

« Only the initiates will have the secrets revealed » : débats entre chimistes autour de l'ouverture des logiciels scientifiques

Alexandre Hocquet & Frédéric Wieber

Séminaire du centre François Viète

Nantes, 28 février 2017



UNIVERSITÉ
DE LORRAINE



Laboratoire
d'Histoire des Sciences
et de Philosophie
Archives Henri Poincaré

Subject: Re: Taxol modellingDate: Fri, 1 Jan 1993 14:11:05 ESTFrom: Franci Michelle M <mfran1@cbrn.mawr.edu>Date: Mon, 4 Jan 93 12:33:36 -0800From: hcj@gull.uncc.edu (Harry C. kfz-heidelberg.de)Subject: energy contribution of hbonds?Date: 06 Jan 1993 10:55:29 -0500 (EST)From: "dwj@lilly.com (Doug Johnson MC7R7)" <JOHNSON_DOUGLAS_W@LILLY.COM>Subject: u : CADD videosDate: Wed, 6 Jan 93 13:03:33 -0800From: chiremv!andromeda!jeffb@uunet.UU.NET (Jeff Blaney)Subject: Re: Polygon-based solvent accessible surfacesDate: Thu, 7 Jan 93 11:00:00 ESTFrom: WILLIAMS%XRAY2@ulkyvxo2.louisiana.edu (William Williams)Date: Wed, 13 Jan 93 13:38:09 ESTFrom: rsefeck@world.std.com (Joe R. Seifeck)Date: 1:18:26 -0500From: jle@world.std.com (Joe R. Seifeck)Date: .OHIO-STATE.EDUSubject: Steric strain at T calcs on diradicals/TSDDate: Mon, 18 Jan 93 16:36 ESTFrom: landman@hal.physics.wayne.edu (Landman)Date: Re: density problemDate: Wed, 20 Jan 93 0:00 ESTFrom: nauss@wrair-emh1.army.mil (Susan Nauss)Date: 22 Jan 93 13:06:32 UTFFrom: "BD 1:46 -0500From: system@alchemy.chem.utoron.amsay@emunix.emich.eduSubject: CHEMICALC: n 93 12:57:08 -0500From: watanabe@tammy.hated Chemistry Course Info RequestedDate: W t: Re: Periodic Table WidgetDate: Wed, 27 opulation Analysis in GaussianFrom: mei@ve Fri, 29 Jan 1993 16:47:38 -0600From: sliu and excited statesDate: Fri, 29 Jan 93 15 ntDate: 01 Feb 93 19:26:49-0500From: SBPM t: DMF dielectricDate: Mon, 1 Feb 93 17:57 e -> 2D drawing?From: QINGSONG@minmet.lan. tware for calculating zeta potential and e vin Moore <kmoore@ncsc.org>Subject: MOLSOL 17:32 +0100From: ivan@gandalf.ciam.unibo.i son.edu>Subject: Memory and CPU for Starde M (Adi Treasurywala)Subject: Coordinates o lel workstations and serversDate: Wed, 10 800From: rickl@biosym.comSubject: DIBUG ad heresa@si.fi.ameslab.gov (Theresa Windus)S er <K360171%EDVZ.UNI-Linz.AC.%@OHSTVMA.ACS 4:47:09 +0100 (WET)From: PREINERT@biolan.u for PC ..Date: Tue, 16 Feb 93 11:29:01 PS 4:15:16 -0500 (EST)From: "DR. DOUGLAS A. S 2 PSTFrom: fisher@scripps.edu (Cindy Fische 3 11:53:05 -0500From: G. Ravishanker <ravi -scale DynamicsDate: Mon, 22 Feb 93 16:49: 6-2300From: "Alexandre MICU P16A" <alex@ba window terminal emulator for MacDate: Thu, X-windows Server on MS-WindowsDate: Thu, 2 wang@kjemi.uio.noSubject: language questio 1 GMT-2:00Date: Mon, 1 Mar 1993 15:25:19 w bmt on computational chemistryDate: Tue ect: Re:Cache on the MacDate: Tue, 2 Mar 9 tion?From: Jan Labanowski <jkl@ccl.net>Dat ender@chemdc1.tau.ac.il (Senderowitz Hanoc brandeis.edu (steve gallion)Subject: GROW 2 Mar 93 04:51:15 GMTFrom: jim@quanta.phy. 3 +0100From: doelz@comp.bioz.unibas.ch (Re .eduSubject: Problem with vibrational freq 3 11:49:33 GMTSubject: Re: MM2 code for UN rheller@asrr.arsusda.gov>Subject: Software ject: First Congress of the ISTCP (pa e: Fri, 19 Mar 93 14:33:18 ESTFrom: Ge lder <AHOLDER@VAX1.UMKC.EDU>Subject: Semic r)Subject: efg in g92From: "Christopher J : Batching files in MOPACFrom: zrfl0128@aw n, 29 Mar 93 15:49:03 BSTFrom: rafapa@obel

Re: Gaussian Flame-wars

- From: "Phil Hultin" <hultin.,at,.cc.UManitoba.CA>
- Subject: Re: Gaussian Flame-wars
- Date: Tue, 11 Dec 2001 16:52:36 -0600

Max Valdez wrote:

>I think Gaussian is a great tool, but it has a LOT of little secrets and
>"bugs", and the support is not so good, I don't think the spend
>that much money on customer support, given the price of the product.

I absolutely agree with this statement. I think I represent a new kind of scientist working with tools like Gaussian. I am an experimentalist, not a computer whiz, [...]

So what's the point of all this? Well, maybe people would be less prone to jump on Gaussian (as they have over this makefile thing) if they didn't feel that the philosophy of Gaussian Inc. was similar to that of the "high priest" - only the initiates will have the secrets revealed and then only after years of study. A nice start would be to make the program work the way the manual implies that it should (or to make the manual describe how the program actually works). Even better: bring the program into the 21st century by providing a user interface that is actually user-friendly. What the priests and initiates regard as "a philosophy and aesthetic of simplicity" looks quite otherwise to many of us on the outside!

Dr. Philip G. Hultin
Associate Department Head and
Associate Professor of Chemistry
University of Manitoba

uk>Subject: Polyrate QueryDate: Tue, 30 Mar 1993 19:31:32 +0200From: c62@aixfile1.urz.uni-heidelberg.de (Andreas Schulz)Subject: NBO analysesDate: Tue, 30 Mar 1993 16:04:42 - ulz)Subject: Stephen R. WilsonDate: Thu, 1 Apr 1993 16:33:29 +0200From: oles@gollum.uio.noSubject: BiosymDate: 01 Apr 1993 11:43:13 -0600 (CST)From: CUNDARIT@memstvx1.memst.edu

Plan

1. Introduction

2. Chimie computationnelle : éléments d'histoire

3. CCL : liste de discussion, exploitation du corpus

4. Débats et tensions dans les deux fils choisis

Contexte

Objet d'étude inscrit dans la thématique générale des rapports entre science et ordinateur (Agar, 2006, et refs. citées)

Objet d'étude s'inscrit aussi dans la question de l'histoire du logiciel (Campbell-Kelly, 2007, 2003) (Mahoney, 2008) (Ensmenger, 2009)

MAIS : histoire du logiciel dans les communautés scientifiques peu été développée (voir tout de même, par ex. Spencer, 2015)

Chimie computationnelle : science qui s'est développée à partir des années 1980, avec ordinateur comme élément constitutif

Objectif

Mettre à jour (certaines) des tensions provoquées par le logiciel dans la communauté scientifique choisie : nécessite d'un corpus qui les révèle

Décrire la complexité de la question (particulièrement actuelle) de l'ouverture / « openness »

Plan

1. Introduction

2. Chimie computationnelle : éléments d'histoire

3. CCL : liste de discussion, exploitation du corpus

4. Débats et tensions dans les deux fichiers

choisis

Computational Chemistry : Un champ disciplinaire en train de s'affirmer dans les années 80

Lié à l'ordinateur, et donc à ses mutations :
-mutations hardware : de big computer à personal computer (de supercomputing à desktop computing)

//Ordinateur 50/60//

//objet big science, lointain, réservé à des spécialistes, difficile d'accès, fédérally funded science, matériel limitant//

//Ordinateur 80/90//

objet banal, personnel, connecté, modulable, accessible, typique de « scientific entrepreneur », matériel de moins en moins limitant,

→ façons de modéliser différentes (par ex. selon Lenhard) : conséquences du matériel sur transformations épistémiques



Au delà de ce « big picture », notre domaine précis est au coeur d'enjeux techniques, industriels et économiques

- constructeurs cherchent à promouvoir le terminal graphique pour le vendre à l'industrie pharma et utilise computational chemistry comme promesse techno-scientifique

- contexte économique mais aussi politique :

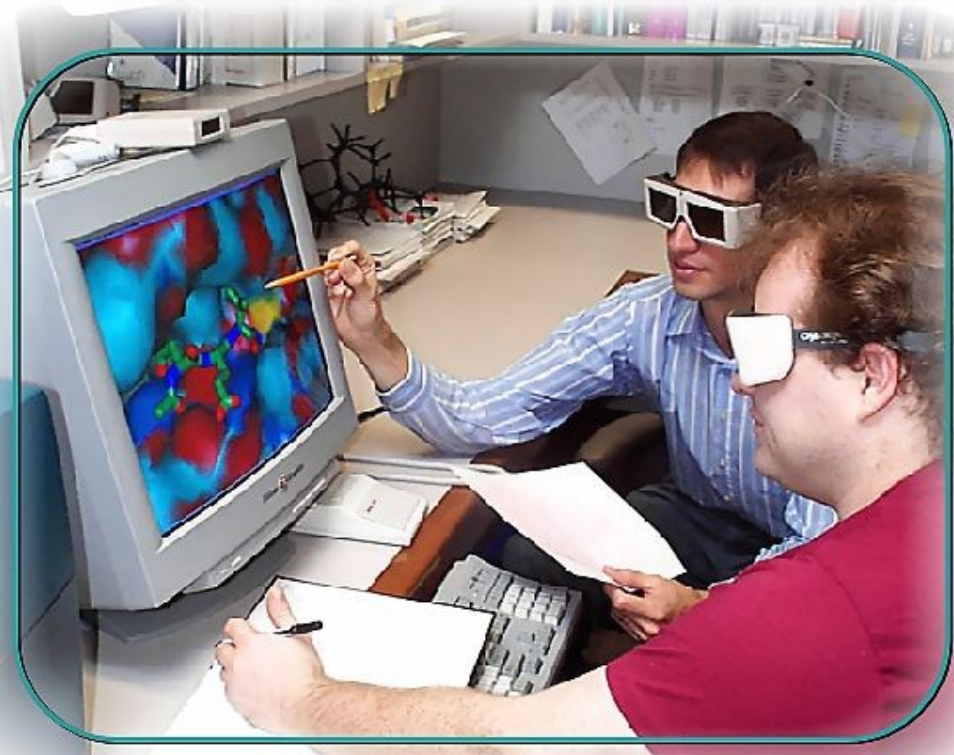
- en pleine époque de la « entrepreneurship science », dans le domaine voisin des biotechnologies (Jones)

