A Curriculum Vitae Like Object describing Sigfrid Lundberg

In brief

Born 1956. Completed upper secondary 1975. B.Sc. in Ecology and life sciences, inorganic and physical chemistry 1979. Ph. D. 1985: Thesis *Five theoretical excursions into evolutionary ecology*, Lund 1985. Docent (approximately associate professor, although untenured) 1995. Since 1995 involved in digital library research and development, concentrating on information retrieval, metadata and text encoding in combination with database and web technology work at NetLab and later on Lund university Libraries.

1. Career

1.1. Positions and research/development projects after graduation

Postdoc (financed by Swedish Natural Sciences Research Council, NFR) at Imperial College, London, in Parasite Epidemiology Research Group visiting Prof. Roy Anderson, FRS. Assistant professor (forskarassistent, NFR) 1 July 1987 to 1 July 1993.

1.2. Teaching and supervision

The first few years as a post-graduate student I had higher teaching load than most students average (initially my position required 50% teaching). Later I have lectured mathematics for post-graduate students in ecology. I have been the supervisor of two post-graduate students who have successfully completed their theses. In addition I have regularly lectured on ecological modelling techniques and evolutionary genetics to students of theoretical ecology.

1.3. NetLab and Lund university Libraries

Since October 1995 scientist at the Lund university library, working on various development projects digital libraries, Internet search and information retrieval and digital preservation. I have contributed to design of our harvesting robot which is also used for archiving of WWW documents. I have large experience of combining full-text retrieval with bibliographic fielded search. More recently I have concentrated on searching more stringently tagged documents.

2. Computing and internetworking

As a natural scientist I have no formal merits in neither computer science, nor in software engineering.

My knowledge stems from the following sources:

- Many years experience of scientific computing within areas where shrink wrapped software is virtually absent, and in particular using numerical analysis.

- Many years experience of system administration (Unix) within the Dept of Theoretical Ecology and NetLab.

- Several years experience of inter-networking, initially at the Dept of Theoretical Ecology, and later at Lund University Library's development depart NetLab.

In the following I give *examples* of achievements within the area of software engineering and inter-networking. (In doing that, I will ignore the area, scientific computing, in which I have gained my

proficiency in computer programming.)

2.1. Operating systems

I have used computers intensely throughout my entire career within research and development. During the first five years (1979-1984) I wrote my applications on the university mainframe running the operating system OS1100. During that period I spent a significant proportion of my time punching cards. Later on I acquired a personal computer and used MS DOS (1985-1987), after that VMS on Digital VAX computers.

Since 1989 I have been using Unix exclusively for all applications and as a Unix user I have been my own system administrator until I joined NetLab. I have experience of System V (AIX, Linux and Solaris) Berkeley UNIX (Ultrix and SunOS). I have been using Linux since release 0.13.

2.2. Programming

I'm programming in Perl, XSLT, PHP, C, pascal and Fortran. I used to be prolific in the last two, but I am starting to forget them. I am best at Perl and XSLT (and XPath) and have a working knowledge on SQL. If needed, I could do work in PROLOG and TCL. (although the first projects would take some time). I can do Javascript as well, but hate it.

2.3. Scientific typesetting

I have a sincere interest in the formatting of scientific text, and use GNU troff for the purpose ("groff"—the printed version of this document is formatted using that system). I am also able to do program macros and customize macro packages in groff. This means that I can process pre-process documents, create indexing scripts. This means that I can handle complete scientific books or journal issues including automatic generation of tables of contents, index of authors, subjects and whatever, backward and forward cross references and the like. I manage to handle TeX and Latex if need be, but I have no experience in writing macros.

An example is the manuscript catalogue downloadable from `http://laurentius.lub.lu.se/`.

I am also interested in the photo typesetting of graphs and charts and I am the author of the gnuplot pic driver, which makes it possible to typeset graphs using groff and annotate those with mathematical symbols. I have contributed the once widely used Linux vgalib driver for the popular ghostscript postscript interpreter.

This document was created as a hypertext document in XML (actually in XHTML) and translated into groff using XSLT and printed on a postscript printer. That is, a typical procedure expected to operate in an electronic publishing system with print on demand facilities.

