

**Report on Humanist  
Digital Humanities, Oulu, 2008**

**W. McCarty, Editor**

**Summary**

In its 21<sup>st</sup> year of operation, since the last report, Humanist gained a significant number of new members, growing by 113, the largest gain by far since I began tracking the yearly figures. The number of messages is up by 70, continuing what now looks to be a fairly regular sinusoidal pattern. Available data is summarized in the accompanying spreadsheet. The extensive repair of Humanist's ageing mechanisms of publication, and its move from Princeton and Virginia to an ADHO server, was not accomplished as promised but is definitely now underway and should be accomplished by this Autumn.

**Management**

Humanist is tended by one person, the Editor, with occasional help from the Assistant Editor (David Gants). In previous years the average amount of time spent daily on running Humanist was ca. 30 minutes per day, 365 days per year. During 2007-8 the workload has gone up for three reasons:

- (1) A dramatic increase in spam through the Virginia account caused an unsupportable amount of manual work, as a result of which I suspended trawling through rejected messages for the few genuine items falsely classified as spam. Consequently members have had to watch for the appearance of their postings, resubmitting as necessary.
- (2) The increasing habit of pasting formatted text from Word into postings has required an increasing amount of manual editing.
- (3) A breakdown in the archiving mechanism at Virginia now means that I must archive messages manually.

The refurbishment of Humanist (for which see below) and its move to an ADHO server promise to reduce the workload to a reasonable level once again.

Until Humanist's move, technical support is supplied by Shayne Brandon (IATH, Virginia), who once again is to be commended for his help at crucial points throughout the year. No help this year has been required from the Office of Information Technology (OIT, Princeton). IATH maintains mail relaying, message digesting and archiving functions, OIT the Listserv software.

---

## Humanist's refurbishment

The refurbishment of Humanist now has a history.

At DH2006 both the ACH and ALLC gave strong support to an initiative for repairing Humanist's ageing mechanisms of publication and for moving these mechanisms from their current institutional homes to an ADHO server. (The problems then visible were summarized in the report for 2005-6.) During 2006-7 a survey of members was carried by Stephen Ramsay, Stéfan Sinclair and Vika Zafrin to discover if Humanist's idiosyncratic style of publication needed to be changed along with the lower-level mechanisms. Members overwhelmingly voted for continuing the current style. With the financial support of ADHO, a programmer was then engaged to carry out the work but did next to nothing, despite repeated enquiries. He resigned in early February 2008. Later that month, after consultation with ADHO, Malgosia Askanas, known widely for her efficiency and skill, was hired. She drew up a detailed plan after consultation with ADHO, began chargeable work 1 May and has been showing steady progress since then. Completion by Autumn seems highly likely.

Both Malgosia's proposal and milestone documents are attached, below. Her estimate for the work is 100 hours at a reduced fee of US \$60/hour (€40, £31) for a total of US \$6000 (€4000, £3100).

## Membership and messages

Statistics regarding membership and messages are given in the accompanying Excel file, `Humanist_200807.xls`, as follows:

- (1) Membership data for 2002-3 to 2007-8
- (2) Number of messages per year from 1993-4 to 2007-8
- (3) Chart of membership from 2002-3 to 2007-8 from (1)
- (4) Chart of messages per year from (2)

It is, as usual, difficult to know what these statistics indicate other than the steady, perhaps even accelerating growth in membership in the U.K., U.S. and perhaps Canada (which continues to show the highest rate of membership per capita. The chart of messages per year continues to show a puzzling sinusoidal pattern, as noted above. I can shed no further light on this variation.

## Publication

Humanist is a publication of the Alliance for Digital Humanities Organizations and the Office for Humanities Communication (King's College London). Since 1995-6, when Stanley N. Katz was its President, Humanist has

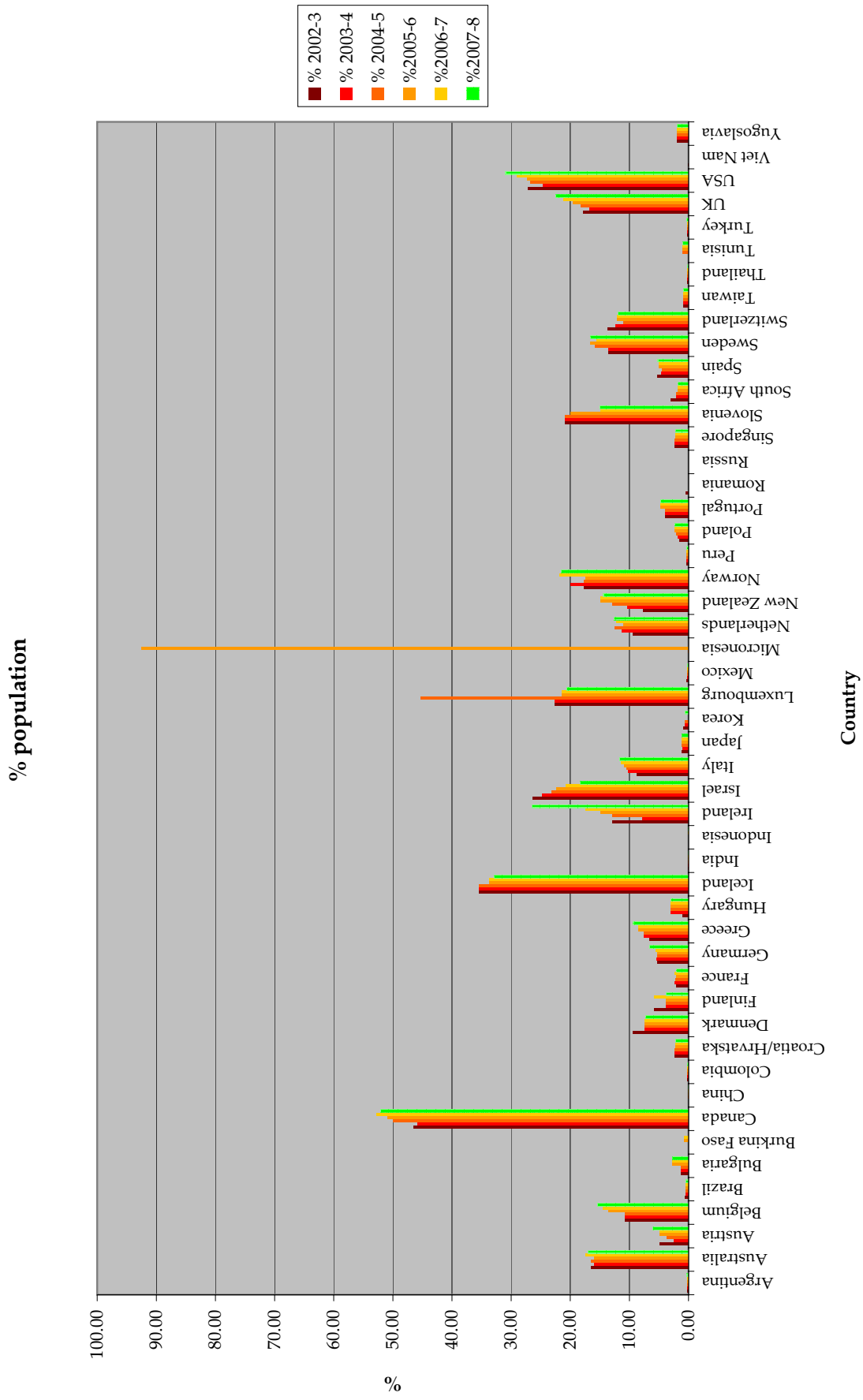
also been an affiliated publication of the American Council of Learned Societies, relations with which continue to be cordial but otherwise inconsequential.