- bouleversement des pratiques de « federally funded science » à « entrepreneurship science » : Bayh Dole act, technology transfer, pression des universités américaines sur la question du patenting, copyright... (Slaughter Roades, Popp Berman)

- dans un domaine scientifique (chimie, pharmacie...) ou culture du brevet et du secret est forte. (contrairement à « computing science », « république des informaticiens »...) (Mowery, Flichy)

Beaucoup se sont intéressés aux transformations de l'activité scientifique liées à ces mutations de l'ordinateur mais :

n'a pas été abordé le software en tant qu'outil du scientifique, outil construit par le scientifique dans ce contexte, à la frontière du scientifique, du technique, de l'économique et du politique. (un objet frontière, Leigh Starr)



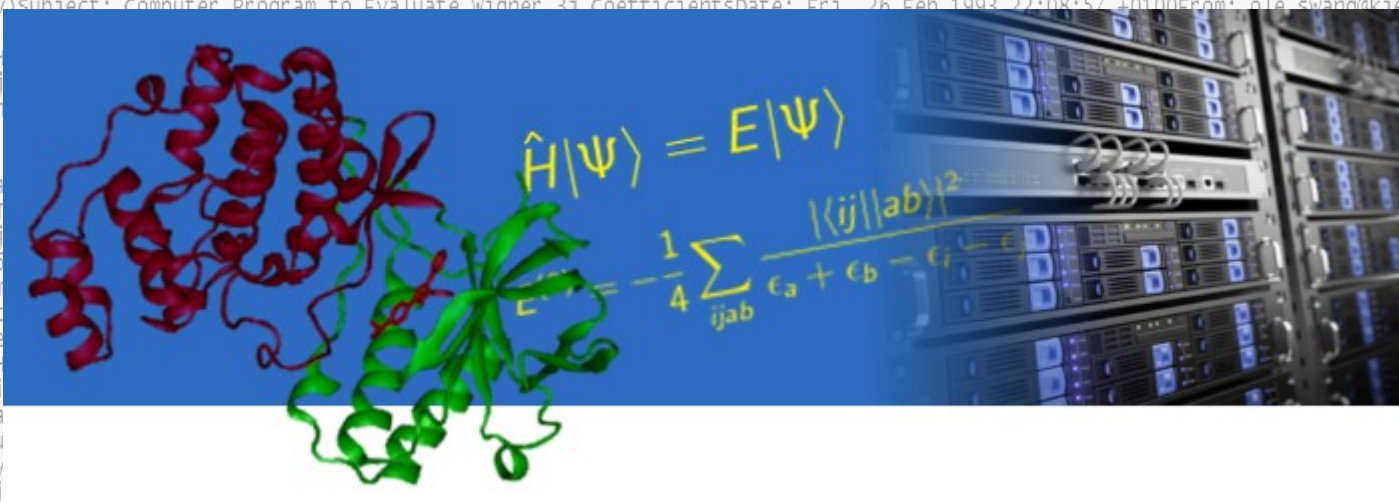
A quelle « culture épistémique » appartient la computational chemistry ? deux cultures différentes mais entremêlées !

une issue de la physique quantique, puis la chimie théorique : complexité mathématique, réductionnisme de la chimie à la physique, narration de la « pureté théorique »

et l'autre de la spectroscopie, chimie physique et chimie organique : description newtonienne de boules et ressorts, pragmatique, épistémiquement attachée à la chimie et à sa façon historique de représenter les molécules, liée plus tôt à l'industrie et dont la complexité épistémique réside dans la paramétrisation et optimisation plutôt que dans le calcul.

Même dans chimie quantique : (discours de Coulson) : tensions entre idéal de calcul "ab initio" et paramétrisation : "méthodes semi empiriques"

Tout un spectre de méthodes, avec des modèles épistémiquement divers, aux origines scientifiques différentes, reliés à des champs disciplinaires différents, mais avec un objet suffisamment commun (et rendu commun par l'outil) pour définir un domaine avec ses journaux, ses conférences dans les années 80s...et sa mailing list en 1991.



Plan

1. Introduction

2. Chimie computationnelle : éléments d'histoire

3. CCL : liste de discussion, exploitation du corpus

4. Débats et tensions dans les deux fichiers choisis

3. Computational Chemistry List (CCL)

CCL.NET
computational chemistry list, ltd.

Wombat Bioactivity Database:



<http://www.ccl.net/index.shtml>

CCL Home

Messages in a queue: 0 (approx. wait time: 00:00 hours)

About CCL

Rules, Instructions,
Supporting CCL,
Uploading file to
CCL Site, Security
Policy, Privacy
Policy, About Us
Contact Us

**Drug Discovery
Training**

Drug Discovery

Page supported by:



Computational Science & Engineering Online
... a laboratory without walls for scientific computing

Welcome to the CCL Website!

ACS COMP Division supports CCL:

1991	CCL is born	E-mail Listserver and FTP archives for computational chemists are established at OSC as a hobby project of Jan Labanowski.
1994	CCL gets Gopher	CCL Introduces Gopher interface to messages and its FTP site. CCL grows in number of subscribers and messages .
1995	CCL gets WEB	CCL Introduces WEB interface to messages and its FTP site.
		
1996	CCL gets money from NSF	After your help with support letters, CCL got a 3-year grant from the National Science Foundation (NSF)
1996	CCL gets domain	Just in case, Jan Labanowski personally registers the domain ccl.net since the good domains are being bought by the millions by domain name investors.
1996	CCL buys small iron	After deliberations and all the advice not to (<i>...buy something solid if you have money...</i>) CCL decided to operate from PCs running Linux and Apache Web Server.

hcgj@gull.uncc.edu (Harry C. V. DOUGLAS_W@LILLY.COM>Subject: u
ible surfacesDate: Thu, 7 Jan 93
out PDB searching methodsDate: F
om: WILLIAMS%XRAY2@ulkyvxo2.loui
an 93 13:38:09 ESTFrom: rsefeck@
500From: jle@world.std.com (Joe
E.EDUSubject: Steric strain at T
diradicals/TSDate: Mon, 18 Jan 9
: landman@hal.physics.wayne.edu
ity problemDate: wed, 20 Jan 93
om: nauss@wrair-emh1.army.milSU
: 22 Jan 93 13:06:32 UTFFrom: "BD
From: system@alchemy.chem.utoron
rix.emich.eduSubject: CHEMICALC:
:08 -0500From: watanabe@tammy.ha
try Course Info RequestedDate: W
iodic Table WidgetDate: wed, 27
analysis in GaussianFrom: mei@ve
an 1993 16:47:38 -0600From: sliu
ed statesDate: Fri, 29 Jan 93 15
Feb 93 19:26:49-0500From: SBPM
lectricDate: Mon, 1 Feb 93 17:57
awing?From: QINGSONG@minmet.lan.
calculating zeta potential and e
<kmoore@ncsc.org>Subject: MOLSOL
From: ivan@gandalf.ciam.unibo.i
bject: Memory and CPU for Starde
asurywala)Subject: Coordinates o
ations and serversDate: wed, 10
ickl@biosym.comSubject: DIBUG ad
fi.ameslab.gov (Theresa windus)S
L%EDVZ.UNI-Linz.AC.%@OHSTVMA.ACS
LOO (WET)From: PREINERT@biolan.u
Date: Tue, 16 Feb 93 11:29:01 PS
500 (EST)From: "DR. DOUGLAS A. S
fisher@scripps.edu (Cindy Fische
-0500From: G. Ravishanker <ravi
Date: Mon, 22 Feb 93 16:49:
: Carl JugDate: Tue, 23 Feb 93 15:08:43 PSTFrom: wsonnen@alnitak.usc.edu (Wayne Sonnen)Subject: dna base chargesDate: 23 Feb 93 15:38:36-2300From: "Alexandre MICU P16A" <alex@ba
g)Subject: Re: Hydrogen bond.Date: Wed, 24 Feb 93 17:21 EDTFrom: <PODOSNNA%NYUACF1.BITNET@OHSTVMA.ACS.OHIO-STATE.EDU>Subject: X-window terminal emulator for MacDate: Thu,
nl.govSubject: X-window terminal emulator for MacDate: Thu, 25 Feb 93 17:19:09 -0500From: feng@sgi.chem.temple.edu (Feng Chen)Subject: X-windows Server on MS-windowsDate: Thu, 2
d.harvard.edu (jill cheney)Subject: Computer Program to Evaluate Wigner 3j CoefficientsDate: Fri, 26 Feb 1993 22:08:57 +0100From: ole.swang@kjemi.uio.noSubject: language questio
semiempirical calculations simple oxidesFrom: sender@chemdcl.tau.ac.il (Senderowitz Hanoch)Subject: MOPAC6.0Date: Sun, 28 Feb 93 13:44:41 GMT-2:00Date: Mon, 1 Mar 1993 15:25:19
n.desubject: CCL Date: Mon, 1 Mar 1993 15:25:19
ate.Edu (Gene Carter)Subject: Co Mb parameters.Date: Tue, 2 Mar 93 17:22:53 CSTFrom: Greg Landrum <landrum@chemres.tn.cornell.edu>Subject: Re:Cache on the MacDate: Tue, 2 Mar 9
lling Software: Does anyone have STELLA?Date: 04 Mar 1993 17:08:31 -0500 (EST)From: MARYJO@northeastern.eduSubject: PDB subdir organization?From: Jan Labanowski <jkl@ccl.net>Dat
ar 93 17:30 Date: 04 Mar 1993 17:08:31 -0500 (EST)From: MARYJO@northeastern.eduSubject: PDB subdir organization?From: Jan Labanowski <jkl@ccl.net>Dat
secc.fi.cnr.itSubject: MO Theory Made VisibleDate: Tue Mar 09 11:23:12 1993Date: Tue, 9 Mar 93 10:52:25 -0500From: gallion@auriga.rose.brandeis.edu (steve gallion)Subject: GROW
uplingDate: Thu, 11 Mar 93 18:54 EDTFrom: <PODOSNNA%NYUACF1.BITNET@OHSTVMA.ACS.OHIO-STATE.EDU>Subject: Software for PCDate: Fri, 12 Mar 93 04:51:15 GMTFrom: jim@quanta.phy.
r 1993 02 Date: 12 Mar 93 04:51:15 GMTFrom: jim@quanta.phy.
MA.ACS.OHIO-STATE.EDU>Subject: partial charges for fictive atoms in g92Date: Mon, 15 Mar 1993 12:55 ESTFrom: HSUS@chemv2.mps.ohio-state.eduSubject: Problem with vibrational freq
, 15 Mar 1993 21:55:05 -0600 (CST)From: Chris Parkinson <mbdtscp@hpc.ch.man.ac.uk>Subject: MM2 code for UNIX systemsDate: Tue, 16 Mar 93 11:49:33 GMTSubject: Re: MM2 code for UN
ar 1993 11 Date: 16 Mar 93 11:49:33 GMTSubject: Re: MM2 code for UN
st Girona Seminar on Molecular SimilarityDate: Thu, 18 Mar 93 10:40:05 CETFrom: ZSYAMP01@EBCESCAL@OHSTVMA.ACS.OHIO-STATE.EDUSubject: First Congress of the ISTCP (pa
:11 -0500Subject: Keep the noise down, pleaseDate: 19 Mar 93 11:37:00 ESTFrom: nauss@wrair-emh1.army.milSubject: Program called GRIDDate: Fri, 19 Mar 93 14:33:18 ESTFrom: Ge
23 -0600 Subject: What about simulated annealing?Date: wed, 24 Mar 93 14:11:32 +0100From: dufner@ws01.pc.chemie.th-darmstadt.de (Hagen Dufner)Subject: efg in g92From: "Christopher J
.milSubject: What about simulated annealing?Date: wed, 24 Mar 93 14:11:32 +0100From: dufner@ws01.pc.chemie.th-darmstadt.de (Hagen Dufner)Subject: efg in g92From: "Christopher J
ject: SC&A Conference speakers - thanksDate: Thu, 25 Mar 93 10:18:48 -0500From: au195@cleveland.freenet.edu (Dr. Gene A. Nelson)Subject: Batching files in MOPACFrom: zrf10128@aw
MO@CHEMNA.DICHI.UNINA.ITSubject: looking for ACES programFrom: Robert Laird <mbdtserl@hpb.ch.man.ac.uk>Subject: Fe-0 calculationDate: Mon, 29 Mar 93 15:49:03 BSTFrom: rafapa@obel
uk>Subject: Polyrate QueryDate: Tue, 30 Mar 1993 19:31:32 +0200From: c62@aixfile1.urz.uni-heidelberg.de (Andreas Schulz)Subject: NBO analysesDate: Tue, 30 Mar 1993 16:04:42 -
ulz)Subject: Stephen R. WilsonDate: Thu, 1 Apr 1993 16:33:29 +0200From: oles@gollum.uio.noSubject: BiosymDate: 01 Apr 1993 11:43:13 -0600 (CST)From: CUNARIT@memstvx1.memst.edus

