university of toronto

computing services
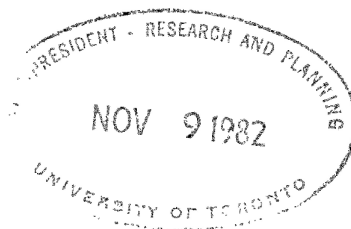
November 5th, 1982

Prof. D. Nowlan,
Vice-President (Research
and Planning) and Registrar,
Simcoe Hall,
University of Toronto.

Dear Prof. Nowlan:

I have participated in several discussions centred on the DOE, RIM, REED and other projects in the Humanities. Most of these groups are attempting to spawn small VAX-11 or VAX-730 systems to support their work. Some indication of cooperative ventures has now surfaced. While it is reasonable for more advanced users to move into these small support systems, I feel the Humanities will require a lot of handholding.

Given the interest that you have expressed in the Humanities and Ian Lancashire's eloquent pleas for support, I would suggest the following as a reasonable Humanities support facility:

1) A VAX-750 or -780 with appropriate storage and peripherals;

2) Software, such as a database package, concordance packages, editors, SCRIBE for formatting, preview devices for formatted text, etc.;

3) A location, such that most of the users could be hardwired at reasonably high speeds;

4) Managed by UTCS, partially subsidized by the University, with real costs reflected to the users;

5) If the remaining LIBRA VAX's are freed up, this would be a good role for one of them.

Perhaps the Review Board would like to think about this.

Sincerely yours,

Alla Heyworth,
Associate Director.

AH:rg

DISTRIBUTION LIST

UNIVERSITY OF TORONTO

Prof. J.Z. Buchwald
Mr. A. Heyworth
Dr. W.C. Jackson
Principal A.F. Johnston
Prof. I. Lancashire
Prof. D.M. Nowlan
Dr. E.V. Swenson

IBM

Mr. M. Brothers
Mr. H. Leiserson
Mr. J. Leppik
Mr. P. LeSeach

MINUTES OF THE MEETING OF THE ADVISORY COMMITTEE

for the IBM/U of T COOPERATIVE AGREEMENT

2 June 1983

Present: Mr. M. Brothers, Prof. J.Z. Buchwald, Mr. A. Heyworth, Dr.
          W.C. Jackson, Principal A.F. Johnston, Prof. I. Lancashire,
          Mr. H. Leiserson, Mr. J. Leppik, Mr. P. LeSeach, Dr. E.V.
          Swenson (Chair)

Regrets: Prof. D.M. Nowlan

1. Needs of Researchers in the Humanities

     Prof. A.F. Johnston, Principal of Victoria College and General
Editor of the REED (Records of Early English Drama) Project and
Prof. I. Lancashire, chairman of the Natural Languages Processing
Steering Committee were invited to describe to the Committee what
researchers in the humanities need by way of computer support. A
summary of Prof. Lancashire's presentation appears as an Appendix.

     The areas in need of support were:

     - data collection and communication links with Oxford and U.S.
universities that have established databases of text,

     (Oxford University - Old English Literature, University of
     Chicago - French literature, Rutgers)

     - data base packages that would assist researchers to organize
and analyze text,

     - ability to handle odd characters and symbols on the terminal
keyboard, screen and printer,

     (IBM has a centre in Toronto that has developed equipment that
     can handle up to 12,000 unique characters. It was also mentioned
     that Sanyo, Osborne and an IBM PC with a graphics board have
     redefinable character generators that will display and print the
     new characters.)

     - printing of high quality documents for publication

     - various software packages, e.g. one that can parse English

     (Researchers are interested in the process of machine translation
      rather than in its product.)

     - improvement of the interface between the researcher and the
microcomputer,

Mr. Brothers will review what IBM can offer in these areas, especially application packages under VM such as SQL.

2. Plans for the Use of the 3033 U16

Action items proposed by M. Brothers, W. Jackson and H. Leiserson with respect to the use of the 3033 U16 were reviewed.

W. Jackson will serve as the primary VM/CMS contact at UTCS. UTCS is drafting a document that will describe the constraints on CMS service. It has also acquired software for back-up purposes and for controlling the use of resources on the system.

H. Leiserson is arranging a trip around the 15th of June. The purpose of the trip would be to understand the operation and management of a CMS service. Possible destinations are the U. of Pennsylvania which has a 4341 running VM/CMS, Princeton which has a 3081 with VM/CMS, and IBM Yorktown Heights. W. Jackson, another UTCS member and J. Buchwald may go.

The feasibility and advisability of introducing PROFS was discussed. The present configuration of the 3033 U16 will not support the addition of PROFS as an offering. There are no ports available. M. Brothers reported that G. Lewis of IBM will be examining further the technical implications.

The advisability of promoting PROFS was also considered problematic. It was considered to be unrelated to the teaching and research purposes stated for the 3033 U16. It was suggested that PROFS should be exchanged with other relevant packages.

A. Heyworth and I. Lancashire stressed that the need that should be addressed is the support required to establish a network of personal computers including communications, a data base facility and a file storage package. J. Leppik offered to investigate whether PROFS was the best offering that IBM had to address this need.

3. Press Release

It was suggested that the press release on the progress of the IBM/UT Cooperative Agreement be prepared for publicaiton in the campus Bulletin in September so as to gain wider exposure.

4. IBM and U.S. Universities

This presentation by P. LeSeach will be given at the next meeting.

5. Next Meeting

The next meeting has been scheduled for Thursday, 23 June 1983 at 9:30 a.m. in Room 245, Simcoe Hall, University of Toronto.
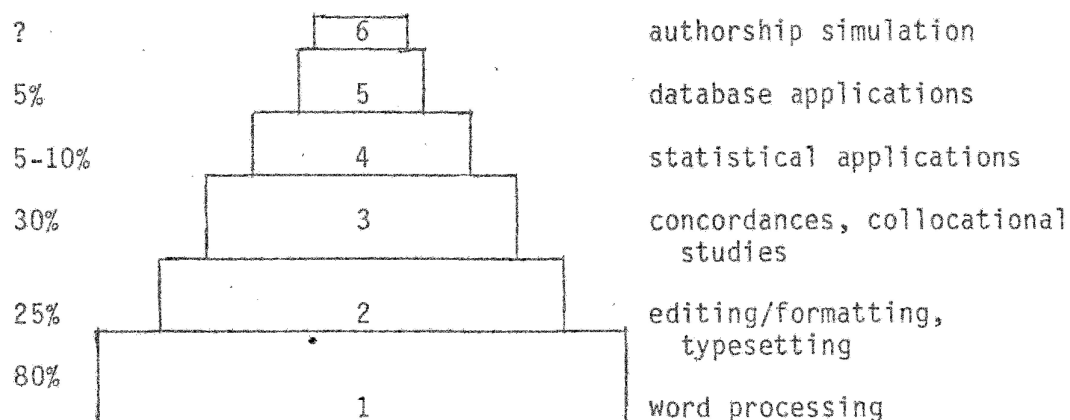
An Overview of Activities of Researchers

in the Humanities that could Benefit

from Computerized Support

presented by Professor I. Lancashire

% of humanists involved in activity	Activities

%	Level	Activity
?	6	authorship simulation
5%	5	database applications
5-10%	4	statistical applications
30%	3	concordances, collocational studies
25%	2	editing/formatting, typesetting
80%	1	word processing

The community has about 1000 members composed of 600-700 faculty members and 300-400 students. There are about 30-40 active users of computerized support tools at present.

Hardware Needs of Each Activity

>   Level 1, some of level 2, a bit of level 3 - this work can be done on a micro.

>   Levels 2 - 6 - this work requires a mainframe

Examples of Software Used by Each Activity

>   Level 2 - WYLBUR, SCRIBE, SOS, TECO, RUNOFF

>   Level 3 - COGS, Oxford Concordance Package (OCP), ARRAS (interactive concordance which spans levels 3-5).

>   Level 4 - STATS (written in SNOBOL), CLUETT

>   Level 5 - FAMULUS (in FORTRAN), MARK IV

Research Areas that would involve participation in the above
activities

Area	Would Use Level	
A	1-3	history of ideas, literary criticism
B	1-4	explication of text
C	1-5	editing text for historical editions, indexes, lexicography
D	1-5	biographies, bibliographies, "literary" history
E	1-5	stylistics (what differentiates one author from another, syntax, style, vocabulary), study of the growth of a language
F	1-6	literary theory, linguistics, cognitive psychology, study of how people write

D76
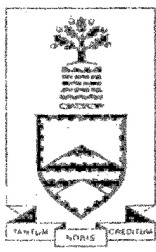
25 September 1983

Professor David Nowlan
Vice-President and Registrar
Simcoe Hall
University of Toronto

Dear David:

Mike Brothers of IBM has offered to arrange for a one-week visit
by Toronto computing humanists and a member of SSHRCC to computer
centres in San Jose, Palo Alto, Stanford, and Berkeley. He has
also offered to put us up at the IBM ranch. You will see from
the attached agenda what he has in mind for us all. With a few
additions (artificial intelligence work at Stanford occurs to
me), this looks like an excellent and exciting set of offerings.
The suggested dates are Oct. 24-28.