### **Coda**

Ending on a note of gratitude continues to be appropriate – to the wider, highly diverse communities of people who sustain Humanist, and so our voice in the world; more specifically, to the organizations of ADHO for their endorsement and support, which have again greatly strengthened this year.

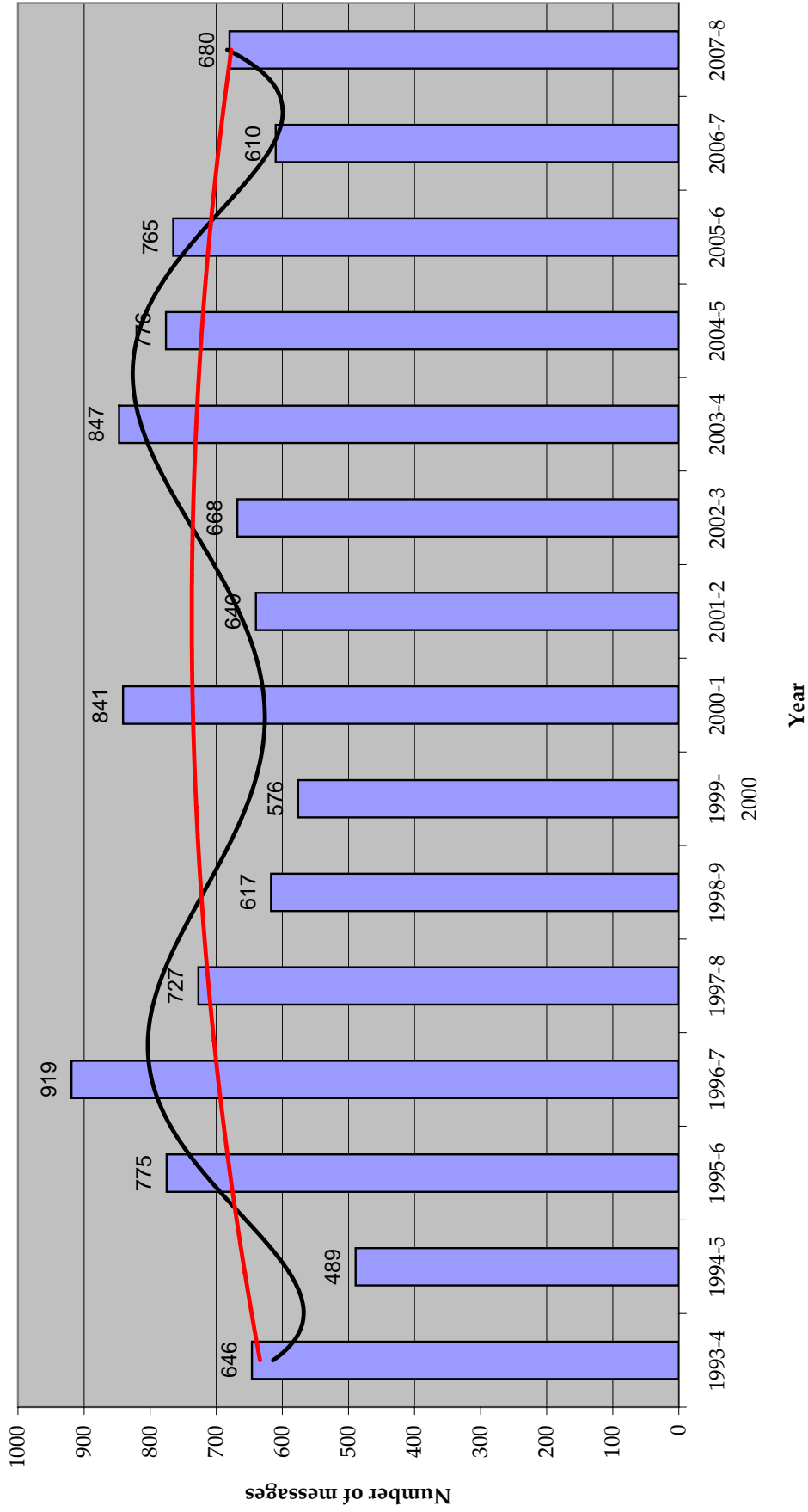
Willard McCarty  
June 2008

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
COUNTRY	POPULATION 2002-5	POPULATION 2005-	POPULATION 2008-	POPULATION 2002-2003	% 2002-3	2003-2004	% 2003-4	2004-5	% 2004-5	2005-6	%2005-6	2006-7	%2006-7	2007-8	%2007-8	Difference from 2006-7
1	Argentina	38537843	40,677,348	1	0.26	1	0.26	1	0.26	1	0.26	1	0.25	1	0.25	0
2	Australia	19489797	20,600,856	32	16.41	31	15.90	32	16.41	32	15.93	35	17.42	35	16.98	0
3	Austria	8,184,691	8,205,533	4	4.89	2	2.44	3	3.67	4	4.89	4	4.89	5	6.09	1
4	Belgium	10,277,731	10,403,951	11	10.71	11	10.71	11	10.71	14	13.51	15	14.47	16	15.38	1
5	Brazil	162,177,623	191,908,958	1	0.60	9	0.49	6	0.44	8	0.43	8	0.43	9	0.47	1
6	Burkina Faso	7,636,666	8,254,745	0	-1.31	0	-1.31	1	1.31	1	1.31	1	0.72	1	0.09	-1
7	Burkina Faso	31,641,184	33,212,698	147	46.46	145	45.83	158	48.33	167	50.91	173	52.74	173	52.08	0
8	Canada	29,812,701	30,933,381	1	0.01	1	0.01	1	0.01	2	0.02	2	0.02	2	0.02	0
9	China	1,281,271,011	1,330,044,605	1	0.24	1	0.24	1	0.24	1	0.23	1	0.23	1	0.22	0
10	Colombia	4,164,435	4,356,427	1	2.33	1	2.33	1	2.33	1	2.22	1	2.22	1	2.23	0
11	Croatia	4,498,504	4,481,543	5	9.34	4	7.47	4	7.47	4	7.36	4	7.36	4	7.29	0
12	Denmark	5,432,335	5,484,723	3	5.79	2	3.83	3	3.83	2	3.83	2	3.83	2	3.81	-1
13	Egypt	52,244,422	52,544,749	12	6.25	12	6.25	12	6.25	12	6.25	12	6.25	12	6.25	0
14	Finland	5,191,989	5,244,749	5	5.25	4	5.37	4	5.25	4	5.25	4	5.25	4	5.25	0
15	France	61,949,221	62,431,939	12	6.25	12	6.25	12	6.25	12	6.25	12	6.25	12	6.25	0
16	Germany	82,431,939	82,369,546	43	5.25	44	5.37	43	5.25	44	5.34	46	5.58	46	6.56	8
17	Greece	10,620,000	10,668,354	7	6.60	8	7.55	8	7.55	9	8.44	9	8.44	10	9.33	1
18	Hungary	10,021,272	10,068,354	1	1.00	3	2.99	3	2.99	3	3.00	3	3.00	3	3.02	0
19	Iceland	292,783	296,737	1	35.36	1	35.36	1	35.36	1	33.70	1	33.70	1	32.86	0
20	India	1,049,792,881	1,147,985,898	5	0.05	5	0.05	6	0.06	6	0.06	6	0.06	6	0.05	0
21	Indonesia	23,162,700	23,752,355	5	12.91	3	7.74	5	12.91	6	14.94	7	17.43	11	26.47	4
22	Ireland	4,015,676	4,195,119	6	2.89	15	4.24	4	2.89	15	4.24	15	4.24	15	4.24	0
23	Israel	6,112,338	6,112,338	6	6.11	6	6.11	6	6.11	6	6.11	6	6.11	6	6.11	0
24	Italy	57,551,551	57,551,551	50	12.91	49	12.91	49	12.91	49	12.91	49	12.91	49	12.91	0