K kjemi <bjerkes@kjemi.unit.no>subject: Summary. Scaled freq for s and HDate: Sun, 21 Feb 1993 20:51:19 +0100 (MET)subject: Long time-scale dynamicsDate: Mon, 22 Feb 93 16:49:
: Carl JugDate: Tue, 23 Feb 93 15:08:43 PSTFrom: wsonnen@alnitak.usc.edu (Wayne Sonnen)Subject: dna base chargesDate: 23 Feb 93 15:38:36-2300From: "Alexandre MICU P16A" <alex@ba
g)Subject: Re: Hydrogen bond.Date: Wed, 24 Feb 93 17:21 EDTFrom: <PODOSNNA%NYUACF1.BITNET@OHSTVMA.ACS.OHIO-STATE.EDU>Subject: X-window terminal emulator for MacDate: Thu,
nl.govSubject: X-window terminal emulator for MacDate: Thu, 25 Feb 93 17:19:09 -0500From: feng@sgi.chem.temple.edu (Feng Chen)Subject: X-windows Server on MS-windowsDate: Thu, 2
d.harvard.edu (jill cheney)Subject: Computer Program to Evaluate Wigner 3j CoefficientsDate: Fri, 26 Feb 1993 22:08:57 +0100From: ole.swang@kjemi.uio.noSubject: language questio
semiempirical calculations simple oxidesFrom: sender@chemdcl.tau.ac.il (Senderowitz Hanoch)Subject: MOPAC6.0Date: Sun, 28 Feb 93 13:44:41 GMT-2:00Date: Mon, 1 Mar 1993 15:25:19
n.desubject: CCL Date: Mon, 1 Mar 1993 15:25:19
ate.Edu (Gene Carter)Subject: Co Mb parameters.Date: Tue, 2 Mar 93 17:22:53 CSTFrom: Greg Landrum <landrum@chemres.tn.cornell.edu>Subject: Re:Cache on the MacDate: Tue, 2 Mar 9
lling Software: Does anyone have STELLA?Date: 04 Mar 1993 17:08:31 -0500 (EST)From: MARYJO@northeastern.eduSubject: PDB subdir organization?From: Jan Labanowski <jkl@ccl.net>Dat
ar 93 17:30 Date: 04 Mar 1993 17:08:31 -0500 (EST)From: MARYJO@northeastern.eduSubject: PDB subdir organization?From: Jan Labanowski <jkl@ccl.net>Dat
secc.fi.cnr.itSubject: MO Theory Made VisibleDate: Tue Mar 09 11:23:12 1993Date: Tue, 9 Mar 93 10:52:25 -0500From: gallion@auriga.rose.brandeis.edu (steve gallion)Subject: GROW
uplingDate: Thu, 11 Mar 93 18:54 EDTFrom: <PODOSNNA%NYUACF1.BITNET@OHSTVMA.ACS.OHIO-STATE.EDU>Subject: Software for PCDate: Fri, 12 Mar 93 04:51:15 GMTFrom: jim@quanta.phy.
r 1993 02 Date: 12 Mar 93 04:51:15 GMTFrom: jim@quanta.phy.
MA.ACS.OHIO-STATE.EDU>Subject: partial charges for fictive atoms in g92Date: Mon, 15 Mar 1993 12:55 ESTFrom: HSUS@chemv2.mps.ohio-state.eduSubject: Problem with vibrational freq
, 15 Mar 1993 21:55:05 -0600 (CST)From: Chris Parkinson <mbdtscp@hpc.ch.man.ac.uk>Subject: MM2 code for UNIX systemsDate: Tue, 16 Mar 93 11:49:33 GMTSubject: Re: MM2 code for UN
ar 1993 11 Date: 16 Mar 93 11:49:33 GMTSubject: Re: MM2 code for UN
st Girona Seminar on Molecular SimilarityDate: Thu, 18 Mar 93 10:40:05 CETFrom: ZSYAMP01@EBCESCAL@OHSTVMA.ACS.OHIO-STATE.EDUSubject: First Congress of the ISTCP (pa
:11 -0500Subject: Keep the noise down, pleaseDate: 19 Mar 93 11:37:00 ESTFrom: nauss@wrair-emh1.army.milSubject: Program called GRIDDate: Fri, 19 Mar 93 14:33:18 ESTFrom: Ge
23 -0600 Subject: What about simulated annealing?Date: wed, 24 Mar 93 14:11:32 +0100From: dufner@ws01.pc.chemie.th-darmstadt.de (Hagen Dufner)Subject: efg in g92From: "Christopher J
.milSubject: What about simulated annealing?Date: wed, 24 Mar 93 14:11:32 +0100From: dufner@ws01.pc.chemie.th-darmstadt.de (Hagen Dufner)Subject: efg in g92From: "Christopher J
ject: SC&A Conference speakers - thanksDate: Thu, 25 Mar 93 10:18:48 -0500From: au195@cleveland.freenet.edu (Dr. Gene A. Nelson)Subject: Batching files in MOPACFrom: zrf10128@aw
MO@CHEMNA.DICHI.UNINA.ITSubject: looking for ACES programFrom: Robert Laird <mbdtserl@hpb.ch.man.ac.uk>Subject: Fe-0 calculationDate: Mon, 29 Mar 93 15:49:03 BSTFrom: rafapa@obel
uk>Subject: Polyrate QueryDate: Tue, 30 Mar 1993 19:31:32 +0200From: c62@aixfile1.urz.uni-heidelberg.de (Andreas Schulz)Subject: NBO analysesDate: Tue, 30 Mar 1993 16:04:42 -
ulz)Subject: Stephen R. WilsonDate: Thu, 1 Apr 1993 16:33:29 +0200From: oles@gollum.uio.noSubject: BiosymDate: 01 Apr 1993 11:43:13 -0600 (CST)From: CUNARIT@memstvx1.memst.edus

Computational Chemistry List (CCL) : bref historique

1991 Ohio Supercomputing Center

Communauté de scientifiques geeks mais pas computer scientists.