We will need travel funds to take advantage of IBM's proposal.
I would think that we should send half a dozen people: three
humanists, Al Heyworth, someone from SSHRCC, and a member of
UTCS. There is no reason why UTCS should not cover this trip
from its own travel budget, but my own department has no travel
funds at all from the faculty (in fact I still have not been
reimbursed for a conference trip I took last June); and this
state of affairs probably holds as well for other humanists and
Al Heyworth. SSHRCC is not willing to spend any budget on such a
trip: it is going to be difficult enough to persuade it to send
a delegate (and I believe IBM is keen to have an SSHRCC person
along). SSHRCC is woefully ignorant in this area. It has refused
to fund several of our major projects in their computing needs, and
seems to know little about what UTCS is now doing to assist SSHRCC-
funded projects with programming.

Travel for five people will cost about $4250.00. We could save
something of this by asking everyone to stay on for two more days,
but then accommodation would likely eat up the savings. If the
university has funds for such purposes, this case seems deserving.
We should get back to Mike Brothers by Friday if at all possible,
since he will want to make arrangements.

I discussed this proposal at last Friday's meeting of the Natural
Language Processing Steering Committee, and Eva Swenson suggested
that I should write to you.

Best wishes,

Sincerely,

Ian Lancashire

HUMANITIES TECHNOLOGICAL APPLICATIONS TOUR

FOR

THE UNIVERSITY OF TORONTO

Tentative Agenda.

MONDAY OCTOBER 31st.

Travel to San Jose   Stay at IBM Homestead.
  (American Airlines   5750 Cottle Road,
  563-Chicago   San Jose, CA 95123,
  231-San Jose)   (408-225-3131)

TUESDAY NOVEMBER 1st.
(IBM Research Lab - San Jose, 5600 Cottle Road. Conf Rm 2C-012)

9:00 - 9:05	Introduction (Mike Brothers - Manager, Coop Pgms)
9:05 - 9:20	IBM General Products Div. (Dave Fitzpatrick- Manager of Marketing Programs)
9:45 - 11:00	Storage Technology Trends. (Dave Fitzpatrick- Manager of Marketing Programs)
11:00 - 12:00	Review of Research Activities at San Jose (Pat Mantey - Senior Manager - Experimental Systems) *Warren Edwards*
12:00 - 1:00pm	Lunch
1:00 - 2:30	Trends in Printer and Display Technology (Pat Mantey - Senior Manager - Experimental Systems)
3:00 - 4:00	Text Processing Research and Development (Pat Mantey)
4:00 - 4:30	IBM Visiting Professors Program (Fred Lochovsky)
4:30 - 5:30	EZ draw using the IBM PC (K.C. Chu)
5:00	Finish

HUMANITIES TECHNOLOGICAL APPLICATIONS TOUR

FOR

THE UNIVERSITY OF TORONTO

Tentative Agenda.

MONDAY OCTOBER 31st.

Travel to San Jose   Stay at IBM Homestead.
  (American Airlines   5750 Cottle Road,
  563-Chicago   San Jose, CA 95123,
  231-San Jose)   (408-225-3131)

TUESDAY NOVEMBER 1st.
(IBM Research Lab - San Jose, 5600 Cottle Road. Conf Rm 2C-012)

9:00 - 9:05	Introduction (Mike Brothers - Manager, Coop Pgms)
9:05 - 9:20	IBM General Products Div. (Dave Fitzpatrick- Manager of Marketing Programs)
9:45 - 11:00	Storage Technology Trends. (Dave Fitzpatrick- Manager of Marketing Programs)
11:00 - 12:00	Review of Research Activities at San Jose (Pat Mantey - Senior Manager - Experimental Systems) _Warren Edwards_
12:00 - 1:00pm	Lunch
1:00 - 2:30	Trends in Printer and Display Technology (Pat Mantey - Senior Manager - Experimental Systems)
3:00 - 4:00	Text Processing Research and Development (Pat Mantey)
4:00 - 4:30	IBM Visiting Professors Program (Fred Lochovsky)
4:30 - 5:30	EZ draw using the IBM PC (K.C. Chu)
5:00	Finish

WEDNESDAY NOVEMBER 2nd.

IBM Research Facility, San Jose, 5600 Cottle Road.Conf. Rm 2C-012)

8:30 - 10:30   Data Systems Trends and Directions
                (Sharon Weinberg-Senior product Consultant)

10:30 - 11:45   Data Base Trends
                (Raymond Lorie?-)

11:45 - 1:00   Lunch - and Drive to IBM Scientific Centre(conf Rm)
                             1530 Page Mill Road,
                             Palo Alto,
                             (415-493-4700)

1:00 - 1:15   Welcome (Dr. H. Kolsky -)

1:15 - 2:30   Graphics & Image Processing
                (Ralph Bernstein - Scientific Staff Member)

3:00 - 4:30   IBM/University project overview.
                (Paul Friedel Palo Alto Sc. Cent.)

4:30 - 5:30   Expert Systems (Peter Hirsch- Mgr. Expert Systems)

4:30   Finish. Return Homestead.

THURSDAY NOVEMBER 3rd.

Travel to Stanford Research Institute (SRI)

9:00 - 9:05   Introduction - (Hector Leiserson - Senior Ind. Spec.)

9:05 - 10:00   Computational Linguistics (Stanford Research Inst)
               (Prof. C. Rae Perreault)

10:00 - 10:20 Travel to Stanford University

10:20   Meet with Maria Bunn (or her rep.)
               Meeting Coordinator, Office for International Visitors

10:20 - 10:30 Walk to assigned rooms

10:30 - 11:00 Panel discussion: " Humanities Computing at Stanford"
               Coordinator: Randy Mellen

11:00 - 12:00 T.I.R.O. Project or "SMILE"
               "Stanford Microcomputer Instructional Lab. Experiment"
               (Randy Melen, Project Coordinator)

12:00 - 1:30   Lunch at Faculty Club

1:30 - 3:30   Humanities Computing at U of T.
               (Prof. Ian Lancashire - Prof. of English)

4:00 - 5:00   Stanford University Academic Computing Strategy.
               (Prof. Bob Street, Vice-Provost of Academic Computing)

5:00   Finish
               Return to Homestead.

FRIDAY NOVEMBER 4th.

Travel to University of California at Berkely- 331 Wheeler Auditorium

9:00 - 10:15   Computing in the humanities at UC Berkley
                   (A. Hernandez and K. Nelson)

10:15 - 10:25   Move to Tolman Hall

10:25 - 11:30   PROFS at UC Berkeley
                   (J. Smith)

11:30   Finish

Travel to Toronto.(Air Canada 756 - Tor. LV 1:15   ARR 8:50)

Proposal for a Humanities Computing Facility

This proposal concerns a computing facility for humanists engaged in graduate research, teaching and text development in about twenty departments, centres and programmes at Toronto. A large IBM mainframe, accessed at low cost by terminals and microcomputers in local discipline work areas, offices and homes, would serve as the hub of a polyglot text processing complex, linking (a) graduate courses teaching new techniques in stylistic, linguistic and content analysis, (b) international research and editorial projects in the Humanities in English, French, Classics, Linguistics, Philosophy, Medieval Studies, Near Eastern Studies and other departments and centres, and (c) individual faculty members undertaking research in Humanities fields.

Toronto is extremely rich in high-level research in the Humanities and has a long history of advanced work in computer applications in these disciplines, but its humanists have also been poor in funds and equipment for computational research. A few examples: Classics has a complete corpus of ancient Greek and Roman texts but no mainframe access; a graduate course in French cannot be offered two years in a row for lack of funds; many faculty engaged in text analysis use personal microcomputers for word-processing but cannot afford to exploit the large, powerful programs developed over many years here to process large texts. Support from IBM would enfranchise many researchers in Humanities computing applications and make possible implementation of exciting new programs that promise to enhance greatly the production of scholarship and criticism.

The general purposes of this facility are to enhance research at Toronto and to share our programs, techniques and texts using the same method, publication, by which we disseminate our other research.

Here are examples of what we would like to do with such a facility.

1. Make available COGS, a polyglot lemmatizing concordance program written here for the IBM 3033 by John Bradley and modified by the UTCS Humanities programmer, Lidio Presutti, as well as a suitable manual, for use at other institutions. (COGS runs six times as fast as the Oxford Concordance Package, the most widely-available package at present.)