25	Japan	127,455,682	127,288,413	15	1.18	13	1.02	14	1.10	14	1.10	14	1.10	14	1.10	0
26	Korea	48,221,340	49,232,844	4	0.83	3	0.62	3	0.62	3	0.62	3	0.62	3	0.61	0
27	Luxembourg	4,411,971	4,486,006	1	22.67	1	22.67	2	45.33	1	21.34	1	21.34	1	20.58	0
28	Mexico	10,558,341	10,820,200	3	0.28	2	0.19	2	0.19	2	0.19	2	0.19	2	0.18	0
29	Micronesia	108,150,000	108,150,000	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
30	Netherlands	16,407,491	16,645,313	15	9.36	18	11.23	20	12.48	18	10.97	20	12.19	21	12.52	1
31	New Zealand	3,885,533	4,035,461	3	7.72	4	10.29	5	12.86	6	14.97	6	14.97	6	14.97	0
32	Norway	4,417,460	4,417,460	0	1.12	0	1.12	0	1.12	0	1.12	0	1.12	0	1.12	0
33	Poland	28,244,665	27,926,628	3	10.35	1	0.35	1	0.35	1	0.36	1	0.36	1	0.36	0
34	Portugal	10,077,168	10,676,910	6	1.55	7	1.81	8	2.07	9	2.33	9	2.33	9	2.34	0
35	Romania	22,329,977	22,246,862	1	0.45	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
36	Russia	140,702,094	140,702,094	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
37	Singapore	4,250,720	4,250,720	0	2.35	1	2.35	1	2.35	1	2.26	1	2.26	1	2.17	0
38	Slovenia	2,017,071	2,017,071	1	2.35	1	2.35	1	2.35	1	2.26	1	2.26	1	2.17	0
39	South Africa	43,220,333	43,220,333	13	2.35	13	2.35	13	2.35	13	2.26	13	2.26	13	2.17	0
40	Spain	43,086,333	43,086,333	22	5.34	20	4.62	19	4.41	20	4.36	20	4.36	21	4.36	0
41	Sweden	8,847,243	9,045,389	12	13.56	12	13.56	14	15.92	15	16.66	14	15.55	15	16.58	1
42	Switzerland	7,489,370	7,581,520	10	13.73	9	12.36	8	10.88	9	12.02	9	12.02	9	11.87	0
43	Taiwan	22,901,088	22,920,946	2	0.88	2	0.88	2	0.88	2	0.87	2	0.87	2	0.87	0
44	Taiwan	65,444,371	65,444,371	1	0.16	1	0.16	1	0.16	1	0.16	1	0.16	1	0.15	0
45	Thailand	65,444,371	65,444,371	1	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
46	Tunisia	10,074,951	10,383,577	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
47	Turkey	68,865,559	71,392,807	1	0.15	1	0.15	1	0.15	1	0.14	1	0.14	1	0.14	0
48	USA	284,869,541	284,869,541	718	17.16	703	16.23	708	16.23	708	16.23	708	16.23	708	16.23	0
49	USA	81,708,541	81,708,541	1	0.12	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0
50	Viet Nam	85,355,576	86,116,556	2	1.87	2	1.87	2	1.87	2	1.87	2	1.87	2	1.87	0
51	Yugoslavia	10,693,953	10,693,953	2	1.87	10	1.87	1	1.87	1	1.87	1	1.87	2	1.87	0
52	Yugoslavia	10,693,953	10,693,953	2	1.87	10	1.87	1	1.87	1	1.87	1	1.87	2	1.87	0
53	Yugoslavia	10,693,953	10,693,953	2	1.87	10	1.87	1	1.87	1	1.87	1	1.87	2	1.87	0
54	TOTAL	41,479,796	41,479,796	1,385	13.85	1,300	13.00	1,383	13.83	1,458	14.58	1,458	14.58	1,458	14.58	0
55	Change			-85		-85		83		75		79		113		