Discipline émergente (journaux, congrès...) : communauté à cimentier.

accessible et ouverte à tous → arène informelle

Une diversité d'acteurs:

-Developers : écrivent le code puis le diffusent

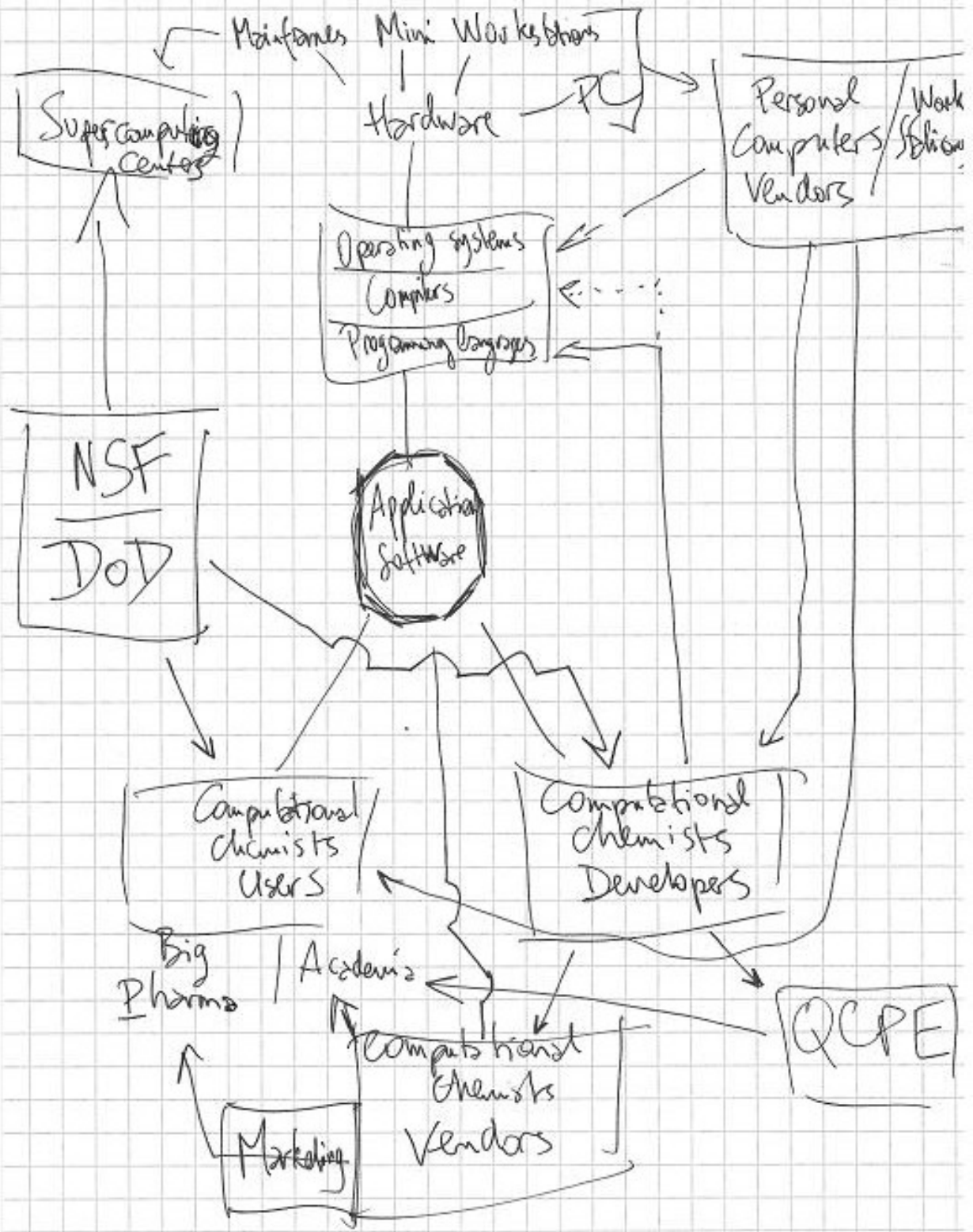
-Users (Academe/Industry): utilisent le code

-Vendors : vendent des logiciels

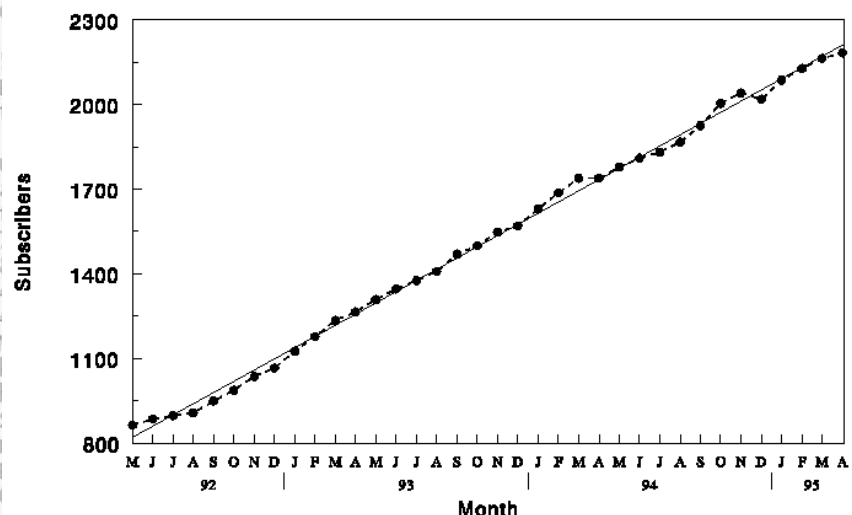
-NSF, DoD : financent des projets

-Supercomputers : allouent des ressources de calcul

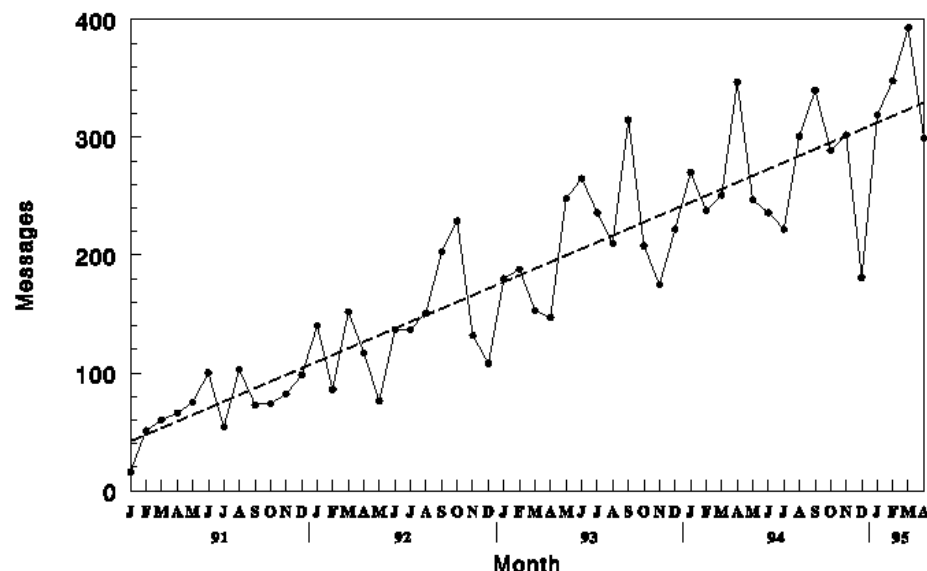
Hardware vendors vendent l'infrastructure



Number of subscribers



Messages per month



-croissance régulière puis stabilisation (approx. 10 messages par jour)

-academic / for profit

-senior positions / graduate students

-developers / users

-USA / pays émergents

90s : Internet=email ; entre attrition et explosion

Rules

By posting to the list, you agree to adhere to the rules outlined in this document.

These rules, regarding what topics are acceptable/unacceptable and commercial announcements discussions. Do not reopen Pandora's box -- please.

Acceptable Topics

Topics which **may appear on the list** include:

- **software/hardware-related** issues
 - bugs: Reporting bugs in chemistry software and possible work-arounds
 - programs and utilities
 - availability: Information on the availability of particular software, data, etc.
 - hardware-related issues relevant to the computational chemistry community
- **announcements** of
 - new chemistry software
 - computational chemistry-related workshops and symposia
- **methods**
 - new methods and techniques in computational chemistry
 - Q&A: questions and answers about solving computational chemistry problems
- **opinions** about services and products

Formes et thèmes du débat liées à l'administration (qui peut poster?) et la modération. Terms of service: quels sont les sujets admis?

Tensions académique/commercial

Commercial Announcements

Commercial software: software sold to anyone for use on the **list**, please respect the following rules:

- **Be brief** -- Short (25 lines or less) commercial posts
- **Be straightforward** -- The subject line should be clear
- **No reposting** -- Reposts of software announcements are not allowed
- **No "vaporware"** -- Announcements of "vaporware" are not allowed.

Definition « threaded conversation »

« threaded conversation » est une forme de conversation asynchrone publique unique et propre à Internet -> mailing list, newsgroups, jusqu'à comms de blogs, murs facebook.

quasi-oralité + topicality : échanges publics autour d'un sujet défini (ref "public sphere" Habermas)

full disclosure of methods?

Re: full disclosure of methods?

Re: full disclosure of methods?

Re: full disclosure of methods?

- From: st-amant{*at*}cgl.ucsf.edu (Alain St-Amant)
- Subject: Re: full disclosure of methods?
- Date: Sat, 26 Jun 93 12:21:56 -0700

Douglas Smith recently wrote in the current discussion on the disclosure of parameters:

> Besides, who ever said we had to reveal all our secrets and make > readily available and accessible? When software copyrights and p > really provide adequate protection, maybe I will agree with that I'll assume that Dr. Smith is referring to the specific algorithms implemented that make the program more efficient but do not affect results. In which case, I might agree.

Of course, I couldn't disagree more if he is referring to some development in the methodology that actually affects the final results in any way. The point that interests me however, is the question of software copyrights and patents to which Dr. Smith alludes. I have been trying to get for what can be copyrighted and patented and I get a different answer everyone. Can only specific code be copyrighted or can the structural algorithms be copyrighted as well? How 'modified' should code be if it can be called legally (and ethically if anyone is interested in an opinion) a new program? Or is it simply forbidden for code to "evolve" into a new program?