- 1 -

2. Use COGS as a basis to create a new system for text analysis of literary concepts and themes (TACT). A mainframe-based program would produce lexical databases that researchers could use on mainframes or download to microcomputers. These techniques, which rely on new developments in textual graphics and collocation, have wide applications in education.

3. Undertake advanced work in analysis of English syntax with IBM's EPISTLE. This program, after testing, should prove helpful in increasing the productivity of English writing and composition courses. Linked to statistical-analysis systems, EPISTLE would supply a needed tool in synchronic and diachronic studies of the English language.

4. Develop an excellent SNOBOL/SPITBOL facility for advanced programming in text analysis.

5. Make available existing programs from other institutions that help humanists with textual collation, the determination of manuscript trees, and source analysis (through text-comparison programs).

6. Explore the use of image processors (such as the experimental IBM Image Editor) in editorial identification of compositors in texts produced by early printing shops and of scribes in manuscripts.

7. Develop a manual in Humanities computing applications for use with the computing facility. (Last September a group wrote such a book for use with the DEC-10 system.)

8. Function as a local centre for computing humanists in this region, hosting, for example, colloquia or visits by leaders in the field and local researchers. Toronto humanists now work with colleagues at McMaster, Western Ontario and York Universities.

9. Extend our good relations with other Humanities computer facilities, as at Oxford and Chicago. (This year Toronto contributed STATS, a textual statistics program, to the Oxford University Computing Center, and the French Department accessed the ARTFL system at Chicago. More interchanges of this sort are very desirable.)

10. Allow local researchers to contribute to international efforts to develop a library of well-edited machine-readable texts.

Our hardware needs are for a self-contained facility: a mainframe and peripheral devices including storage, tape backup and communication links, a service printer as well as a

high-quality printer (producing copy suitable for publication),
additional terminals for department and centre offices and
perhaps libraries such as Robarts, Fisher and Pontifical
Institute, several microcomputers (for program development), and
graphics plotter. Software needs include the Image Editor and
EPISTLE. Technical support will be as important here as it was
in the last IBM-Toronto agreement. IBM participation in
colloquia and research visits would also be invaluable.

Ian Lancashire
11 June 1984

12 June 1984

Professor David Nowlan
Vice-President (Research and Planning)
Simcoe Hall

Dear David:

I understand that you will be meeting with representatives
from IBM to discuss ways in which IBM can lend support to
university research in the future. In the hope that the
Humanities may have some role in this extended partnership,
I have written the enclosed "Proposal for a Humanities
Computing Facility." It is intended to be an exploratory
statement of some directions that computing humanists would
like to follow in the near future.

I have also sent a copy to Mike Brothers for his
information.

                              Sincerely,

                              Ian Lancashire
                              Professor of English

Proposal for a Humanities Computing Facility

This proposal is to set up a dedicated computing facility for humanists engaged in graduate research, teaching and text development in about twenty departments, centres and programmes at Toronto. An IBM mainframe at an Arts and Science Humanities computing site, accessed by terminals and microcomputers from a central site, departmental work areas, and faculty offices and homes, would serve as the hub of a multi-lingual text processing and logic-analysis complex, linking (a) courses teaching symbolic logic, statistical uses of data bases, and linguistic, stylistic and content analysis, (b) international research and editorial projects in the Humanities in English, French, Classics, Linguistics, Philosophy, Near Eastern Studies and other departments and centres, and (c) individual faculty members undertaking research in all Humanities subjects.

Humanities research should be interpreted broadly as including much current work in Faculties such as Library and Information Science, and Music, and in special centres and programmes of the School of Graduate Studies (such as Medieval Studies).

Here are examples of what we would like to do with such a facility.

1. Teach undergraduates, and do graduate research on symbolic logic and theorum proving (a pilot project for 1985-86, about 100 students in PHL 245H [Symbolic Logic] using BERTIE, a program developed at Dartmouth College; research in automatic theorum-proving by A. Urquhart).

2. Teach graduate students in History (especially in HIS 1002H "Introduction to Historical Methods") the use of statistical analysis of databases for descriptive purposes in filling out historical argument; and provide historians with a general research facility.

3. Do advanced work in analysis of English syntax with IBM's EPISTLE (a pilot project for about 100 students in English composition courses in 1985-86). EPISTLE would also be useful in research in natural-language processing and, linked to statistical-analysis systems, would provide a needed tool in studies of the English language.

4. Make available editorial utilities for advanced scholarship by editors of bibliographies, literary texts, concordances, journals, dictionaries and monographs. Provide a general-purpose database management system for text

- 1 -

applications, preferably one that works on both a mainframe
and micros.

5. On the basis of COGS, a multi-lingual lemmatizing
    concordance program running on the 3033 and written at
    Toronto, develop an interactive ARRAS-like system for text
    analysis of concept and theme (TACT), one that could
    support advanced word-distribution graphing and
    collocational analysis on both mainframe and
    microcomputers.

6. Give humanists good text-processing facilities for the
    writing and publication of articles, books and theses.

7. Put researchers at Toronto, especially those in SSHRCC
    projects, in touch with their colleagues in other cities by
    means of BITNET-NETNORTH and other networks to facilitate
    exchange of such written research.

8. Put in place input and output devices for effective
    handling of Hebrew and Greek character sets for disciplines
    such as Near Eastern Studies.

9. Make a good LISP and SPITBOL facility available for
    Humanities programming in logic and text applications.

10. Develop documentation in Humanities computing applications
    for use with the computing facility.

11. Function as a local centre for computing humanists in this
    region, and cooperate with humanists at other Ontario
    computer centres (e.g., Waterloo, Guelph, York, Queen's) in
    software use and development.

12. Set up a tape library of machine-readable texts at Toronto
    with the aid of an optical scanner, allowing local
    researchers to contribute to an international library of
    well-edited machine-readable texts.

The general purpose of this facility would be to enhance
research and permit teaching innovations at Toronto. A valuable
way to do so is to share the programs, techniques and texts
produced by it using the same method, publication, by which we
disseminate our work generally.

Equipment and Facilities

   An Arts and Science Humanities Computing Centre, managed by
UTCS, with an IBM mainframe for substantial programs such as
COGS, EPISTLE, BERTIE, OCP and a dbms system, accessed by nearby
terminals and microcomputers, organized into
partially-independent local-area networks. This site might be in
the Robarts Library and have both undergraduate and research
rooms.

   A useful model is Waterloo's Micronet, in which groups of PCs
communicate with an IBM mainframe through Series-I peripherals.

   For departments such as Philosophy, English and other
departments, a central site like the Robarts Library makes sense
because they have no local space for a distributed facility, are
already convenient to the Robarts Library, or find most of their
graduate students already doing their research in Robarts Library
carrels.

   Other research centres, however, might well prefer their own
sites, either because they have their own libraries or because
special input-output requirements for character sets would
require unusual software or hardware (e.g., History and perhaps
Medieval Studies, Near Eastern Studies, and Classics).
Departments and centres should be allowed to choose whether they
wish a central or distributed site. A typical configuration
(e.g., for History) would have several PCs clustered about a
PC/AT, itself linked by a Series-I to the mainframe.

   The central site should have a printer, a tape library, a
general-purpose phototypesetter and an optical scanner.

   Humanists will probably need several kinds of operating
systems. UNIX provides many utilities for text analysis by an
experienced programmer (e.g., awk, grep, sort) but, as yet, lacks
sophisticated program "packages" to do multi-lingual concording,
collocation and database work. MVS or CMS, on the other hand,
are necessary hosts for important "word-crunching" packages like
EPISTLE, COGS-3 and OCP operating on very large texts.

   IBM micros running UNIX would provide a useful link with the
many UNIX-based SSHRCC projects in the Humanities (DOE,
Athenians, RIM).

      Ian Lancashire
      2 November 1984

12 November 1984

Ken H. Fockler
Manager
Scientific and Education Programs
IBM Canada Ltd
3500 Steeles Ave East
Markham, Ont L3R 2Z1

Dear Ken:

Last Thursday David Nowlan announced that IBM expressed interest in
natural language processing as one possible area of cooperation
with the university in the near future. My colleagues and I were
delighted to hear that this is so.

When we last talked over that enjoyable dinner with Mike Brothers
and Lance Miller, you mentioned that we would like to hear more about
computing in the Humanities at Toronto. For that reason I have
enclosed two papers for you.

The first is my recent proposal to Professor Nowlan on the possible
directions that this cooperation might take. The important ones are
natural-languages systems like EPISTLE (Graeme Hirst of Computer
Science and myself are most concerned with this), text-analysis
software (my Natural Language Processing Committee and Barron
Brainerd of Mathematic Linguistics), editorial tools (the SSHRCC
projects as a group) and symbolic logic and theorum proving
(John Slater, chairman of the Philosophy Department).