Year	Messages
1993-4	646
1994-5	489
1995-6	775
1996-7	919
1997-8	727
1998-9	617
1999-2000	576
2000-1	841
2001-2	640
2002-3	668
2003-4	847
2004-5	776
2005-6	765
2006-7	610
2007-8	680

Messages/year 1993 to 2008



## **Proposal for the new Humanist List framework**

Version 2; April 18, 2008

Malgosia Askanas

(ed. W. McCarty 15/6/08)

The present document constitutes a proposal for the framework within which the Humanist list will operate at its new location at digitalhumanities.org.

### **I. Overview**

#### *I.1 New location of the Humanist List and the software*

This proposal calls for the Humanist list itself to be moved from its current listserver at Princeton to a listserver at digitalhumanities.org, and for the list archives to be moved from their current location at Virginia to digitalhumanities.org. Similarly, all the software, web pages, databases, etc. described in the present proposal are intended to reside at digitalhumanities.org.

### **I. 2 The proposed Humanist List email flow**

Members of the list submit posts by sending them either to Willard or to the Humanist list address. In both cases, the posts will be collected at a single location on the server, and Willard will be able to perform interactive editing and grouping of these posts, as well as send the final results to the Humanist listserver, via a special web interface (referred to as the **Editorial Interface**). The Editorial Interface will offer a number of interactive functions for the control of the editing and sending process, and the underlying software will aid in the editorial process by automatically performing certain routine transformations on the messages.

### **I. 3 The Humanist List subscription process and the Membership Database**

Those wishing to subscribe to the Humanist list will be requested to fill out a form (the **Membership Application**) located on the Humanist website, which will be moved from its present location at Princeton to digitalhumanities.org. The contents of each filled-out form will be automatically collected at a single location at the server, as well as optionally emailed to Willard. Willard will be

able to interactively approve or reject these membership applications, as well as edit the information supplied in each form, via a special web interface (referred to as the **Membership Approval Interface**). Approval will cause the underlying software to automatically send a subscription request to the Humanist listserver, and to deposit the information from the form in a special database (referred to as the **Membership Database**). This is proposed to be a MySQL database residing at digitalhumanities.org. Every member of the Humanist list will be able to see the (public parts of) all the other members' stored information, edit his own information, and control the details of his list subscription, via a web interface (referred to as the **Listmember Interface**). Access to the interface will be login-and-password controlled. If the login has administrative privileges, the interface will permit viewing and editing of all information stored in the database.

#### *I. 4 Integration of the administrative web interfaces*

For Willard's convenience, it is proposed that the three administrative interfaces described in the present document – the Editorial Interface, the Membership Approval Interface, and the administrative aspect of the Listmember Interface – be integrated as three separate options under a single URL.

#### *II. The new setup of the Humanist List*

The Humanist list will be moved from its current listserver at Princeton to a listserver at digitalhumanities.org. The Princeton listserver is Listserv, whereas the listserver at digitalhumanities.org is Mailman. Given our integration of the listserver settings with the Membership Database (see below) the porting of the subscriber list (which needs to preserve the “no mail” flag settings) will be done via a script which will create for each current member, at one and the same time, a Membership Database entry and a subscription to the list itself. The full move will involve (1) creating the list in Mailman and creating a proper redirect mechanism for the listaddress; (2) configuring the list using Mailman's administrative web interface; (3) disabling Mailman's own web interface for subscribers; and (4) subscribing all the current members to the relocated list (while making sure their “no mail” flags are set properly).

#### *II. 1 Redirection of mail sent to the listaddress*

Mail sent to the listaddress (i.e. to the address conventionally assigned to the list when it is created in Mailman) needs to be automatically redirected as follows: If it is one of the edited digests prepared by Willard and forwarded

to the list by the Editorial Interface, it should go directly to the listserver for distribution to the list. If it is from anywhere else, it should first go through a good spam filter and a good virus filter. If the message survives filtering, two copies of it (one of them optional) will be generated. One copy will be piped to the **Preliminary Cleanup Script** (part of the **Editorial Software**) which performs some preliminary cleanup on the message and then deposits it in a special directory, where it will be accessible from the **Editorial Interface**. The other, optional copy will be forwarded to Willard for his information and preview. Willard, via the Editorial Interface, will be able to turn on or off the generation of these preview copies.