I will summarize to the net any e-mail sent to me, but I think that make for an interesting discussion and it would be as interesting to

Initial Message

Reply

Reply

Reply

Reply

Reply

Analyse quantitative: Small data mining

Dizaines de milliers de messages organisés: texte + date + auteur (métadonnées) permet:

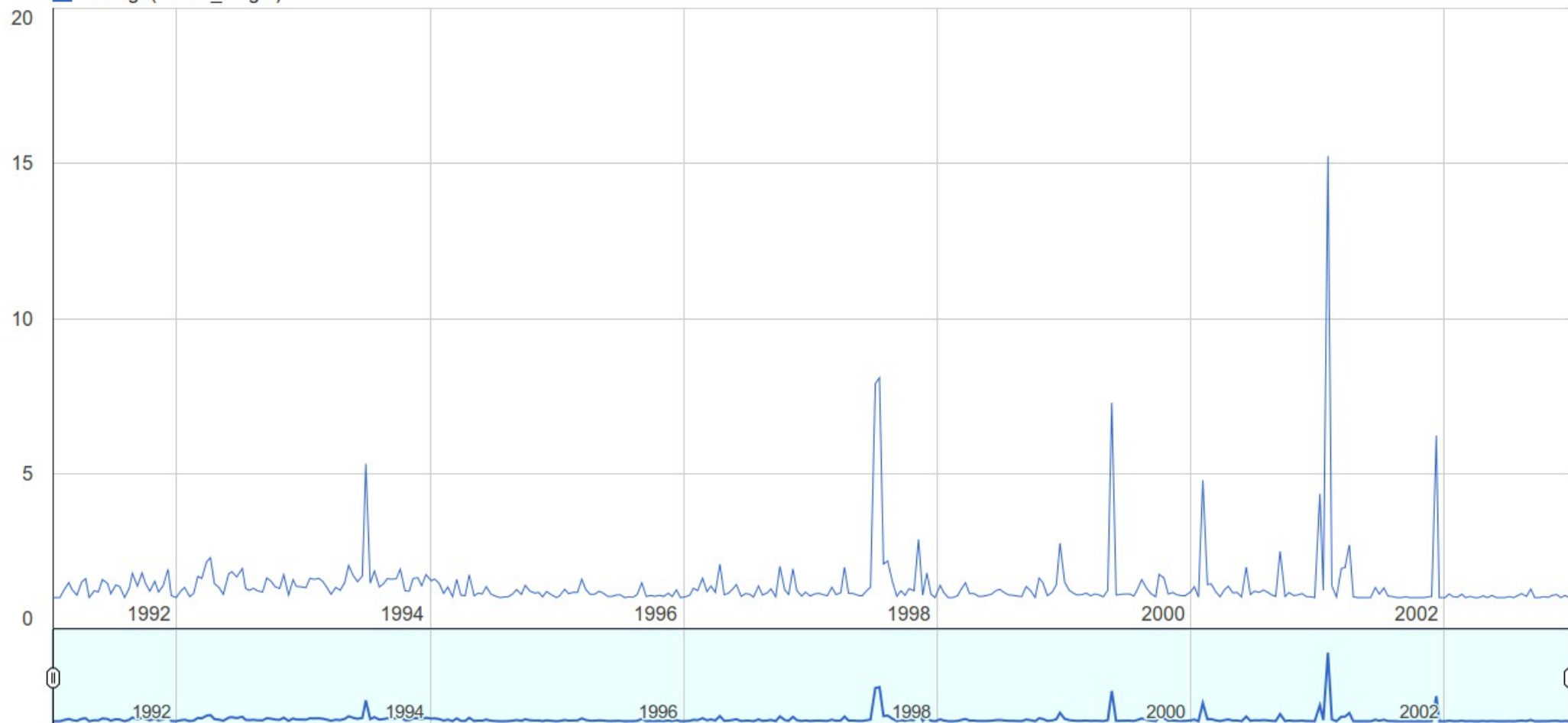
Time series

Social network analysis

Text mining

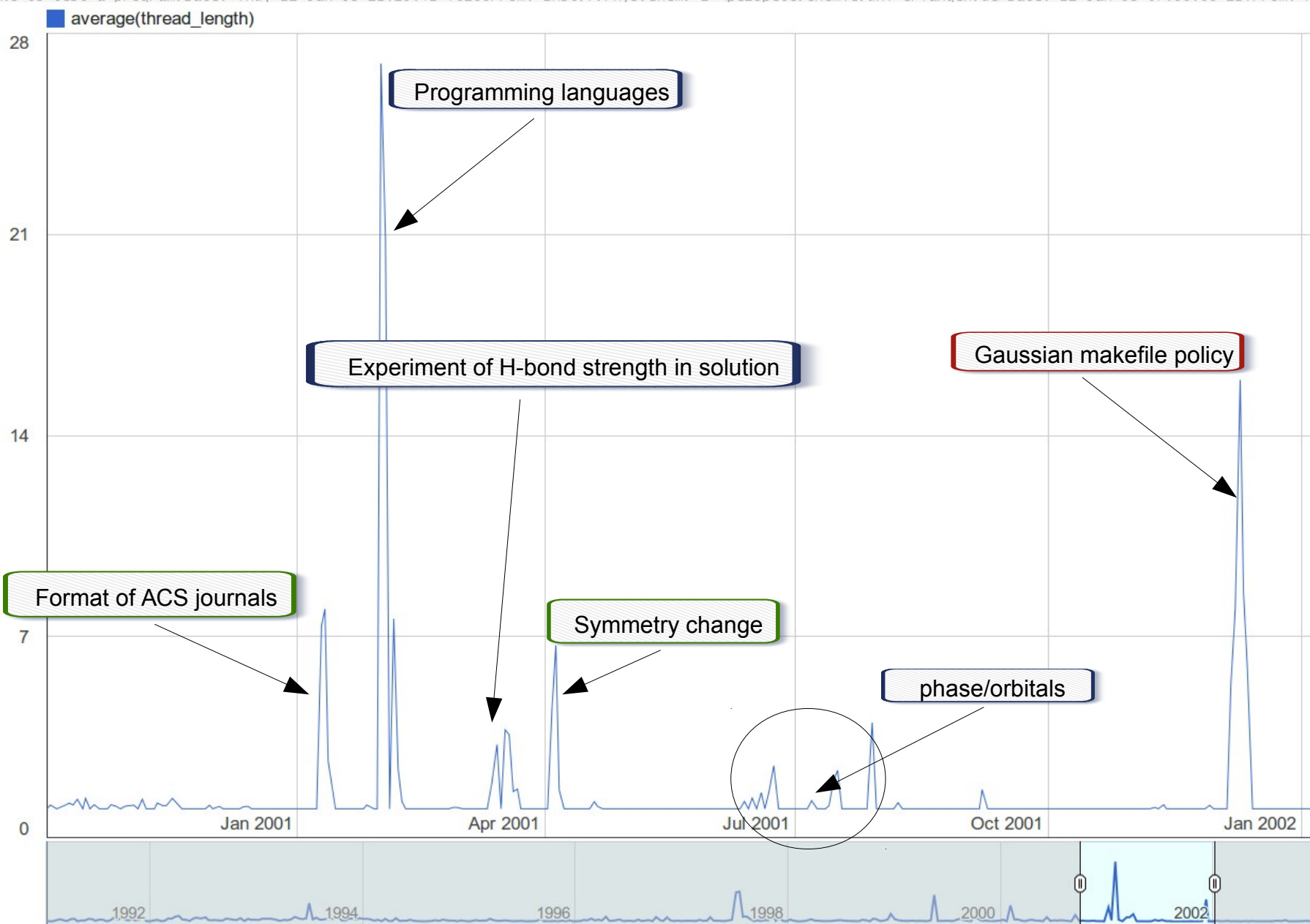
Nouveauté: utiliser ces techniques avec le « thread » comme unité de mesure

pas de flammes avant 1993 (netiquette et bande passante)

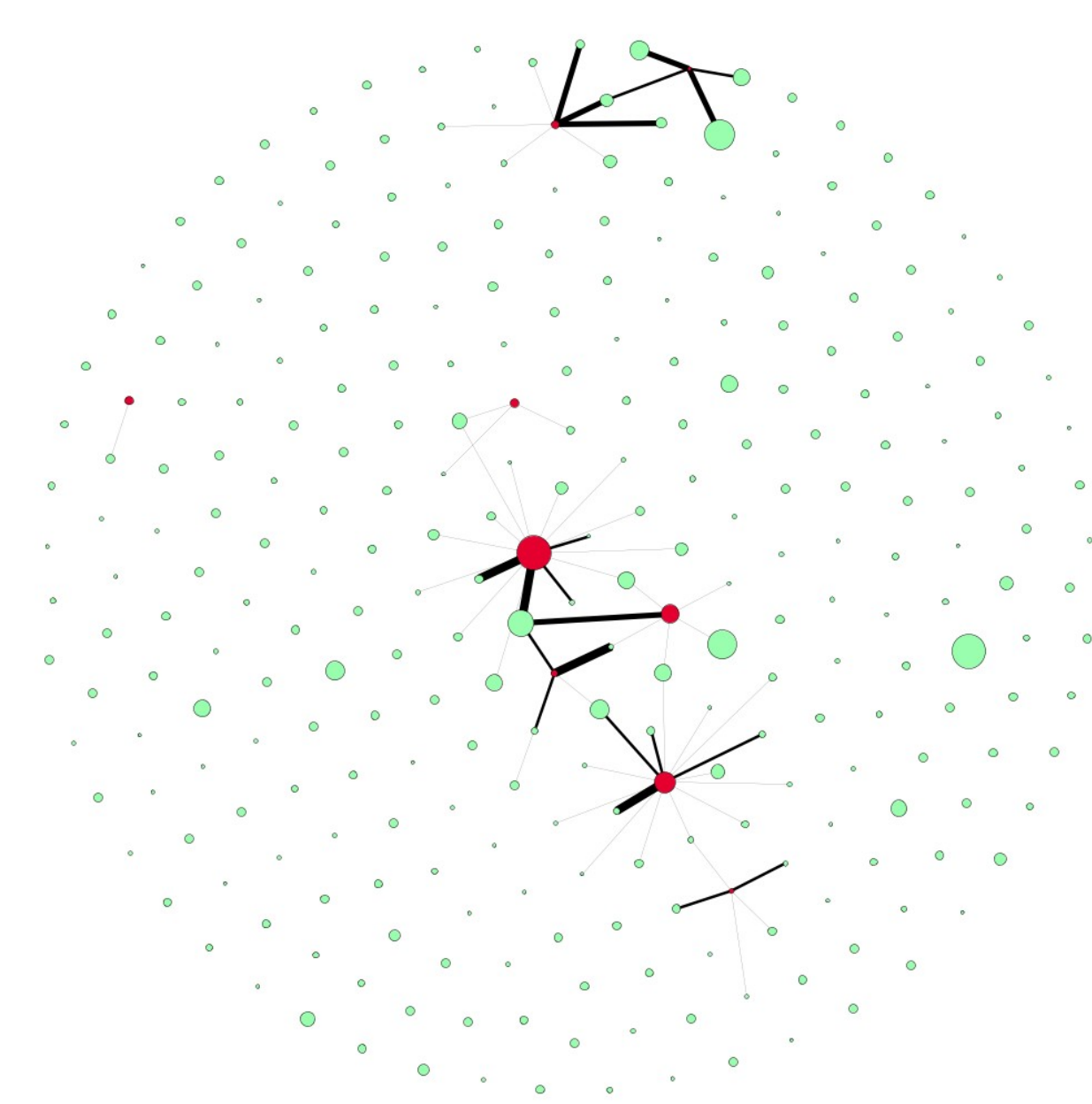


Small data mining: Time series

2001 zoom : Révéler les topics intéressants d'une certaine année



Small data mining: Social Network Analysis



Qui poste dans quels fils ?