The second is microCHARTIT and documentation on it, a very modest
piece of freeware developed by an Engineering undergraduate and
myself last summer. David Barnard of Queen's has expressed interest
in seeing it and recently I sent a copy off to him. Next month
UTCS will be enhancing the program for teaching purposes.

If I can be of any help to you in sorting through these things,
please let me know.

With all good wishes,

Truly yours,

Ian Lancashire
Professor of English
978-5270 (office)
231-2659 (home)

c. Mike Brothers
   David Nowlan

7 January 1985

Mr Ken H. Fockler
Manager
Scientific and Education Programs
IBM Canada Ltd.
3500 Steeles Ave East
Markham, Ont. L3R 2Z1

Dear Ken:

Our meetings on Tuesday the 15th begin at 9 am in the office of
Principal A.F. Johnston, Room 101, New Academic Building (now
called Northrop Frye Hall, I think), Victoria College. (This
modern building lies on the east side of Queen's Park Crescent
where it turns north into Avenue Road; its office windows face
north and south.) Administrators attending include

> Acting Dean R. Craig Brown (School of Graduate Studies)
> Ms Carol M. Gillin (Office of Research Administration)
> Mr Al Heyworth (from David Nowlan's office)
> Dean Ann H. Schabas (Faculty of Library and Information
> Science)
> Dean J.J.B. Smith (Arts and Science) -- Berrie Smith
> Principal A.F. Johnston (Victoria College) -- Sandy Johnston

Questions about Humanities research and funding can be directed
to these people.

From 10:10 to 11:30 am we will be meeting in the Senior Common
Room, Burwash Hall, Victoria College, with Humanities faculty
involved in Humanities computing. They will be able to describe
their interests and needs. Here are some of those attending.

> Barron Brainerd (Linguistics)
> Jim F. Burke (Spanish and Portuguese)
> Ernie G. Clarke (Near Eastern Studies)
> John Hurd (Religious Studies)
> A.A. Iannucci (Italian) *bi'62 + S&N 926-1350 ext 3295 (962-*
> Christopher Jones (Classical Studies)
> Richard G. Landon (Fisher Rare Book Library)
> Paul Rutherford (History)
> Jack T. Stevenson (Philosophy)
> T. Russ Wooldridge (French)

I'm sure that you will enjoy the opportunity to talk with all these
researchers. They will give you the flavour of the subject better
than the enclosed brief can.

Best wishes,

Sincerely,

Ian Lancashire

Memorandum

From: Ian Lancashire

To: Humanities computing faculty members

Ken H. Fockler, Manager, Scientific and Education Programs, IBM Canada Ltd., will meet with us on Tuesday morning, January 15, from 10:10 to 11:30 am in the Senior Common Room, Burwash Hall, Victoria College, to learn about possible avenues of IBM support for Humanities computing over the next few years. Part of Mr. Fockler's responsibilities include negotiating IBM partnership agreements with Canadian colleges and universities.

You will find enclosed a brief describing one possible scenario, but this has no special status and should not limit discussion.

cc. B. Brainerd (Linguistics)
    J.F. Burke (Spanish and Portuguese)
    E.G. Clarke (Near Eastern Studies)
    J.C. Hurd (Religious Studies)
    A.A. Iannucci (Italian)
    M. Israel (South Asian Studies)
    C.P. Jones (Classical Studies)
    R.G. Landon (Fisher Library)
    P. Rutherford (History)
    J.T. Stevenson (Philosophy)
    T.R. Wooldridge (French)

Proposal for a Humanities Computing Facility

This proposal concerns an editorial and research computing facility for humanists in SSHRCC-funded research projects and about twenty departments, centres and programmes at Toronto. In a proposed Humanities Building to be erected by Victoria and St Michael's Colleges, a large machine, accessed by terminals and microcomputers from a Humanities computing centre, departmental work areas, and faculty offices and homes, would serve as a multi-lingual text-processing, typesetting and logic-analysis complex, linking (a) international research and editorial projects in English, French, Classics, Linguistics, Philosophy, Near Eastern Studies and other departments and centres, (b) individuals undertaking computer-aided research in all Humanities subjects, and (c) courses teaching symbolic logic, statistical uses of data bases, and linguistic, stylistic and content analysis.

Humanities research should be interpreted broadly as including much current work in Faculties such as Library and Information Science, and Music, and in special centres and programmes of the School of Graduate Studies (such as Medieval Studies).

Here are examples of what such a facility could provide.

1. Editorial and typesetting utilities for historical and literary texts, bibliographies, concordances, journals, dictionaries and monographs.

2. Effective handling of Roman and non-Roman (e.g., Hebrew, Greek, Russian) character sets for disciplines such as Near Eastern Studies, Classics, French, German and Slavic.

3. A general-purpose database management system for bibliographical and text applications, preferably one that works on both a mainframe and micros.

4. Access to international networks to facilitate exchange of research information.

5. A tape library of machine-readable texts at Toronto developed with the aid of an optical scanner, allowing local researchers to contribute to an international library of well-edited machine-readable texts.

6. Programs analyzing English syntax and style for English composition courses and for research in natural-language processing.

7. Multi-lingual lemmatizing concordance programs for text
    analysis, word-distribution graphing and collocational
    analysis.

8. Text-analysis utility programs written in SPITBOL and in a
    UNIX environment (C, LISP, ICON) for the teaching of
    Humanities programming techniques.

9. Programs in symbolic logic and theorem proving for PHL 245H
    [Symbolic Logic] and graduate research.

10. Programs for the statistical analysis of databases to
    provide quantitative data that supplements historical
    argument, for use in HIS 1002H and other research.

11. Resources for simulation of large library automation
    systems as part of a new degree program in the Faculty of
    Library and Information Science.

12. Documentation for use with the facility.

13. A local centre for other computing humanists in this
    region.

The general purpose of a facility would be to sustain and enhance
research, permit teaching innovations, and provide economical
alternatives to traditional methods of academic publishing.

Equipment and Facilities

At first, a facility managed by UTCS at a central site, to be
housed eventually, we hope, in the proposed Victoria-St.
Michael's Humanities Building: one or more machines to handle
substantial programs such as COGS-3, Writer's Workbench, BERTIE,
OCP and a dbms system, accessed by nearby terminals and various
commonly-owned microcomputers, organized into
partially-independent local-area networks.

Support would be needed in two stages.

1. Development of editorial and typesetting tools, a
    multi-lingual text editor and formatter, and applications
    in various discipline areas.

2. Integration of this system in a computing facility within a
    Humanities Building.

For Toronto's SSHRCC Humanities projects, now dispersed across the university at the Robarts Library (14th floor), Victoria and St Michael's Colleges, the south-west campus, and other buildings, a central site is desirable to pool resources and build on present faculty experience.

For departments such as Philosophy and English, a central site makes sense because they have no local space for a distributed facility and are already convenient to the federated colleges, or find many of their graduate students already doing their research in the Medieval Studies and Reformation and Renaissance Libraries.

Other units (such as FLIS), however, will prefer their own sites, convenient to discipline offices, classrooms and library. Departments and centres should be allowed to choose whether to rely on a central site or to be self-sufficient.

The central site needs a high-quality printer, tape storage facilities, a general-purpose phototypesetter, and an optical scanner.

Graduate and undergraduate terminal areas should be available.

Humanists need several kinds of operating systems. UNIX provides many utilities for text analysis of small texts but does not handle multi-lingual concording of a large corpus. Current IBM machines have "word-crunching" packages like COGS-3 and OCP, which do well with very large texts.

Possible IBM Assistance

A 4361 mainframe, surrounded by groups of PCs and terminals communicating through Series-I peripherals, would provide a good host for COGS-3, OCP, BERTIE (a symbolic logic package), a database system, and a high-quality formatting system.

If IBM donated the machinery and agreed to maintain it, humanists would benefit from very cheap, powerful computing.

An IBM 4250 would provide high-quality printing, yet not up to the standards of typesetting now being done in projects like REED (a Linotron 100) or RIM (a Compugraphic 8400).

IBM has image scanners (these capture pictures in digitized form) but not, as far as I know, conventional optical scanners for automatic input of text.

At present the most friendly IBM operating system seems to be VM/CMS. IBM does not yet offer UNIX on a mainframe and so could not support utilities like Writer's Workbench, although the PC/AT has XENIX. Communicating PC/ATs might be a suitable way to linking major research projects and departments in the university.

Attractive IBM software includes the Document Composition Facility, EPISTLE (a package in development that corrects faulty grammar and analyses style) and a dbms like STAIRS. PROFS, an environment for business computing that has many comfortable features, would probably not be used widely by humanists, who do most word-processing on micros.