We also need some mechanism whereby Willard can forward to the Preliminary Cleanup Script – and thereby to the Editorial Interface – messages that are intended for the list but sent to him personally. It seems to me that the safest and least complicated way of doing this is to create a separate email address for such forwarded posts, with a separate redirection mechanism. This way, we can configure the list (see below) so that posts from Willard can go through freely, i.e. without requiring approval, and we will not have to worry about possible breakdowns of the redirection mechanism and whether the redirection mechanism can properly distinguish between, on the one hand, posts that Willard forwards to the list from his own mailbox and, on the other, batched digests he creates via the Editorial Interface. [The problem here is that what such forwarded posts look like depends entirely on the mail client one is using and on the exact function (“Forward”, “Redirect”, “Bounce”, etc.) one chooses to achieve the forwarding.]

## **II. 2 List configuration considerations**

Since the mail sent to the Humanist listaddress will be redirected as described above, in principle no posts will ever reach the listserver except the edited digests that Willard creates using the Editorial Interface – which are also the only posts allowed to be distributed to the subscribers. If the principle could be completely realized in practice, there would be no need to configure any posting restrictions (such as moderation) into the list itself. In practice, however, it is possible for the redirection mechanism to break down, which would cause posts that are usually redirected to the Editorial Interface to go, instead, to the listserver. Consequently, the list needs to be configured with appropriate posting restrictions, to prevent such stray emails from going out to the subscribers. The optimal strategy is to configure the list so that only Willard can post – this way his own posts sent to the listaddress (i.e. the batched digests) will go directly to the list, while all the stray posts that somehow evade the redirection mechanism will be held for Willard’s approval. The possible risk of this strategy is that any “spoofed” email which

uses Willard's address as the "From: " address and gets sent to the listaddress will be let through to the subscribers. I think this risk is pretty low, and if we find otherwise we will reconfigure the list accordingly. Consequently, I propose that we configure the list so that all the existing members are set to "moderated" all new members are also set to "moderated" (these are two separate configuration parameters in Mailman) but Willard's "moderated" bit is turned off. This way, Willard will not have to use Mailman's administrative interface to manually approve each of his own edited digests after it has been sent out.

Subscription to the list will in principle be accomplished exclusively via the **Membership Application Form** (see Section VI below). However, since there might occasionally be people attempting to subscribe by sending email directly to the listserver, the list should be configured so that all subscriptions have to be approved by Willard.

## II. 3 What to do about Mailman's web interface for subscribers

Every Mailman list offers a standard web interface whereby people can subscribe, unsubscribe, change their subscription settings, see the subscription list, etc. Anybody can use the interface to attempt to subscribe. Access to the other functions can be configured to be restricted to listmembers, and is then controlled by login ID and password, where the login ID is the email address under which the person is subscribed, and the password is either chosen by the subscriber (at the time of subscription or via a later update) or randomly assigned by the listserver.

In our scheme, however, list subscriptions will be accomplished exclusively via the Membership Application Form (which will solicit from the candidates the additional information needed by Willard and intended ultimately for deposit in the **Membership Database**). The Membership Application Form, in addition to the candidate's email address, will solicit from the candidate a desired password. After the candidate's membership is approved, he or she will be able to use this email address/password combination to access the Membership Database. It stands to reason, therefore, that the very same combination should also allow the member to configure all of his Mailman list options – rather than each member being required to maintain two separate passwords, one for the Membership Database and another one for the listserver. When Willard approves a candidate using the membership Approval Interface, the underlying software will issue an email message to the listserver requesting the subscription of the candidate under the specified email address and with the specified password. And the two passwords need to remain in sync: if the subscriber wants at some point to change his password, the change needs to be effected both at the Membership Database

and at the listserver. Note also that the Membership Interface needs to include a password reminder function.

Consequently, we need to disable Mailman's subscriber web interface for the Humanist list, and incorporate all of its functions into our own Listmember's Interface – which will become the sole interface whereby Humanist members manage both their data in the Membership Database and their subscription to the list. Unfortunately, Mailman does not seem to offer any configuration parameter that disables the subscriber interface, so we will have to find some other way of getting rid of it.

## II. 4 Porting the subscription list to the new listserver

Mailman has a “batch subscription” interface and, if there were no other considerations, it would not be difficult - given a current membership list from the Listserv at Princeton - to batch-subscribe all the current members to the Mailman list. The only necessary adjustment I can think of would be to make sure that members whose subscription is set to “no mail” in Listserv get set to “no mail” in Mailman. All other individual preferences could be configured by the members themselves via the Listmember Interface.

The problem, however, is that in order to access the Listmember Interface the subscribers need to be recorded in the Membership Database, and they need passwords – which, I assume, they do not currently have on the Listserv at Princeton. It seems to me that the way to port the membership list, therefore, is not to use Mailman's “batch subscription” interface, but to write a script that will take as input the current subscription list from Princeton (including the “no mail” flags), and do the following for each entry on the list: (1) generate a random password; (2) create an entry for this member in the Membership Database; (3) generate a Mailman command to subscribe the given address to the list and assign to it the given password; and (4) send to the subscriber an email informing hir of hir ID and password and providing all other requisite information pertaining to the new setup for the list.

## II. 5 Notes

**Important:** It follows from the above that the Membership Database structure and the Listmember Interface needs to be implemented before the actual porting of the Humanist list to the new server.

## III. The Editorial Software and Interface

The Editorial Interface will be a software platform and web interface that will replace the existing editorial software that Willard uses to format and

process the messages he sends to the Humanist listserver. Access to the Interface will be by an ID/password combination (and will be integrated with access to the other administrative interfaces described in the present document).

The existing software, residing at Virginia, performs a certain amount of automatic formatting on the Humanist messages. Some of this formatting is applied to the individual messages before they are digested, and some to the digested messages.

The proposed new software at digitalhumanities.org will perform all of the formatting functions which are performed by the existing software, as well as some additional ones which Willard either performs manually now or would like to have. These additional functions include: (1) replacement of "junk" control strings (generated by copying and pasting into email from typesetting programs) by appropriate ascii characters; (2) deletion (but not automatic detection – see below) of redundant quoted text; (3) scrubbing attachments. This causes the attachments to be detached from the message and placed in a web directory; the URL to the scrubbed attachments is then automatically inserted into the message; (4) conversion of pure HTML into text; and (5) collapsing of multipart/alternative messages to just the first (i.e. text) part. Note that functions (3)-(5) are offered by Mailman itself - but we will not be able to take advantage of this fact, since what we send to Mailman for final distribution to the list are not individual messages but batched digests, and the batching destroys the original MIME structure of the individual messages.