Analyse de réseaux sociaux permet cartographier les fils et les acteurs (ici 2001).

point rouge = un fil
point vert = un posteur
(gros = gros nb total de messages)

liens : "poste dans un fil"
(gros = gros nb de messages dans ce fil)

majorité des fils = petits

Small data mining: Social Network Analysis

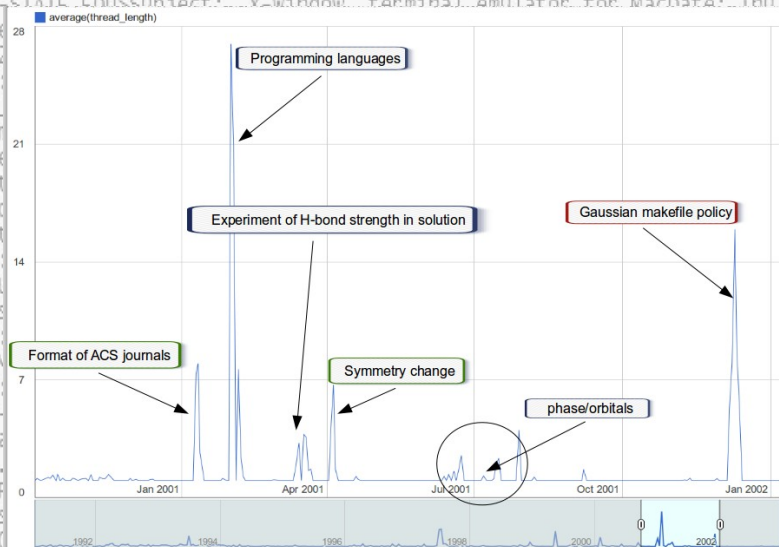
Programming languages

Format of ACS journals

Gaussian policy

vue filtrée : les principaux fils de 2001 apparaissent et leur proximité correspond à des posteurs en commun
les principaux acteurs sont aussi caractérisés par les fils auxquels ils contribuent

familles thématiques de threads, et de posteurs



Small data mining: Text mining

Evolution de trois keywords (graphics, linux, supercomputer) en fonction du temps

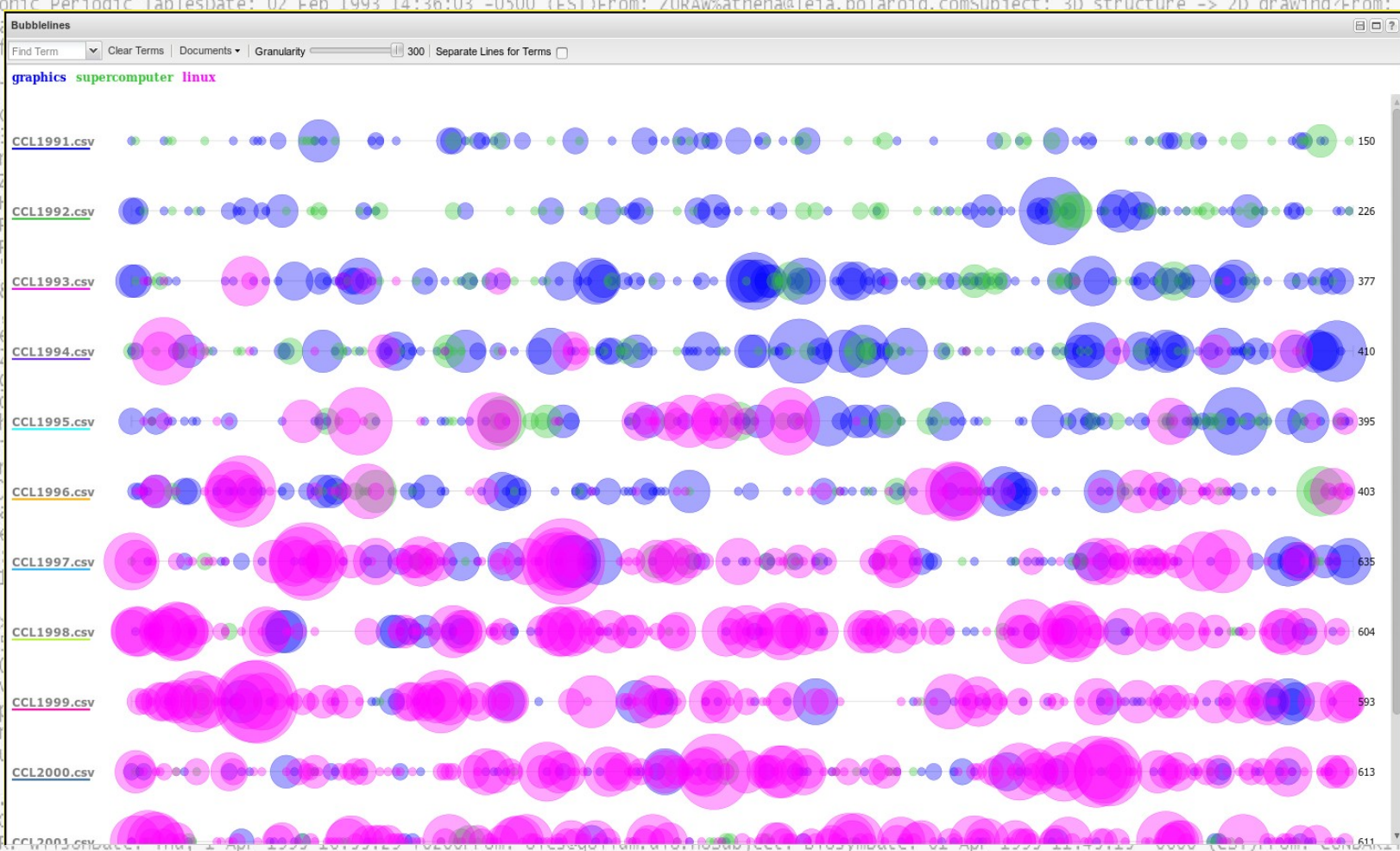
« graphics » and « supercomputer » : époque du supercomputing

« linux » : époque du desktop computing

→ évolution des débats dans la liste

exemple du hardware : disparition de "graphics" et "supercomputer" au milieu des années 90, corrélativement à l'apparition de "linux"

-> passage de "supercomputing" à "desktop modelling"



Small data mining: Text mining

Four of the six threads highlighted in 2001:

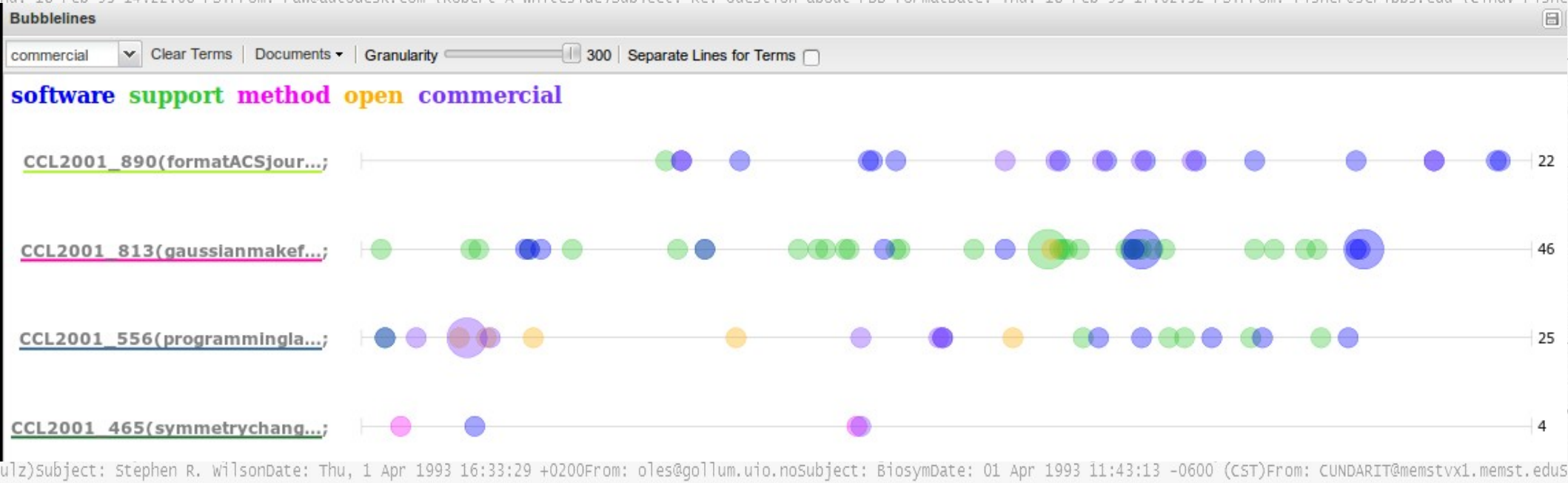
Ici une ligne correspond à un thread plutôt qu'une année

Classer les fils (ici ceux de l'année 2001) par keywords

"software", "commercial", "support" or "open" permet :

de révéler ceux qui sont les plus intéressants

- de les catégoriser par style de topic



1993 : débuts de la CCL : étiquette prohibe flammes
war (la bandwidth est rare) -> ce fil est la première
de l'histoire, contexte de communauté scientifique
en train de se définir, de stratégies éditoriales,
commerciales, de stratégies bureaucratiques
nouvelles -> qui font réagir.