The international office of IBM research and support for natural language applications has been established in Toronto and would probably be of help in implementing software and hardware for handling multiple character sets. (At present Gutenberg, running on Apple II, handles word-processing needs for non-Roman alphabets, and the Athenians project has Greek on a modified UNIX system.) Although currently available IBM equipment at Toronto does not meet our needs, IBM supports all important character sets in the world now and presumably has access to machines sold in Europe, the Middle and Far East, etc., that might do so.

Ian Lancashire
Chairman
Natural Language Processing Steering Committee
7 January 1985

At present the most friendly IBM operating system seems to be
VM/CMS. IBM does not yet offer UNIX on a mainframe and so could
not support utilities like Writer's Workbench, although the PC/AT
has XENIX. Communicating PC/ATs might be a suitable way to
linking major research projects and departments in the
university.

Attractive IBM software includes the Document Composition
Facility, EPISTLE (a package in development that corrects faulty
grammar and analyses style) and a dbms like STAIRS. PROFS, an
environment for business computing that has many comfortable
features, would probably not be used widely by humanists, who do
most word-processing on micros.

The international office of IBM research and support for
natural language applications has been established in Toronto and
would probably be of help in implementing software and hardware
for handling multiple character sets. (At present Gutenberg,
running on Apple II, handles word-processing needs for non-Roman
alphabets, and the Athenians project has Greek on a modified UNIX
system.) Although currently available IBM equipment at Toronto
does not meet our needs, IBM supports all important character
sets in the world now and presumably has access to machines sold
in Europe, the Middle and Far East, etc., that might do so.

> Ian Lancashire
> Chairman
> Natural Language Processing Steering Committee
> 7 January 1985

17 January 1985

Professor David Nowlan
Vice-President, Research and Government Relations
Simcoe Hall

Dear David:

I have enclosed the current version of our proposal for
a Humanities Computing Facility. Copies are going out to
all members of the Humanities (Computing) Interest Group for
their comments.

Negotiations now should pass directly to you. Ken Fockler
said that he will prepare a brief of his own, based on his own
perceptions and on the information we have given him during
the Tuesday meetings and on paper.

The notion of a Humanistic Computing Research Group creates
some additional problems. I am prepared to head this but need
release time to do so. Erindale will have to approve my
involvement and may demand part of my salary in return.
(A separate Erindale Humanities facility is built into the
proposal, on the other hand.) Appearances to the contrary,
I would not be offended or unhappy if someone else became head.
John Hurd and Jack Stevenson are good candidates.

Berry Smith of Arts and Science spoke out strongly for an HCRG
on Tuesday morning. I think that he should be asked to
do whatever administrative shepherding has to be done
after you and IBM set the parameters.

Ann Schabas has excellent facilities and space to offer.
The best terminal/micro room, however, is controlled by the
university and is now used as teaching space. She cannot
deliver that by herself. FLIS can give us a good machine room
with a raised floor and cabling and air-conditioning.

Best wishes,

                    Sincerely,

                    Ian Lancashire

17 January 1985

To: Humanities Interest Group

From: Ian Lancashire
     Chairman
     Natural Language Processing Steering
  Committee

On Tuesday January 15 Mr Ken Fockler, Manager of Corporate and Scientific Programs at IBM Canada, met with a group of university administrators and also with a representative group of Humanities computer users in the areas of mainframe and microcomputer applications.

We discussed the possibilities in setting up a Humanities Computing Facility with the support of both IBM Canada and the university.

A proposal for this facility and for a Humanistic Computing Research Group to develop it is attached.

Please read it and pass along your comments either to Professor Russ Wooldridge, Chairman of the Humanities Interest Group, or to myself. Although the Natural Language Processing Committee and many of you have already advised us about your needs, I have written the proposal and handled discussions with both the university and IBM up to the two meetings on Tuesday.

It is possible that the proposal may not reflect the hopes and needs of all of you. I am certain that it can be improved. If we work together, we can make a good case for support.

PROPOSAL FOR A HUMANITIES COMPUTING FACILITY
AT THE UNIVERSITY OF TORONTO

This proposal outlines an instructional, editorial and research
computing facility shared by Humanities faculty and students in
about twenty departments, centres and programmes on two campuses,
and by members of a dozen SSHRCC-funded research projects.

   The university's offer of its current IBM administrative
machine, a 4341, as part of next year's planned configuration of
computing machinery for central services, is one factor. IBM's
interest in developing a three-year partnership with the
Humanities is another. The possibility that the Faculty of
Library and Information Science may have space for a Humanities
computing laboratory is a third. The permanent (and unusual)
Humanities programming staff already at UTCS, serving a diverse
group of humanists, is also an important factor.

   This proposal took shape in the Natural Language Processing
Steering Committee last fall. Since then, many faculty members
and administrators have contributed to it. As we all become
accustomed to the idea, this proposal will be refined and
modified.

The Proposed Facility

   An IBM 4341, accessed by terminals and microcomputers at the
St. George Humanities computing site, and from departmental work
areas, offices and homes, has already been set aside next year,
by the university, for some Humanities' needs and would serve as
the hub of a multi-lingual text-processing, typesetting and
logic-analysis complex, linking (a) undergraduate and graduate
courses teaching symbolic logic, effective writing, statistical
uses of data bases, and linguistic, stylistic and content
analysis, (b) university members undertaking computer-aided
studies in most Humanities subjects, and (c) international
research and editorial projects in Classics, English, French,
Linguistics, Near Eastern Studies, Philosophy and other
departments and centres.

   To be effective, however, the 4341 needs additional hardware
and software support from both IBM and the university, as well as
a site. Two sites are possible: a central facility on the St.
George Campus; and a smaller Erindale one. The Faculty of
Library and Information Science may have space suitable for a
central machine room, convenient to areas in the building that
might serve as a public terminal site: it is unclear how soon
this space can be made available. (In the long term, the
facility would best be harboured by a Humanities Research

- 1 -

Building under consideration by members of Victoria and St
Michael's Colleges.) The North Building Language Lab would
probably be suitable for an Erindale Campus site.

For departments such as Philosophy, French and English, a
central site makes sense because they have no local space for a
distributed facility and find many of their graduate and
undergraduate students already doing work in the Fisher and
Robarts Libraries.

Some disciplines, however, may prefer their own sites,
convenient to faculty offices, classrooms and library.
Departments and centres should be allowed to choose whether to
rely on a central site or to have their own site. Not every
SSHRCC project, also, will need to use a central Humanities
facility.

Humanities' activities, also, should be interpreted broadly as
including much current work in Faculties such as Library and
Information Science, and Music, in special centres and programmes
of the School of Graduate Studies (such as Medieval Studies), in
computational linguistics and natural language study in Computer
Science, and in allied institutions such as the Pontifical
Institute of Medieval Studies.

To administer these facilities, the university should consider
setting up a Humanistic Computing Research Group (HCRG), with an
academic director, an office and a modest operating budget. An
HCRG might operate under the Vice-President, Research and
Government Relations, the Faculty of Arts and Science, or the
School of Graduate Studies. The Group could draw its membership
from the Natural Language Processing Steering Committee and the
Humanities (Computing) Interest Group.

Current UTCS support for the Humanities should be subsumed
under a HCRG: a dedicated programmer (Lidio Presutti), his
part-time manager (John Bradley), operating systems support, and
secretarial support for documentation. UTCS should be asked to
manage the site and provide communications support.

The central site needs a high-speed printer, ample disk and
tape storage facilities, communications peripherals, and access
to a high-quality printer, a suitable phototypesetter and an
optical scanner.

Both graduate and undergraduate terminal areas should be
available.

Humanists need several kinds of operating systems. UNIX
provides many utilities for text analysis of small texts but does
not handle multi-lingual concording of a large corpus.
Currently, native IBM systems (MVS, VM/CMS) seem to do
"word-crunching" tasks better with very large texts, although

- 2 -

that may not be the case in the future.

Likewise, both mainframe and microcomputer facilities are needed: the former for power and capacity; the latter for many basic instructional services and faculty research.

Humanities Computing Facility Services

Here are examples of services that a good facility might provide.

1. Editorial and typesetting utilities for historical and literary texts, bibliographies, concordances, journals, dictionaries and monographs.

2. A general-purpose database management system for bibliographical and text applications, preferably one that works on both a mainframe and micros.

3. Access to international networks to facilitate exchange of research information and to use library databases.

4. Availability of crucial data-sets on disk: e.g., dictionaries for English and French (and other languages).

5. A tape library of machine-readable texts at Toronto developed with the aid of an optical scanner, allowing local researchers to contribute to an international library of well-edited machine-readable texts. (Toronto should probably specialize in a given interdisciplinary field.)

6. A "local habitation and a name" where public lectures, symposia and demonstration of academic software could take place.