The **Editorial Software** will consist of a set of scripts. One of these scripts, the **Preliminary Cleanup Script**, will be the recipient of the mail redirected from the Humanist listaddress. The script will save the original of each received message in a backup directory (which should be cleaned up every day to remove messages more than, say, a month old), and then perform on the message any cleanup that the current Virginia software performs on individual list messages, as well as the junk-to-ascii translation, attachment scrubbing, HTML-to-text conversion, and collapsing of multipart/alternative messages; then it will deposit the message in a special directory for use by the Editorial Interface. Each message will be put into a separate file in this directory, and the filenames will be generated in such a way that a file containing a message with an earlier "Sent-date" will precede in the lexicographic file-order a file containing a message with a later "Sent-date". The junk-to-text translation will be guided by a replacement table which can be interactively modified from the Editorial Interface – so that if new "junk" sequences are seen in a message, the software's repertoire of replacements can be appropriately updated (after

which Willard will be able to interactively initiate another round of replacements).

The remaining scripts fall into three classes, corresponding to the three functional aspects of the Editorial Interface: (1) those which permit control of the settings used by the automated part of the editorial process; (2) those which aid in the process of inspecting, editing and grouping together the individual messages received from list subscribers; and (3) those which aid in the final inspection and editing of the batched messages and in sending them out to the Humanist list. Although all of these aspects will be presented in a single web interface, they are conceptually and functionally different, and are best described separately.

### III.1 Control of editorial parameters

The following parameters (and any others Willard can think of) will be adjustable from the Editorial Interface:

- The sending (or not sending) of the preview copies of individual list posts to Willard
- The table of junk-character-to-text replacements. This table, used by the editorial software, will be initialized with a number of replacement-pairs, and new ones can be added as the need arises.

After adding new entries to the junk-character-to-text translation table, the Editorial Interface will offer the option of re-applying the translation process to all messages which have already been batched. Since the batching process itself involves an automatic round of junk-character-to-text translation, there should never be any need to reapply the translation to as-yet-unbatched messages.

### III.2 Message grouping

The Message Grouping subsystem of the Editorial Interface will display a Message Queue window listing all the individual messages - in a selected category - that have been received from Humanist listmembers and have not yet been batched (i.e. conglomerated into a "digest" ready to be sent to the list). The "selected category" can be either a previously created thematic group, or the queue of all messages still waiting to be thematically grouped (or rejected). Irrespective of the category selected, the listing of the messages will show the sender, date and subject of each message. For each message, the following choice of actions will be offered:

- Put (or move) the message into a specific thematic group. The interface will offer a choice of previously created thematic groups, as well as the option of creating a new one.
- Reject the message.
- Rerun the Preliminary Cleanup Script on the original of this message (and replace the current cleaned-up version with the new one). This will be useful if the preliminary Cleanup Script needs to be updated to accommodate some new feature found in the given message.
- Display the full contents of the message in a separate **Message Display** window. These contents will already have been edited by the Preliminary Cleanup Script. The Message Display window will offer the following choice of actions upon the given message:
  - Edit the message interactively.
  - Mark a point in the body of the message and ask that all text following that point be deleted. This is proposed as an improvement over manual elimination of redundant quotes, and saves us from having to teach the software how to recognize either quoting or redundancy.
  - Rerun the Preliminary Cleanup Script on the original of this message (see above).
  - Put (or move) the message into a specific thematic group (existing or new).
  - Reject the message (i.e. delete it from the system).

It is proposed that whenever a message gets either rejected or moved to a different category than the one that is currently being displayed, it will disappear from the current Message Queue window (so that only messages still in the given category will appear in the window).

Rejected messages will be deleted from the directory used by the Editorial Interface, and their (“scrubbed”) attachments will be deleted from the web directory in which attachments are stored. The originals of these messages will remain in the backup directory until they become old enough to be removed by the daily cleanup script.

Messages on which no action has been taken will stay wherever they were, for future access via the Editorial Interface.

### III.3 Batch creation and dispatching

The Batch Creation subsystem of the Editorial Interface will display a list of all currently active thematic groups – i.e. all existing thematic groups which contain some messages (either consolidated into a digest or not). The list will indicate, for each thematic category, whether or not the

messages contained in it have already been consolidated. For each group, the following choice of actions will be offered:

- Consolidate the grouped messages into a single digest. This consolidation will include some automatic formatting of the resulting digest, much along the lines of the existing software at Virginia (e.g. building a table of contents, adding a header containing the volume and issue number, etc.), as well as a round of junk-character-to-text translation.
- Send the digest to the Humanist list.
- Display the consolidated digest in a separate **Digest Display** window. The Digest Display window will offer the following choice of actions upon the given digest:
  - Edit the digest interactively.
  - Rerun junk-character-to-text replacement on the digest.
  - Send the digest to the Humanist list.

In addition, the interface will also offer the option of initiating a round of junk-character-to-text replacement on all batched messages (to accommodate additions to the translation table).

It is proposed that whenever the contents of a thematic group get consolidated and the resulting digest sent out, this thematic group will disappear from the current Batch Creation window (so that only thematic groups still requiring some action will appear in the window).

#### III.4 Notes

The more recent archives at Virginia (see, for example, [http://www.iath.virginia.edu/lists\\_archive/Humanist/v21/subject.html](http://www.iath.virginia.edu/lists_archive/Humanist/v21/subject.html)) have the annoying characteristic that the authors of the individual messages are not featured in the message index - all messages look uniformly as if they were authored by "Humanist Discussion Group (by way of Willard McCarty)". Whatever the cause of this, I propose that the editorial software should format the messages in such a way as to permit the original senders to figure properly in the batched messages and in the archives.