2.4. Markup languages

The last five years or so, I have been interested in manipulating structured text. Originally, I was forced into this by the fact that I was involved in extracting data from HTML in the harvesting systems we have developed at NetLab (see below). More recently I have been working with XML (and to some extent SGML) as a transfer syntax for metadata as well as in the context of scientific publishing. I am fluent in Text Encoding Initiative (TEI), Master and DocBook.

In both contexts I have experience of building as well as of the parsing and transformation of XML using XML Dom, the XPath language and XSLT. As regards metadata, I have used XSLT for translation between internal formats and different exchange formats as well as for translating XML for presentation and human consumption. Also, I am using XSLT for automatic indexing av textual documents, to transform them suitable for search and retrieval.

2.5. TCP/IP and inter-networking

I have experience of using and maintaining servers for most popular TCP/IP protocols appearing in an Unix environment, including SMTP, WAIS and Z39.50 databases, HTTP, ftp and gopher. In addition I have experience of maintaining mailing-

lists, line printing using Berkeley LPD, NFS etc.

Z39.50 is a protocol for information retrieval, and the protocol I am most familiar with beside the usual suite (mostly the ones mentioned above that are lumped together and labelled Worldwide Web — WWW is in principle those protocols that permit access through URLs). The public interfaces for most databases I have been involved in, the access protocol have been Z39.50. This is for compatibility between web search engines and library OPACs.

One of my specialties is harvesting of web resources for building resource discovery databases. At NetLab I have been engaged in coding the Combine harvesting robot. It is well suited for up to medium size harvesting projects (leading to databases up to 5.000.000 records). The Combine is written in perl and C++, and is a package of programs using custom TCP/IP protocols for inter-process communication. Through its architecture it may be run on a cluster of UNIX systems for enhanced performance. In practice we usually have a monolithic installation on a single system.

3. Activities, Projects and services

3.1 Projects and services

Theretical Ecology Simulation Server (1994–1995). See above.

Fråga en ekolog, Ask an ecologist (1995–1996). A service for the general public, where a panel of ecologists answer question on ecological and environmental issues. Comprised scripts that received questions via HTML forms, posted these on a mailing list. Panel members answered questions again using forms. The answers were automatically sent via e-mail to the person and likewise entered into a search and browsable database on the Web.

NWI. Nordic Web Index (funded by Nordinfo, BTJ and BIBsam). Initially meant to become a Nordic search engine build on public service principles. A distributed search service using individual harvesters in all Nordic countries. Each member had its own Z39.50 target. We started this before AltaVista, but after Lycos, which at the time did not support European character sets.

NWI II (funded by Nordinfo). A continuation of NWI. Working on Metadata, Z39.50 and archiving the Web

Desire (EU project) An infrastructure for European researchers. We wrote the Combine here, generalizing the NWI harvester.

Desire II (EU project) We did further improvements on Combine and the search systems for better metadata support.

SAFARI (The Council for Higher Education, Högskoleverket, transferred to The Swedish Research Council) I built this — not very popular but widely known — system for research information. The system in now closed.

ETB (EU project) Building a metadata exchange system for a network of information gateways, all of which exchanging records by posting them to a NNTP network

STUDERA.nu (Council for Higher Education, Högskoleverket) A database for academic courses. Metadata and full text is collected using harvesting robot, and made accessible using Z39.50. The user interface was built using a Z39.50 gateway.

S:t Laurentius digital manuscript library A digital library based upon complete digitalization and a complete catalogue built using XML technologies.

E-publishing I am involved in building cataloguing and publishing tools for Lund university Libraries, in particular forthcoming refurbished dissertation database and an entirely new services containing undergraduate theses.

3.2 Standarisation

I have following the standards development in information retrieval and metadata area since entering this business. In particular I have actively participated in

the Dublin Core Metadata Initiative (DCMI) since 1996 as working group member in several groups, and was for a couple of years co-chair for the implementors SIG and co-chair with Dan Brickley, W3C, for the DCMI architecture WG. Likewise I'm following the development of the Z39.50 protocol and the work on RDF/XML and the semantic web.

