However this risk is assessed as small.

i) For the duration of this 4D4Life project, each participant will have signed both the Grant Agreement and the Consortium Agreement, signalling their willingness to take part, as well as, in all cases where eligible, to receive funds appropriate to their tasks.
ii) Outside the paperwork of this 4D4Life project, many of these participants have also signed Access Agreements, and, by becoming members of the Species 2000 residual legal body, become part of the ownership and governance of the ongoing wider Species 2000 programme. They are thus stakeholders in the programme.
iii) The 'Species 2000 model' has in general proved attractive to participants, and since its start in 1996, 66 organisations have come to participate. So far none has withdrawn from the programme.

Mitigation
Some of the taxonomic sectors provided by supplier GSD databases are available from alternative suppliers, although we have used peer review to select what are thought to be the most suitable in quality and coverage. In a few cases we even have alternative suppliers within the programme, because of overlaps. So – in a minority of cases the withdrawal of a sector could be followed by a replacement from a different source.

Failure
A low level of withdrawal of data suppliers would be unlikely to have severe consequences, as there are anyway substantial gaps in the present coverage. However any large scale withdrawal of data suppliers would be extremely serious for the ongoing programme. In such a case it would be important to identify possible problems with the participation model, and to attempt to rectify this quickly.

Section B.4: Ethical Issues

B.4.1 Data protection issues

B.4.1.1 Metadata

4D4Life will use and further develop the Species 2000 Metadatabase that contains, or will contain, technical metadata concerning the structure, management and content of each participant database. This metadatabase contains, or will contain, metadata from participants in all previous and current components of the Species 2000 programme, in addition to those joining the programme through this 4D4Life Project.

At the start of the 4D4Life Project the Directors of Species 2000 (partner 8 in the Consortium) will give permission for the 4D4Life Project to access and work on the metadata as an integral part of the wider Species 2000 Programme. At the end of the 4D4Life Project, the custodianship of the metadatabase will remain with Species 2000, the residual legal body for this and other components of the wider Species 2000 Programme.

The Metadatabase contains information supplied by the custodians of participating databases. The initial data for each database is taken from a questionnaire response provided by the custodian, and at the end of the questionnaire the custodian is required to give their permission for this data to be used and published by Species 2000. For databases actively contributing to the Catalogue, the data is refreshed each year, and some components of this are explicitly intended for publication or use in the credit fields of the Catalogue.

Public display and publication of some parts of the metadata (including logos and the names of experts marked in the “latest taxonomic scrutiny” field) is also regulated by the Access Agreement as a compulsory component of material published by Species 2000, and a compulsory component of datasets licensed to third parties. The Access Agreement is made between the database custodian and the Species 2000 legal body.

Partners in 4D4Life and their employees working on 4D4Life will be required to observe the confidentiality of certain items that relate to metadata, such as for instance the content of peer reviews. The content of peer reviews is confidential to the wider Species 2000 Programme, and shared only with the custodian of the database and staff and appointed officers of the wider Species 2000 Programme.

B.4.1.2 Data protection – personal data

Personal data, such as names, addresses, and contact details of database custodians and technical staff are held in the metadatabase and at the Species 2000 Secretariat, and are subject to the UK Data Protection Act. These are and will be kept to the minimum needed for the adequate functioning of the 4D4Life and wider Species 2000 Programme.

B.4.2 Issues relevant to ICT

As already listed above ethical and legal issues relating to intellectual property, and relating to data protection have been adequately addressed within the Species 2000 Programme, and will be adopted in this 4D4Life project. These arrangements are compliant with both the Charter

of Fundamental Rights of the European Union and the opinions of the European Group on Ethics in Science and New Technologies (EGE).

The 4D4Life Project does not make use of:

i) Data gathered on human volunteers
ii) Animals used in ICT Research
iii) ICT implants or wearable computing
iv) Personal data on health or genetics
v) Bio/Nano-electronics

B.4.3 Other issues

Within 4D4Life there are no issues relating to:

i) Informed Consent
ii) Use of animals
iii) Human embryonic stem cells

Ethical Issues table

	YES	PAGE
Informed Consent		
• Does the proposal involve Children?		
• Does the proposal involve patients or persons not able to give consent?		
• Does the proposal involve adult healthy volunteers?		
• Does the proposal involve Human Genetic Material?		
• Does the proposal involve Human biological samples?		
• Does the proposal involve Human data collection?		
Research on Human embryo/foetus		
• Does the proposal involve Human Embryos?		
• Does the proposal involve Human Foetal Tissue / Cells?		
• Does the proposal involve Human Embryonic Stem Cells?		
Privacy		
• Does the proposal involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)		
• Does the proposal involve tracking the location or observation of people?		
Research on Animals		
• Does the proposal involve research on animals?		
• Are those animals transgenic small laboratory animals?		
• Are those animals transgenic farm animals?		
• Are those animals cloned farm animals?		
• Are those animals non-human primates?		
Research Involving Developing Countries		

• Use of local resources (genetic, animal, plant etc)		
• Impact on local community		
Dual Use		
• Research having direct military application		
• Research having the potential for terrorist abuse		
ICT Implants		
• Does the proposal involve clinical trials of ICT implants?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO THIS PROJECT	**YES**	

End of DoW