The messages sent to the list need to be sanitized before they are displayed in any of the Web interfaces, so as to guard against the possibility of the message containing malicious code.

The batch-creation software will automatically change the volume and issue numbers on May 7<sup>th</sup> of each year, as per Willard's current manual practice.

#### IV. Moving the Humanist webpage

The Humanist webpage will be moved from its current location at <http://www.princeton.edu/humanist/> to [digitalhumanities.org](http://digitalhumanities.org). It looks to me like most of this move should be pretty straightforward. There are two obvious exceptions to this overall straightforwardness: the parts pertaining to the archives, and the Humanist Subscription Form. More about them below.

#### V. Moving the Humanist list archives

The list archives, which currently reside at Virginia, will be moved to [digitalhumanities.org](http://digitalhumanities.org). This move has four aspects:

- (1) An archiving infrastructure for the “live” list – i.e. the ongoing list activity - will have to be designed. I don’t know what the current infrastructure is at Virginia, and whether it is desirable and feasible to imitate it. Mailman keeps an archive for every list, but Willard will have to decide if he likes it enough to make it the “real” archive for the ongoing activity at Humanist. To view an example of a Mailman archive, go to <http://lists.puptcrit.org/mailman/private/puptcrit/>, and log in with the email address “[ma@panix.com](mailto:ma@panix.com)” and the password “azesadex”. Depending on what is decided, appropriate archiving software may have to be installed or written.
- (2) A display strategy for the archives – both the old ones and the “live” one - will have to be decided upon. This aspect and the preceding aspect are very intimately tied together. Depending on what is decided, appropriate display-supporting software may have to be installed or written.
- (3) The archive files currently residing at Virginia will have to be tar-zipped, brought over to a designated location at the [digitalhumanities.org](http://digitalhumanities.org) server, and made available for display.
- (4) The Humanist website will have to be modified so as to display the archives – both the old ones and the “live” ones - in their new location. The nature of these modifications will depend on the decisions made with respect to the archiving and display infrastructure at the [digitalhumanities.org](http://digitalhumanities.org) server.

##### V.1 Notes

Whatever archiving software we decide upon, we should make available for download the raw archives in some sizable chunks – e.g. one year’s worth of posts.

Some component of the archive software will automatically change the archive volume on May 7<sup>th</sup> of each year, as per Shayne Brandon’s current

manual practice. The same or a different component will also appropriately modify the archive-link section of the Humanist website to include a link to the new archive volume.

**Budget Note:** The move of the archive from Virginia to digitalhumanities.org was not mentioned in the original specification, and therefore was not included in my original rough estimate of hours-of-work required.

## VI. The Humanist subscription process and the Membership Database

The subscription process to the Humanist list will begin with the candidate filling out the **Membership Application Form**, which will be located on the Humanist website. The following data will be solicited by the form full name, email address, desired password, biographical blurb and (optional) URL. The email/password combination will, after the subscription is finalized, enable the subscriber to log into the **Listmember Interface** (access-controlled by Apache's mod\_access mechanism), from which the subscriber will be able to view the other subscribers' "public" data in the **Membership Database**, modify hir own data, and control the Mailman settings for hir subscription.

The data entered in the form will be received by a cgi script, the **Membership Application Script**. The script will perform a sufficiency and sensicality check on the data, prompt the user for any required corrections, and sanitize the completed contents. It will then email one copy of the application to Willard (for his information and preview), and place another copy in a special directory on the server, for use by the **Membership Approval Interface**. This is the web interface whereby Willard will be able to edit, approve or reject the accumulated membership applications.

When an application is approved, the script underlying the Membership Approval Interface will make a new record for this member in the Membership Database. It will also send an email to the Humanist listserver with a request to subscribe the given applicant (under the specified email and with the specified password) to the Humanist list. This will cause the listserver to send to the new member a "Welcome to the list" message, which should be crafted so as to provide all the information Willard wishes to be received by new listmembers.

Every listmember will be able to log in to the Listmember Interface and view the public fields (everything except the password and the listserver settings) of other members' records in the Membership Database, modify hir own (full) record, including hir listserver settings. In addition, if the login is detected to be that of the administrator, the interface will permit the viewing of all fields

(not just the “public” ones), editing or deletion of any record, and creation of new records from scratch. The script underlying the Listmember Interface will detect any modification to information that concerns the listserver, and will automatically dispatch to the listserver an email requesting the appropriate setting adjustments.

#### VI.1 The Membership Database

It is proposed that the Membership Database be a MySQL database. The following information will be stored for each member: full name, email address , password, biographical blurb, URL, and the Mailman settings for hir account. The database will also contain a separate Deleted Members table for archiving of deleted records.

#### VI.2 The Membership Application Form

The Membership Application form will contain fields for the full name, email address , password, biographical blurb, and URL. All the data except for the URL will be obligatory. The script underlying the Membership Application Form will perform completeness and sensicality checks on the data entered by the prospective listmember, will prompt for missing data and corrections, and will sanitize the user’s input to protect against malicious exploits. Then the script will email one copy of the completed form to Willard, and deposit another copy in a separate file in a special directory for use by the **Membership Approval Interface**.

#### VI.3 The Membership Approval Interface

Access to the Membership Approval Interface will be by an ID/password combination (and will be integrated with access to the other administrative interfaces described in the present document). The Membership Approval Interface will display a compact list of all membership applications currently waiting to be approved. The list will indicate, for each application, the full name and email address of the applicant. For each application, the following choice of actions will be offered:

- Approve the application.
- Reject the application.
- Display the application’s contents in a separate **Application Display** window. The Application Display window will offer the following choice of actions upon the given application:

- Edit the contents interactively.
- Send an email to the applicant. This will simply act like a “mailto” field in a webpage – i.e. open a composition window in Willard’s mail client.
- Approve the application
- Reject the application.

It is proposed that whenever an application is either approved or rejected, the corresponding entry will disappear from the Membership Approval window (so that only application still waiting for a decision will appear in the window).