2001 : champ bien établi, a traversé plusieurs
crises (liées au marché de l'emploi et à l'apparition
d'autres promesses technoscientifiques, comme la
bioinformatique), un logiciel particulier (Gaussian)
accapare l'attention par son succès et les débats
qu'il provoque

Plan

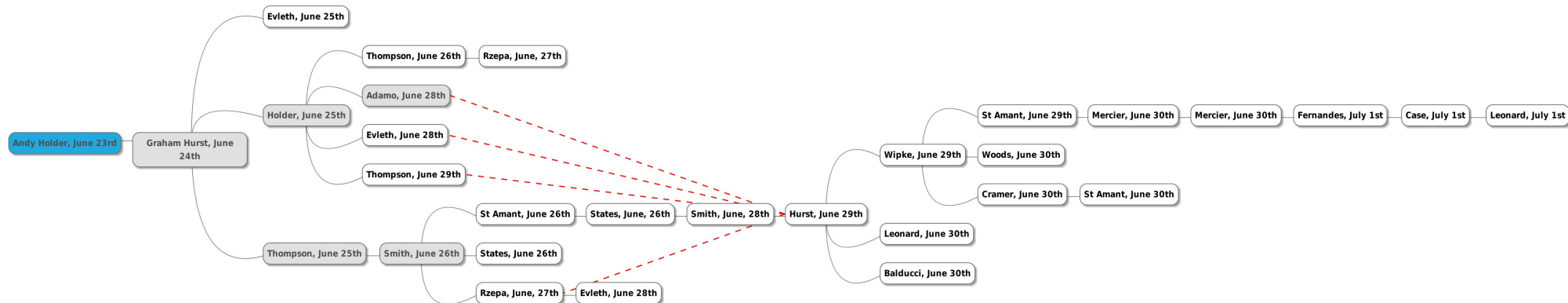
1. Introduction

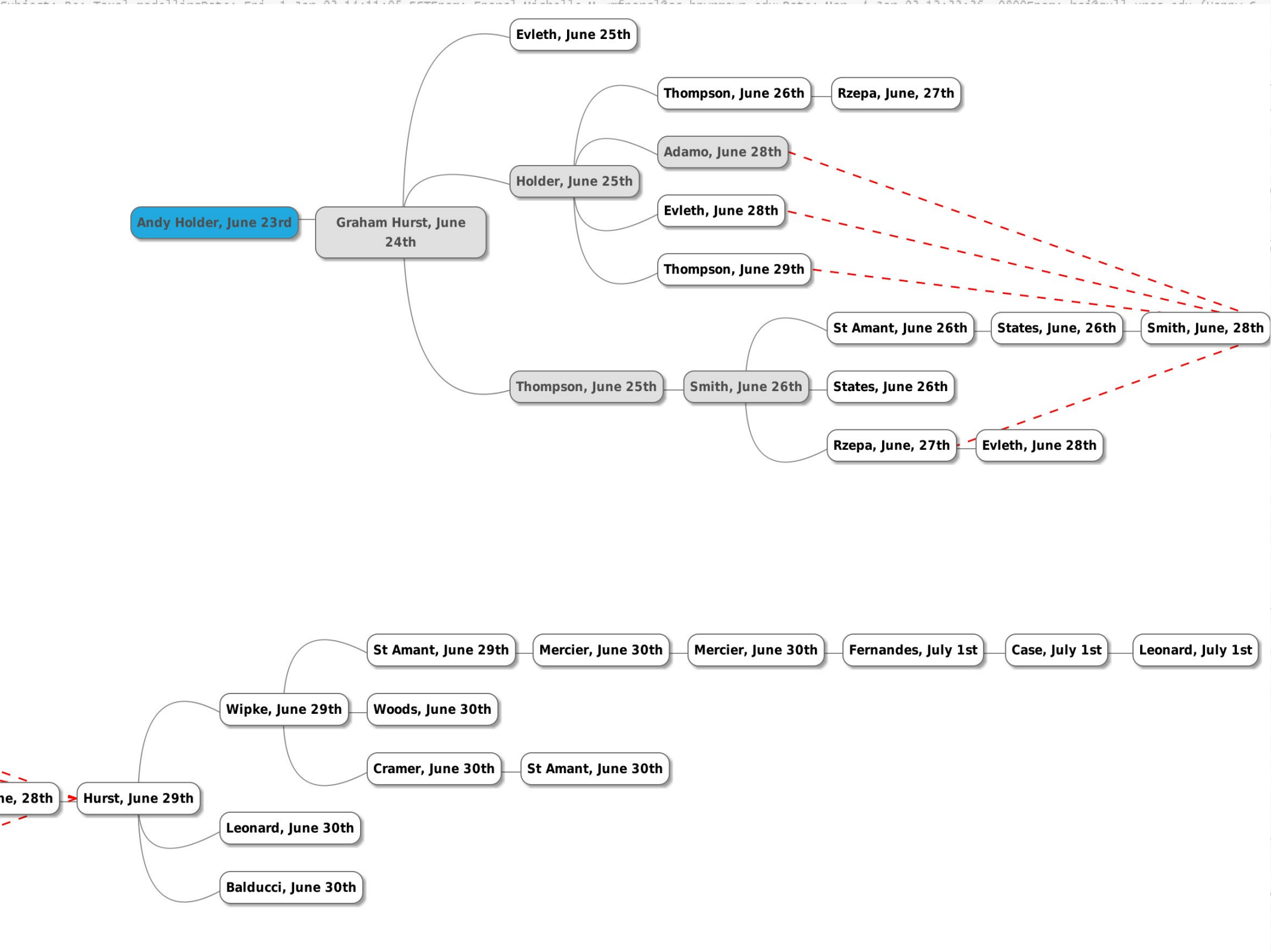
2. Chimie computationnelle : éléments d'histoire

3. CCL : liste de discussion, exploitation du corpus

4. Débats et tensions dans les deux fils choisis

Structure de l'enfilade de Juin 1993





Fil de discussion 1993

- du 23/06/93 au 12/07/93

- 29 messages

- 18 contributeurs

- commence par un message anodin : Andy Holder annonce la publication d'un article donnant les résultats obtenus, pour différents systèmes moléculaires, avec une nouvelle méthode semi-empirique de chimie quantique « SAM1 »

SAM1 reference

- *From:* Andy Holder <[AHOLDER%!
!% VAX1.UMKC.I](mailto:AHOLDER%!
!% VAX1.UMKC.I)>
- *Subject:* SAM1 reference
- *Date:* Wed, 23 Jun 1993 17:44:29 -0500 (CDT)

Hi netters. For those of you interested, Michael Dewar's new SAM1 method now has an official reference. It is as follows:

Dewar, M. J. S.; Jie, C.; Yu, G. Tet. 1993, 23, 5003.

This is primarily a listing of results for the new method for a vast array of systems. It also compares quite favorably with AM1 and PM3. A more complete paper describing the model will be forthcoming from my lab in the next year. Professor Dewar felt it important to

Fil de discussion 1993

- commence par un message anodin : Andy Holder annonce la publication d'un article donnant les résultats obtenus, pour différents systèmes moléculaires, avec une nouvelle méthode semi-empirique de chimie quantique « SAM1 »

- troll : Graham Hurst : « this post disturbs me... »

- problème : « it will be impossible to independently reproduce these results [...] If the method has not yet been published, then the results should not have been accepted for publication since they cannot be verified. [...] Is the SAM1 method sufficiently described in this paper (or publications it references) to allow independent verification of the reported results? »

Fil de discussion 1993

- problème de possibilité de reproduire les résultats car modèle non décrit ; problème de vérification indépendante, problème d'éthique de publication

- la discussion se développe en suivant ces thèmes : tensions, positions, interrogations des acteurs émergent

- dynamique de la discussion : qui répond à qui (voir schéma 1)

- contenu discussion : les thèmes, entremêlés, qui se développent

Fil de discussion 1993

Trois directions (trois premières réponses à Hurst)

- vérifier : possibilité de reprogrammer soi-même

- travail scientifique (ouvert et reproductibilité) VS implication dans entreprise (développement de logiciels)

- éthique publication : importance peer-review ; les résultats de méthodes propriétaires ne font pas partie de « open scientific literature »

>>> ces trois directions sont ensuite développées

Fil de discussion 1993

Un message pivot : Douglas Smith, 26/06/93

- répond point par point à Thompson : monde idéal / réel

- T. : "good science is that of reproducibility and independent verification" OK
mais rarement suivi : paramètres en MM, nota MM2, MacroModel, Hyperchem...

- T. : "I would never have recommended such work for publication" OK "real
problem and a good argument for standardization" mais "utopian and most likely
not practical" car "codes and the proprietary nature of commercial software"

- "Besides, who ever said we had to reveal all our secrets and make them
readily available and accessible? When software copyrights and patents really
provide adequate protection, maybe I will agree with that attitude"

- T. : "results of proprietary methods do not belong in the open scientific
literature" ALORS : "where do they belong? Comparison of these results with
"standard" and commonly available "academic" results is healthy and stimulating.
And not to tweak Mark Thompson, who freely distributes Argus, what about
Gaussian? [accès ou pas au code source]

Fil de discussion 1993

Un message pivot : Douglas Smith, 26/06/93

- manque transparence méthode (paramètres)

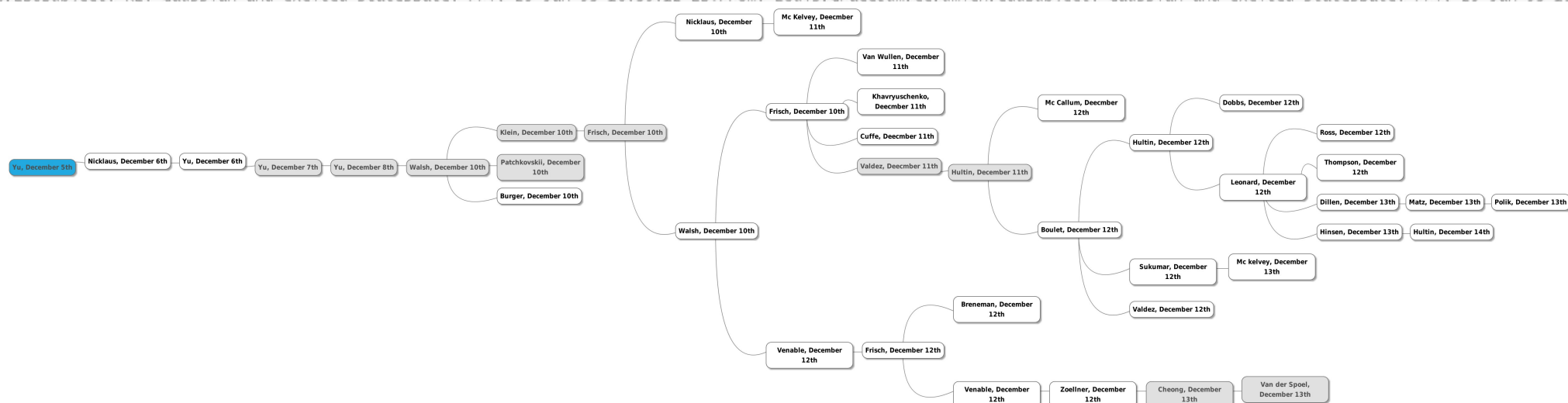
- logiciels propriétaires : diffusion et code source