7. Resources for instruction in (at present) about six courses on the St. George campus, and three courses on the Erindale Campus. For instance,

    1. CAI (computer-assisted instruction) programs to help in the teaching of languages such as French, Spanish and Greek;

    2. "pre-writing" programs that help students learn how to structure a paper, and "post-writing" problems that analyze English syntax and style for English composition courses (and for research in natural-language processing);

3. text-analysis programs and software tools (such as Toronto's COGS-3 and microText Analysis System, the Oxford Concordance Program, ARRAS, and UNIX-like utilities), and a variety of programming languages (including SPITBOL and ICON), for advanced instruction in the natural languages research (and for research and program development on both the 4341 and microcomputers);

4. programs for PHL 245H (Symbolic Logic) such as Bertie, and for graduate research in theorem proving and aspects of epistemology relating on AI (artificial intelligence);

5. programs for the statistical analysis of databases to provide graph-displayed quantitative data that can supplement historical argument, for use in HIS 1002H and other research; and

6. an environment where students not previously "computer-literate" can gain confidence with computer tools in their field.

Basic handling of Roman and non-Roman (e.g., Hebrew, Greek, Russian, Sanskrit) character sets for disciplines such as Near Eastern Studies, Classics, French, German and Slavic.

8. A computer environment for a wide variety of investigations into the design and evaluation of information storage and retrieval systems and of information networks (with emphasis on content and user characteristics). This should be a separate laboratory belonging to FLIS (as part of a new proposed degree program in Information Science), but clearly its activities would overlap to some extent with the HCRG facility.

9. Bibliographical support for Humanities research into Fisher Rare Book Library collections.

10. Published documentation for use with the Humanities facility.

11. A centre with expertise to offer other computing humanists in this region.

The general purposes of these facilities would be to permit computer-aided instruction, to sustain and enhance research, and to provide economical alternatives to traditional methods of academic publishing.

Toronto Humanities Computing Background

This proposal arises from a group of faculty members
representing all large (and most small) Humanities departments,
and two faculties. Although the Humanities come late to
computing, its faculty now participate significantly in the
administration of university computing services and facilities
and know well what can be accomplished for their disciplines with
computing tools. The following things might be said in support
of these members of the university.

1. They have a 20-year history in the field and have published
    the first textbooks on concordance software and on
    quantitative analysis of language.

2. They are active in software development: e.g., a
    Greek-language-instruction system, Hebrew/Greek concording
    and typesetting programs, a textbook in computer-aided
    literary analysis.

3. With faculty advice and software, UTCS Humanities
    programming support staff are actively developing
    text-analysis programs (e.g., COGS-3, microText Analysis
    System) now being requested by other Ontario universities
    (Western Ontario, Queen's).

4. They represent a heavy concentration of international
    research projects: 9 SSHRCC projects here are actively
    using and developing computer software (Dictionary of Old
    English [DOE], Records of Early English Drama [REED], Royal
    Inscriptions of Mesopotamia [RIM], the Athenians, the
    Correspondence of Madame de Graffigny, the Estienne-Nicot
    Dictionaries, the Leibniz Lexicon, the Documents of Essex
    England Data Set [DEEDS], and the Bibliography of
    Philosophy in Canada).

5. They have a good record of attracting university support:
    e.g., Presidents' Committee grants to DOE and REED (1982);
    set-up of Humanities Support Group at UTCS (1982); English
    Department effective-writing teaching machine (1983);
    various discipline micro facilities.

6. They innovate in teaching: currently six graduate and
    undergraduate courses exist in this area.

7. They are numerous: well over 100 people are listed with the
    Humanities (Computing) Interest Group.

8. They enjoy strength in allied subject areas such as mathematical and computational linguistics, library and information science.

9. They have good links to important Humanities computing centers at Waterloo (Toronto-Waterloo Cooperative on Information Technology), Oxford, Tubingen, Chapel Hill, and Berkeley.

IBM Support

Discussions have been ongoing for some months about a possible three-year partnership between the university's humanists and IBM. Our knowledge of IBM hardware and, especially, software is not what it should be, but the following describes one configuration that might serve many of our needs.

1. IBM 4341 (U of T administrative machine) or 4361 (model 3) running VM/CMS and possibly UNIX if it becomes available

2. large disk storage: 2100 Mb needed for large text concordances (typically requiring 1000 Mb and more), on-line national dictionaries, and data bases such as the Thesaurus Linguae Graecae

3. several fast tape drives

4. a good high-speed printer such as the IBM 6670 (or the SHERPA printer)

5. Series I (typesetter interface for TE/E) and 3277 graphics workstation

6. IBM maintenance of the 4341 and related IBM hardware and software

7. about the mainframe site: 10 terminals (graduate research), 20 standard PCs or PCjrs (undergraduate instructional), and 5 advanced PC/ATs

8. 5 non-standard PCs made for Israel, Saudi Arabia, Japan, etc., and 5 advanced PC/ATs (for language departments, centres and research projects on the St. George campus)

9. 5 advanced PC/ATs and 5 PCs or PCjrs at an Erindale Humanities Language Lab, communicating with the St. George laboratory

10. communications peripherals to UTCS PACX; a
    local-area-network for microcomputers

11. IBM software such as language development tools (PASCAL,
    FORTRAN, PL/1), a database management system such as SQL,
    the Document Composition Facility (DCF), and the Integrated
    Publishing System (IPS) typesetting programs

12. co-operation and support of IBM personnel, as at the
    National Language Technical Center

13. 4250 high-quality printer for academic publications
    (journals, monographs, festschriften, etc.) and
    documentation

14. pre-commercial-release access to EPISTLE, an
    error-recognition program for business documents, for "beta
    testing"

University "Deliverables"

Under a cooperative agreement, in return for its support, IBM
will expect to see the Humanities deliver agreed-on academic
services or resources to a community of computer users. These
could include such things as a student computer laboratory,
software for teaching or research, published reports and
documentation, and participation in public activities such as
lecture series and conferences. Where such facilities,
publications and activities coincide with what the Humanities
would in any case undertake if it had the necessary resources, an
agreement of this sort seems consistent with academic goals at
this university.

"Deliverables" on the part of the university might include
various services described above under "Humanities Computing
Services."

Impact on Current Humanities Computing

IBM support could not provide all services that the Humanities
needs for its current computing.

Management of the facility will fall to UTCS for the time
being.

Licensing instructional software not in the public domain for microcomputers in the number required for a laboratory will require some funds. (One license appears to be necessary for each user machine.)

CMS under VM does not have the powerful Sort package needed by COGS-3. A non-IBM program would have to be used. Some sorting operations -- ones on a million words are not uncommon -- might require a larger mainframe.
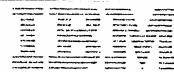
Typesetters compatible with DCF and IPS appear to be made by Autologic (APS-5 or Micro-5). Several Humanities projects already own other makes (REED, a Linotron 100; RIM, a Compugraphic 8400) and it would be desirable to build on existing strength.

Some humanists use LISP and other programs running under UNIX (such as Writer's Workbench in the English Department) that do not appear to work on IBM machines. One humanist uses a version of APL that is not IBM's.

Departments such as Linguistics operate Apple microcomputers to take advantage of Gutenberg, which handles word-processing (but not data processing) needs for non-Roman alphabets. No major manufacturer has faced up to the problem of allowing users to define their own character sets and collating sequences. (IBM, however, does support all important character sets in the world now and presumably has access to machines sold in Europe, the Middle and Far East, etc.)

Like other parts of the university, then, the Humanities will continue to compute in a multi-vendor environment and will need the wherewithal to develop additional standalone resources of its own and to share in general-purpose UTCS systems.

        Ian Lancashire
        Chairman
        Natural Language Processing Steering Committee
        28 January 1985

January 29, 1985

Professor Ian Lancashire
Professor of English
University of Toronto
C/O REED Office
85 Charles St. W.
Toronto, Ontario
M5S 1K5

Dear Ian

Ian, thank you very much for setting up the meetings on January 15 at Victoria College. They were very informative and helpful.

I will be reviewing my notes and the material that you have provided me over the next few weeks. I would then like to meet with you to discuss my interpretation of the statement of the objective, activities and deliverables of the proposed project.

The Humanities area of the University of Toronto is very impressive and diverse in its interests and activities. I was impressed with the number of projects already underway and the eagerness to get on with other creative ways of using computing.

It will be exciting for me to become a "humanist" over the next few months as we work on this proposal. I met with David Nowlan last Friday and he offered much encouragement and a few good pointers on how to sell the project. David wants to be kept informed of our progress and I would suggest that he join us for some of our future meetings.

I will call soon to set up our next meeting.