4. List of publications

Lundberg, S., 1985. The importance of egg hatchability and nest predation in clutch size evolution in altricial birds. *Oikos* 45:110–217.

Lundberg, S. and N. C. Stenseth, 1985. Coevolution of competing species: Ecological character displacement. *Theoretical Population Biology* 27:105–219.

Lundberg, S. and T. Alerstam, 1986. Bird migration patterns: Conditions for stable geographical population segregation. *Journal of Theoretical Biology* 123:403–414.

Lundberg, S. and C. Löfstedt, 1986. Intraspecific communication in the sex communication channel: A selective force int the evolution of moth pheromones. *Journal of Theoretical Biology* 125:15–24.

Holmgren, N. and S. Lundberg, 1990. Evolution of migration patterns in dominance structured populations. *Ring* 13:219–220.

Lundberg, S., 1990. Evolution of increased susceptibility to infectious diseases in age structured populations. *Journal of Theoretical Biology* 147:133–142.

Holmgren, N. and S. Lundberg, 1993. Despotic behaviour and the evolution of migration patterns in birds. *Ornis Scandinavica* 24:103–109.

Massad E, S. Lundberg and H. M. Yang, 1993. Modeling and simulating the evolution of resistance against antibiotics. *International Journal of Biomedical Computing.* 33:65–81.

Holmgren, N., S. Lundberg, and P. Yodzis, 1993. Functional responses and heterogeneities: reanalysis of an experiment with cladocerans. *Oikos* 76:196–199.

Ingvarsson, P. K. and S. Lundberg, 1993. The effect of a vector-borne disease on the dynamics of natural plant populations: a model for Ustilago violacea infection of Lychnis viscaria. *Journal of Ecology* 81: 263–270.

Lundberg, S. and L. Persson, 1993. Optimal body size and resource density. *Journal of Theoretical Biology* 164:163–181.

Åström, M., P. Lundberg, and S. Lundberg, 1994. Population dynamics with sequential density dependencies. *Oikos* 75:174–181.

Lundberg, S., J. Järemo, and P. Nilsson, 1994. Herbivory, inducible defence and population oscillations: A preliminary theoretical analysis. *Oikos* 71:537–540.

Lundberg, S. and H. Smith, 1994. Parent-offspring conflicts over reproductive efforts: Variations upon a theme by Charnov. *Journal of Theoretical Biology* 170:215–218.

Ingvarsson, P. K. and S. Lundberg, 1995. Pollinator functional response and plant population dynamics: Pollinators as a limiting resource. *Evolutionary Ecology* 9:421–428.

Lundberg, S., P. Nilsson, and T. Fagerström, 1996. Seed dormancy and frequency dependent selection due to sib competition: The effect of age specific gene expression. *Journal of Theoretical Biology* 183:9–17.

Lundberg, Sigfrid and Pär K. Ingvarsson, 1998. Population Dynamics of Resource Limited Plants and Their Pollinators. *Theoretical Population Biology* 54:44–49

Ardö A., Arvidsson A., Hammer S., Holmlund K. and Lundberg S., 1998. NWI II, An Enhanced Nordic Web index: Final report. *Nordinfo-Nytt* 3–4/98

Ardö A., S. Lundberg, 1998. A Regional Distributed WWW Search and Indexing Service — the DESIRE Way. *Computer Networks*, 30:173–183

Lundberg, S., 1999. En arkitektur för ett distribuerat system för för spridning av forskningsinformation. `http://sigge.lub.lu.se/1998/safari/cris.pdf`

Lundberg, S., 2001. Technical description of the Studera.Nu search engine Harvesting policies, maintenance and software used.
`http://sigge.lub.lu.se/2001/studera.nu/documentation/studera.pdf`

Lundberg, S., 2002. S:t Laurentius Digital Manuscript Library: An excursion along the border between resource discovery and resource description. Paper presented at DRH2002, cf.: `http://www.drh2002.lib.ed.ac.uk/abstracts.html#170`). The paper is at `http://laurentius.lub.lu.se/search/presentation/laurentius.pdf` or (in source format) `http://sigge.lub.lu.se/2002/Laurentius/laurentius.xml?passthru=1`