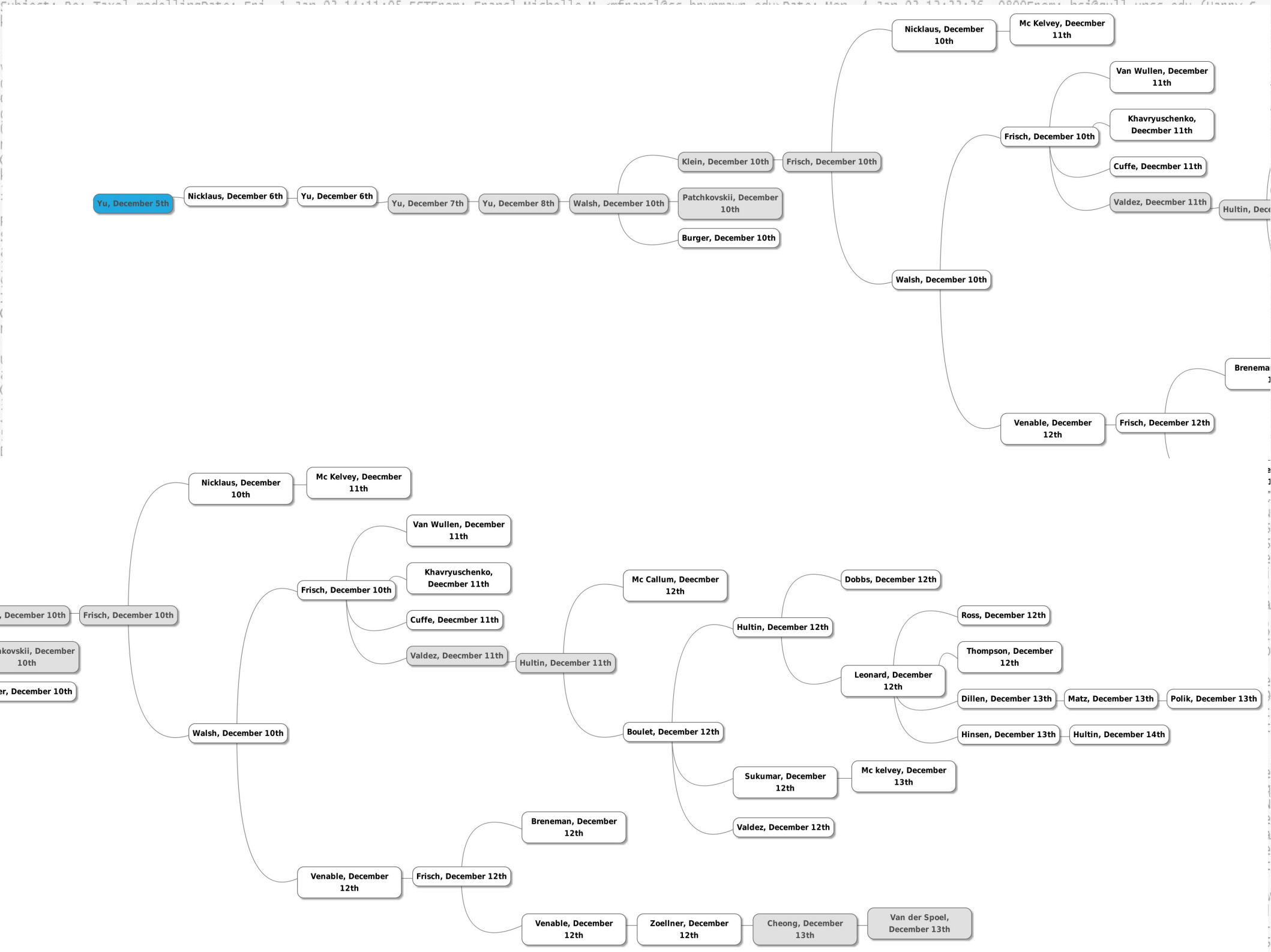
- protection et rétribution travail de programmation

>>> vers questions sur « software copyrights and patents » ; « actual trend in commercializing computational packages without source codes »

>>> message de synthèse de Hurst, apaise la controverse

Structure de l'enfilade de Décembre 2001





Fil de discussion 2001 : GAUSSIAN

- du 8/12/01 au 14/12/01

- 45 messages

- 33 contributeurs en une semaine (rythme + élevé que fil 93)

- annonce d'un « benchmark » (benchmark pour PC : époque où on bricole les machines - desktop computing)

- enchaînement des premiers messages (voir schéma 2)

- problème (4^{ème} et dernier message de Yu) : « We have got the information from Gaussian Inc. that distributing the modified version of makefile or the instructions is violation to the licence agreement. »

- déroulement : ce que la licence d'un logiciel scientifique peut ou doit permettre ; fil s'envenime par l'intervention du président de Gaussian, Inc. (M. Frisch).

Fil de discussion 2001 : GAUSSIAN

- ouverture du code source en tant que garantie de reproductibilité : tensions avec la licence, différences de Gaussian avec le monde du libre

- Gaussian est un logiciel propriétaire qui laisse son code source lisible (par souci de transparence épistémique) mais la licence est TRES restrictive :

- pas d'initiative dans la compilation

- pas de publication d'output

- on ne peut compiler Gaussian que sur le matériel prévu par Gaussian (argument garantie de reproductibilité) au détriment du benchmarking par la communauté scientifique.

- performance du calcul (software) et reproductibilité (méthode) sont entremêlées

- refus de la contribution de la communauté // question du support par la communauté

- différentes catégories d'utilisateurs (lead users, end users...)

Citation du titre

« a LOT of little secrets and "bugs" »

« the support is not so good »

computer whiz VS experimentalist

people [...] prone to jump on Gaussian »

« to make the program work the way the manual implies that it should »

« a user interface that is actually user-friendly »

Re: Gaussian Flame-wars

From: "Phil Hultin" <hultin.,at,.cc.UManitoba.CA>

Subject: Re: Gaussian Flame-wars

Date: Tue, 11 Dec 2001 16:52:36 -0600

Max Valdez wrote:

>I think Gaussian is a great tool, but it has a LOT of little secrets and
>"bugs", and the support is not so good, I don't think the spend
>that much money on customer support, given the price of the product.

I absolutely agree with this statement. I think I represent a new kind of scientist working with tools like Gaussian. I am an experimentalist, not a computer whiz, [...]

So what's the point of all this? Well, maybe people would be less prone to jump on Gaussian (as they have over this makefile thing) if they didn't feel that the philosophy of Gaussian Inc. was similar to that of the "high priest" - only the initiates will have the secrets revealed and then only after years of study. A nice start would be to make the program work the way the manual implies that it should (or to make the manual describe how the program actually works). Even better: bring the program into the 21st century by providing a user interface that is actually user-friendly. What the priests and initiates regard as "a philosophy and aesthetic of simplicity" looks quite otherwise to many of us on the outside!

Dr. Philip G. Hultin
Associate Department Head and
Associate Professor of Chemistry
University of Manitoba
Winnipeg, MB, Canada R3T 2N2
(vox) 204-474-9814
(fax) 204-474-7608
<mailto:hultin.,at,.cc.umanitoba.ca>
<http://www.umanitoba.ca/chemistry/>

Fil de discussion 2001 : GAUSSIAN

- la question de la maintenance (Frisch) avive les tensions : logiciels sont "market oriented" et sont maintenant dans un monde qui suscitent des attentes de clients : le fil embraye sur le sujet de la maintenance (ou absence de maintenance), payante (ou pas), de qualité (ou pas)

- ceci conduit à discuter la question de l'accessibilité (Hultin) : « user-friendliness », « black box » >>> QUI utilise les logiciels (différents types d'utilisateurs : développeurs ou utilisateurs)

>>> différences entre ouverture et accessibilité, compétences nécessaires pour modéliser, parallèle avec instruments de mesure

- empowerment of users

Comparaison fil 1993 et fil 2001

- Tensions : partent de la méthode en 1993, partent du logiciel en 2001

- Articulation logiciel avec : vérification et publication en 1993, hardware et maintenance (utilisateurs) en 2001

- Copyright, brevets, licences : en 1993, des questions se posent (incertitude), diversité des prises de position ; en 2001, conversation sur un plan juridico-technique (invocation de la propriété intellectuelle, de la licence) et industriel (métaphore de la voiture)

- Gaussian : un classique mais n'est pas central en 1993 ; l'objet de la discussion en 2001

>>> Sentiment d'injustice de Gaussian
Gaussian est très préoccupé par la transparence (code source), mais la critique ne se fait plus sur la transparence (d'où l'absence de références à la reproductibilité et la publication) mais sur les formes d'utilisation possibles et permises (benchmark, licence, maintenance, user-friendliness)

Bilan

Multiples tensions :

- au niveau de la construction des logiciels

- de leur distribution

- de leur maintenance

- de leur utilisation

Questions des acteurs :

- sur ce que peut et doit être l'activité scientifique (normes

« idéales » : idéal d'ouverture de la science, transparence,

vérification et reproductibilité, publications et peer-

reviewing, productions scientifiques comme bien commun)

- sur les rapports entre méthodes et logiciels

implémentant

- sur la forme de rétribution du travail de codage

- sur le brevet, la propriété intellectuelle, la licence

- sur la marchandisation des produits de leurs recherches

Bilan

Multiples tensions :

- au niveau de la construction des logiciels

- de leur distribution

- de leur maintenance

- de leur utilisation

Questions des acteurs :

- sur ce que peut et doit être l'activité scientifique (normes

« idéales » : idéal d'ouverture de la science, transparence,

vérification et reproductibilité, publications et peer-

reviewing, productions scientifiques comme bien commun)

- sur les rapports entre méthodes et logiciels

implémentant

- sur la forme de rétribution du travail de codage

- sur le brevet, la propriété intellectuelle, la licence

- sur la marchandisation des produits de leurs recherches

Perspectives au-delà de l'étude de cas :

- ouverture comme valeur problématique : dépend contexte

et arène

- biographies de logiciels (autres sources)