Yours truly

K.H. Fockler:kh
Scientific and Education Programs Manager
Corporate and Scientific Programs

cc: Professor D.M. Nowlan - University of Toronto - Simcoe Hall

Department of English
7 King's College Circle
7 May 1985

Professor David M. Nowlan
Vice-President, Research and Government Relations
Simcoe Hall

Dear David:

   I want to bring you up-to-date on my dealings with IBM and the
proposed Humanities partnership. We last touched base shortly
after the January 15 meetings between Ken Fockler and a group of
administrators and humanists at Victoria College. You received
my "Proposal for a Humanities Computing Facility at the
University of Toronto" (28 January), which emerged from my survey
of Humanities faculty. I tried to incorporate everything of
possible interest there to focus discussion on planning.

   Since then, I have done three things: engage faculty in the
business of planning; continue discussions with IBM on their
deliverables; and take the issues of Humanities computing to
universities outside Toronto. We can now refine my January
proposal.

   Engaging the Faculty. At my request Russ Wooldridge of French
convened three meetings of the Humanities Interest Group: one on
research typesetting (the REED solution of IBM PC/XT and Linotron
100), another on COGS-3 (the major UTCS text-analysis program);
and a third on character sets, with papers by John Hurd and a
representative of IBM's National Languages Technical Center. I
then prepared a two-hour seminar on computing in Humanities
instruction for the Arts and Sciences Teaching Workshops on 21
March, with cooperation from Professor Glyn Holmes (French,
Western Ontario), who with Professor Paul Bratley (Computer
Science, Montreal) has edited all four issues of Computers and
the Humanities for 1985. On April 15 I also talked to a Slavic
Workshop about Humanities computing opportunities.

   I also joined Sandy Johnston's planning committee for the
Humanities Research Building proposed for land belonging to
Victoria and St. Michael's Colleges. Unfortunately, real
political opposition to this project has cropped up, as you must
know, but I will continue to work with Sandy on it. There is no
prospect of linking this proposal with the discussed
IBM/Humanities partnership, at least in the next few years.

   We learned several things from all these discussions. (1)
While IBM cannot yet give a general solution for all character
sets, it can help with many, especially our two national Canadian
languages, English and French, as well as Latin, Greek, Hispanic,

German, Hebrew, Arabic and Japanese. A closer association with IBM's National Languages Technical Center would be worthwhile. Negotiations between the Dictionary of Old English and Xerox however, may be a preferable way of meeting the specialized needs of Asian languages.

(2) REED's method of typesetting has favour with colleagues in both French and Classics: we should encourage the growth of an editorial and typesetting centre along these lines. A central typesetting facility, beyond what UTCS is already doing, might confuse things.

(3) COGS-3, developed by John Bradley and Lidio Presutti, is now a first-class product and should be developed further. Confirmation came in just last week from a conference in New York, attended by a member of the French department, that we are doing advanced work with COGS.

(4) French, Philosophy and English have very active interests in an undergraduate computer laboratory with IBM PCs: Philosophy has begun developing a program named SymLog for its large Symbolic Logic courses; English and French have cooperated to introduce MTAS (Micro Text-analysis System) into literature teaching; French is interested in supplementing its language courses with Canadian CAI programs; and English hopes to expand the computing component of its effective-writing program beyond UNIX, a difficult environment for non-scientists, into DOS. The large attendance at the Arts and Science seminar, and the enthusiasm of departments as varied as Philosophy and Near Eastern Studies, contrasted in my mind with the low level of enthusiasm at Erindale College last year (where the Humanities discipline representatives rejected a proposal for a computer component in teaching that came from a 6-person task force chaired by myself).

You will remember my own interest in an English parsing system. A new member of DCS, Graeme Hirst, taught a course in Computational Linguistics that I attended last term, and we recently discussed ways and means of creating a parser within a proposed IBM partnership. We have agreed to ask IBM for Penelope, the parsing component of EPISTLE/CRITIQUE, as a basis for developing a semi-automatic parser for historical English (from Old English onwards). DOE is interested in this work, and Graeme knows the Yorktown Heights people and they have already agreed to participate in a conference on Humanities software that I am planning here next February.

Continuing Discussions with IBM. On April 1 I met with Ken Fockler and Mike Brother's assistant, Bob Smithson, to review the January proposal and the Humanities' role in the current partnership. I agreed to include the texts of four of my talks at conferences in the current deliverables and to add MTAS once Lidio Presutti has rewritten parts of it on the PC/AT this

- 2 -

summer. (None of MTAS has so far been done with IBM assistance.) Ken Fockler seemed to have no problem with my proposal, except for EPISTLE, about which IBM is having some internal discussion (this was before my discussions with Graeme Hirst). I added one possible university deliverable: the proposed conference on Humanities software here in February (to this Ken expressed warm interest). Bob Smithson listened coolly but he has since given me, at my request, documentation on IBM's VM/IX and PC/IX operating systems.

Don Avery, another of Mike Brother's new assistants, asked me to speak at WAT85 last Friday, where I talked about micro networking, Humanities departments, MTAS and SymLog. Don will be producing the IBM newsletter on its partnership agreements. At WAT85 I distributed free diskettes containing MTAS (not source code, however) and descriptions of it and SymLog. IBM enclosed both in a special binder with a University of Toronto cover.

This coming Thursday (May 9) I will be visiting Ken Fockler and Bob Smithson at IBM to talk about Humanities deliverables.

External Affairs. This term I have lectured at Queens (January 28), Waterloo twice (March 28, for the New Oxford English Dictionary project; and May 3, at WAT85), and York (April 15). Although the talks were different, I did stress the many applications of Humanities computing and the need for inter-university co-operation in software development. A long talk with John Leyerle (my former thesis supervisor and an old friend) helped me with a strategy towards these ends. Russ Wooldridge has agreed to act as co-chair of a Humanities software conference here next February, and we intend to ask the Waterloo-Toronto Co-operative to sponsor us. I now have the active support of Glyn Holmes (Western Ontario), Paul Bratley (Montreal), Grace Logan (Waterloo) and Graeme Hirst. Paul will be paying air fares for Glyn and myself to Montreal at the end of May for a planning session. I think that we have an excellent opportunity here to make Toronto a center of a Canada-wide Humanities software enterprise.

Now, for some specifics of an IBM/Toronto Humanities partnership. These are only suggestions, David, not "demands"; they represent the latest in my thinking, written down quickly, and without the nuts-and-bolts expertise that others have. So far, people have not been breaking down my door to help me with this! (I even had some trouble getting information on VM/CMS.)

IBM Deliverables.

- Software and maintenance for the 4341, Series I and DASD: both VM/CMS and VM/IX, with preference perhaps to the latter, since it is new (John Bossins and I have chatted about a mutual interest in UNIX under VM)

- New disk drives and tape drives for the 4341

- High-quality printer (IBM 6670) for the 4341

- 50 PC/XTs (number and type to depend on sites and teaching plans)

- Micros for graduate use: 5 PC/ATs

- Printers to serve this micro laboratory

- Additional DOS and VM software (e.g., Penelope)

- Co-operation and support of IBM programmers and personnel, as at the National Language Technical Centre

University Deliverables

- 4341 and Series I, at UTCS, serving graduate-student and faculty research, administered under the Natural Language Processing Committee (and for that purpose, given a modest annual operating budget outside UTCS)

- Terminals (perhaps some of the gift 3101s) to the 4341 in departments and research projects

- UTCS operator for the 4341

- A major student microcomputer facility using the WATSTAR network, for 30/50 PC/XTs, preferably at FLIS: about $50,000.00 for a 30-micro network, and about $75,000.00 for a 50-micro network (developed by the University of Waterloo)

- UTCS Micro Lab part-time operator for WATSTAR, under the Natural Language Processing Committee (aided by that independent operating budget)

- Small student micro language lab at Erindale (possibly without network), probably operated by the French department

- Annual conference on Humanities software: SSHRCC and university support required, to be organized by the Natural Language Processing Committee

- Student user's guide to the Toronto WATSTAR laboratory, by Humanities faculty under the Natural Language Processing Committee

- Implementation report on the Toronto WATSTAR facility, by myself and UTCS

- Software development on micros: this might include MTAS (public domain) and SymLog (I do not know whether Philosophy

would develop this commercially; if so, SymLog falls under
"Technology Transfer"); a part of Lidio Presutti's time
could be targeted to this

- Software development on the 4341 under VM: this might
  include COGS-3, other text/logic-analysis utilities, and an
  English parser; again, a part of Lidio Presutti's time
  belongs here, but SSHRCC support would be needed to develop
  a parser

A few final points.

Toronto must develop a better ongoing relation with Waterloo
than it now has if we want to get serious about working closely
with Canadian universities on Humanities software. WATSTAR seems
to make sense not only as a network for undergraduate Humanities
students but also as an "outward and visible sign" that the two
universities can operate in parallel. I am visiting Waterloo
tomorrow and will report back to you on WATSTAR.

Last, without wishing to push myself forward, David, if the
university needs me in all this business, then I have a problem.