When an application is approved, the script underlying the Membership Approval Interface will make a new record for this member in the Membership Database, copy into this record the information supplied within the application, and initialize to the Mailman defaults the fields pertaining to the Mailman settings for this member’s account. It will also dispatch an email to the listserver with a request to subscribe the given applicant (under the specified email address and with the specified password) to the Humanist list. Lastly, it will add an entry for this member to the .htpasswd file which controls access to the Listserver Interface.

Unless Willard would like something different, rejected applications will simply be deleted from the system.

#### VI.4 The Listmember Interface

Access to the Listmember Interface will be by an ID/password combination. In the case of the Humanist administrator, this access will be integrated with access to the other administrative interfaces described in the present document.

The “top level” interface that solicits the ID/ password combination will have a “password reminder” option that will send to the specified email address (which has to match an address in the database) the corresponding password.

The Listmember Interface will display a compact list of all current members of the Humanist list – say, 100 entries per page. The list will be ordered alphabetically by last name, and will consist of full names, email addresses, and institutional associations. An option will be provided to jump to names that begin with a specific letter of the alphabet. For each member listed, the following choice of actions will be offered:

- (Administrator or given member only) Delete the member from Humanist. The result of this action will be three-fold: the member’s

record will be moved to a “Deleted Members” table in the Membership Database, the corresponding entry will be deleted from the .htaccess file controlling access to the Listmember Interface, and an “unsubscribe” command will be dispatched, via email, to the listserver.

- Send an email to the given member. This will simply act like a “mailto” field in a webpage – i.e. open a composition window in the user’s mail client.
- Display the contents of this member’s database record in a separate **Database Record Display** window. Depending on whether or not the logged-in person is either the administrator or the member whose record is being displayed, the window will display either the full record for this member, or only the “public” fields. The window will offer the following choice of actions upon the given record:
  - (Administrator or given listmember only) Edit the contents of the record interactively. The changed contents will be subject to the same checks and sanitation as the Membership Application Form. Performing any changes that have an effect on the Mailman settings will result in an appropriate email being dispatched to the listserver and, if applicable, an appropriate update to the .htaccess file which controls access to the Listmember Interface.
  - (Administrator only) Delete the member from Humanist. The results will be as described above.
  - Send an email to the given member (as above).

Additionally, the administrator will be offered the option to create a new record in the Membership Database. Selection of this option will open up a new window containing a form with fields corresponding to those in the Membership Database. Completion of the form will result in the same automatic actions as those described above for the a Membership Application Form.

## VI. 5 Notes

Willard currently has a large backlog of membership applications that need processing. Once the Humanist framework is fully operational on the new server, a script can be written to directly processes these applications as if they had all been approved.

The database needs to be periodically backed up. Should this be done by a cron job, or initiated manually from one of the administrative interfaces?

Should we save the approved membership applications somewhere in the system, in case we need to recreate the database? Is there in fact any

likelihood that the database may ever have to be (partially or fully) recreated?

A cron job will periodically remind every members to review hir database entry and update it as desired.

# **The new Humanist List framework – Implementation Milestones**

Malgosia Askanas  
(ed Willard McCarty 15/6/08)

## **1. Website port and directory structure**

- Copy the existing Humanist website to the new server (without making it public as yet).
- Create the underlying directory structure for the new framework.

## **2. Administrative Interface: Top-level web structures**

- Create the top-level webpage of the Administrative Interface, and mockups (i.e. web pages without the underlying scripts) of the Editorial Interface, the Membership Approval Interface, and the administrative aspect of the Listmember Interface.

## **3. Membership Database**

- Create the database itself and the requisite table structures within it.

## **4. Membership Application Form**

- Create the web form and implement the underlying script.

## **5. Membership Approval Interface**

- Implement the underlying script (the web form has been created earlier).

## **6. Listmember Interface**

- Create the web forms and implement the underlying scripts.

## **7. Humanist list – creation and configuration**

- Create the list in Mailman
- Create a proper redirect mechanism (through a virus filter, a spam filter, and a Preliminary Cleanup Script stub) for the listaddress, and another

- one (just through the Preliminary Cleanup Script stub) for the address that Willard will use to forward posts from his own mailbox to the list
- Configure the list using Mailman's administrative web interface
  - Disable Mailman's own web interface for subscribers

#### **8. Preliminary Cleanup Script – 1: Duplication of existing functionality**

- This version of the script will perform the functions performed by the corresponding script at Virginia, and deposit the outputs, in the proper order, in the Editorial Interface directory.

#### **9. Preliminary Cleanup Script – 2: Addition of new functions**

- The additional functions are: junk-to-ascii translation, attachment scrubbing, HTML-to-text conversion, and collapsing of multipart/alternative messages.

#### **10. Editorial Interface – Message Grouping**

- Implement the still-missing web aspects (e.g. the Message Display Window) and the underlying script.

#### **11. Editorial Interface – Batch Creation**

- Implement the still-missing web aspects (e.g. the Digest Display Window) and the underlying script.

#### **12. Editorial Interface - Parameter Control**

- The web aspects are done at this point, so this is just the script.

#### **13. List archives**

- Move the existing list archives and implement a scheme for the "live" archive. Scheme has to be decided upon. Implementation may require help from John Unsworth and Shayne Brandon.
- Make appropriate modifications to the Archive page of the new Humanist website.
- At this point, the new website is ready to be made public.

#### **14. Subscription list move**

- Obtain from the Princeton ListServe a list of all current Humanist subscribers (full names, emails addresses, and "no mail" status)
- Create a script which accepts this list as input, creates for each current member a random password, creates an entry for this member in the Membership Database, subscribes this member to the new Mailman list, updates the .htaccess file that controls access to the Membership Interface, and sends the member an informational email.
- Use this script to populate the Membership database and the new Mailman list.
- At this point, as soon as Willard has released any still-waiting posts to the old list, the list address can be switched from Princeton to the new server.

#### **15. Odds and ends**

- Includes implementation of the various cron jobs, any tying up of any other loose ends.