My present status as an Erindale College faculty member on
temporary loan (as Professional Faculties Coordinator) to the
Department of English at St. George leaves me "up in the air" and
rather vulnerable to arbitrary pressures from Erindale. Several
important colleagues there have made clear to me that my services
are needed there and that my heavy involvement downtown gives
them problems. (I know that English staffing at Erindale is
difficult from my three years as Discipline Representative there,
and this year the College lost one member to the Department
chairmanship, while another, whose services the College had not
secured when it had the chance, decided to go back downtown.)
Twice recently, pressure has been put on me to take over the
editing of an Erindale journal in Renaissance studies. (I have
now refused twice.) As well, I have been urged twice recently to
negotiate with DEC on behalf of Erindale. While I could not
refuse, I am reluctant to do so now, for obvious reasons.
Embattled English colleagues at Erindale last fall asked me to
teach an undergraduate course in computer applications in
English, despite the College's rejection of my Humanities
computing proposal last year. (When I agreed, my course
description barely survived College hostility on a shaky road to
approval.) My future in Humanities computing teaching at
Erindale does not seem to me very rosy.

My assessment is that Erindale's need for English staff is
greater than its need for someone in my research and teaching
area, but that out of desperation the College will fight hard to
keep any live body in its budget. This aspect of the IBM
partnership process you will have no control over, but it is as
well that you should know of the difficulties of my position, and

- 5 -

of my relative helplessness to do much about them. I would
appreciate advice, if you have any.

    Apologies for the length of this, and thanks for your warm
support for my efforts and the Humanities.

                        Sincerely,

                        Ian Lancashire

        P.S. Copies of my CASE '84 talk and of
        three other talks are now with Eva Swenson
        for your use.

20 May 1985

Professor R.C. Brown
Dean
School of Graduate Studies
University of Toronto

Dear Craig:

You may remember that we briefly talked about the possibility of
a graduate program in Humanities computing several months ago.
Developments have advanced and I can now write to you with a
tentative proposal.

The University has designated an IBM 4341 mainframe, a set of
disk drives, and a Series 1 unit as a "seed" for a future
partnership between the Humanities and IBM Canada Ltd.   My recent
talks with IBM indicate an agreement in late 1985 and the
installation of equipment by early 1986. We have in mind a
Humanities Computing Resource Centre with a student laboratory of
networked microcomputers, linked (where possible) with the 4341.
The partnership would call on humanists to run symposia, give
presentations and develop documentation for a graduate and
undergraduate computer laboratory.   Plans are underway for the
first symposium in Reading Week.   Other humanists are now moving
towards cooperation with different manufacturers (e.g., Xerox).

One difficulty experienced by Humanities faculty in this area
is that they do not have an administrative canopy under which to
pursue their work.   We have managed so far by extending the
responsibilities of the Natural Language Processing Steering
Committee, but this really only exists to recommend activities
for Lidio Presutti, the UTCS Humanities Support Programmer.   The
NLPSC is not an academic body.   Presumably the University will
want to create a partnership committee to oversee the Humanities
arrangements, but this would exclude humanists computing on
non-IBM equipment.

I have explored a number of options.   One would be to set up a
Humanistic Computing Research Group, along the lines of CSRI.
Unfortunately, Humanities faculty have no common ties to a
suitable host department (we have no equivalent of Computer
Science) and, collectively, we do not have enough external
operating funds to sustain a HCRG.   A second option would have
Humanities computing administered under an Arts and Science
committee: this would suit some of the undergraduate
instructional computing we hope to undertake (and I hope such an
arrangement will be possible) but has two disadvantages at
present.   Current strength in Humanities computing is in graduate
research, not undergraduate instruction.   We have half a dozen
graduate courses on the books in English, French, Linguistics,
Philosophy and Religious Studies.   The second problem is that

- 1 -

several members of our group belong outside Arts and Science
(e.g., the suburban campuses, the Faculty of Library and
Information Science, and Medieval Studies).

The third option known to me is a graduate program in
Humanities Computing.   After some struggling with the
complexities of the St. George campus (I have taught at Erindale
for many years), I sense that the solution best suited to our
diverse group would be an association with the School of Graduate
Studies, along lines sketched in the tentative calendar
description enclosed.   You will see that it is modelled on the
McLuhan Program in Culture and Technology, a Social Sciences
program to which we have some affinity.

I would appreciate knowing your reaction to this.   If you think
it has enough merit for public debate, I would prepared to take
the proposal further.

<div align="right">
Yours truly,
</div>

<div align="right">
Ian Lancashire
Professor of English
</div>

cc. D. M. Nowlan
    A. H. Schabas
    J. J. B. Smith
    R. A. Taylor
    members of the N.L.P.S.C.

Department of English
7 King's College Circle
8 June 1985

Professor David M. Nowlan
Vice-President, Research and Government Relations
Simcoe Hall

Dear David:

   IBM-Humanities negotiations are going nicely with Al Heyworth's
help on communications issues, which turn out to be crucial to
the success of the proposal.

   George Lewis, Bob Smithson, Al Heyworth, Russ Wooldridge and I
met in late May to sketch out IBM needs in carrying a proposal
ahead internally. Then, after consulting John Bradley about the
disk storage requirements, everyone except Russ Wooldridge met
again on June 7 for three hours to review four diagrams created
by Al to represent possible scenarios. We reached agreement on
one. Al's diagram -- he is revising it for us -- shows best what
we will go forward with: one main graduate and undergraduate site
(at FLIS, if possible); one small similar site at Victoria; a
third, graduate only, at the Pontifical Institute of Mediaeval
Studies; and a fourth, undergraduate only, at the Erindale
campus.

   You know my personal views on how IBM systems do not always
conform to Humanities' needs: e.g., UNIX as a good small-text
operating environment (IX/370), WATSTAR rather than PC Network as
a LAN, and REED's Linotron rather than the 4250 as a suitable
typesetting machine. In the past few weeks it has become clear
that if we are to negotiate successfully with IBM, I must give
way temporarily on a couple of things. Any IX/370 and PC/IX
system will be harder for IBM to sell internally than a CMS and
DOS system; and we will not be able to communicate effectively
from micro to mainframe if we run IX. Al also thinks IBM is very
keen on having a showcase for its 4250 printer as well as a good
site where PC Network links PCs with the 4341.

   I have tried to make the proposal attractive to IBM by making
VM/CMS the main 4341 operating system, and DOS the dominant micro
one. Small, essentially micro-based facilities will run IX part
of the time. We can import UNIX-type utilities under VM/CMS to
gain the advantages of both. Although "giving in" here, I have
enriched our gift request list by asking for micros in support of
individual researchers and by adding some nice peripherals, such
as the 3820 page printer and a 37XX communications device that Al
thinks important. My argument to IBM has been (a) that it can
best persuade faculty of the value of a 4341 mainframe if it puts
the power of one within reach of certain "key" faculty (the best
way to do so being to give them a "gateway" in the form of PC/ATs

and PC/XTs), and (b) that it can reap a wider harvest of "presenters" at IBM conferences, and of papers or reports that might swell the deliverables, than it now has by targeting gift machines at individual researchers.

With some misgivings, I have accepted PC Network over WATSTAR for the time being. WATSTAR's price-tag for the university in comparison to that of PC Network, and the apparent direction IBM evidently intends to take with its own LAN, suggest that we ought to give the latter a chance. The 1986 Humanities software conference will give me an opportunity to make a case for WATSTAR in a public forum. Then the Humanities can apply to funding if it wishes. Al did not think WATSTAR extraordinary.

Several important issues still remain to be settled before the partnership goes ahead.

1. the location of the main resource center: IBM wants a named
    site to bring forward in the process of getting internal
    agreement

2. the number of humanist research projects to receive IBM
    micros: this resolves into how rich IBM can make their
    donation

3. funding to hire UTCS to manage the facilities, which in
    turn depends on

4. the administrative structure for the facility, and its
    budget;

5. whether or not we ask IBM for a database system, SQL (we
    have too little information now about its limitations: Al
    has agreed to consult a humanist working at Yorktown
    Heights with SQL about these);

6. the availability of Penelope, an IBM parsing program that
    is part of Epistle;

7. the availability of the Sherpa scanner we saw at Palo
    Alto.

IBM will be making its first presentation about this proposal in August, I think. A signed agreement late this year seems to be its intent. We would be wise to nail down all these issues before late July.

Plans for the software conference, now scheduled in mid-April, also go well. We have speakers already lined up, and Waterloo's strong interest in becoming involved.

I wonder if UTLAS might also consider participating. For the time being I have put library inquiry systems down as part of the

conference agenda. The presentation from DataPhase struck me as the best so far. If Humanities computing can help in the UTLAS-Toronto software development enterprise, or in any other respect, by all means ask.

I'll get in touch with you after returning from Washington on June 29th.

Sincerely,

Ian Lancashire

8 June 1985