Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2005),
Dallas, TX, September 2005, pp. 207-214.

Estimating the Numbers of End Users and End User Programmers

Christopher Scaffidi
School of Computer Science
Carnegie Mellon University
cscaffid+isri@cs.cmu.edu

Mary Shaw
Sloan Software Industry Center
Carnegie Mellon University
mary.shaw@cs.cmu.edu

Brad Myers
School of Computer Science
Carnegie Mellon University
bam@cs.cmu.edu

Abstract

In 1995, Boehm predicted that by 2005, there would be "55 million performers" of "end user programming" in the United States. The original context and method which generated this number had two weaknesses, both of which we address. First, it relies on undocumented, judgment based factors to estimate the number of end user programmers based on the total number of end users; we address this weakness by identifying specific end user sub populations and then estimating their sizes. Second, Boehm's estimate relies on additional undocumented, judgment based factors to adjust for rising computer usage rates; we address this weakness by integrating fresh Bureau of Labor Statistics (BLS) data and projections as well as a richer estimation method.

With these improvements to Boehm's method, we estimate that in 2012 there will be 90 million end users in American workplaces. Of these, we anticipate that over 55 million will use spreadsheets or databases (and therefore may potentially program), while over 13 million will describe themselves as programmers, compared to BLS projections of fewer than 3 million professional programmers. We have validated our improved method by generating estimates for 2001 and 2003, then verifying that our estimates are consistent with existing estimates from other sources.

1. Introduction

As first reported in 1995 [2] and widely disseminated as part of the Construction Cost Model (COCOMO) version 2.0 [3], Boehm et al. predicted that the number of "end user programming performers" would reach 55 million by 2005. This number originally functioned as an estimate of the number of people who would not benefit from COCOMO 2.0, thereby bounding the applicability of that proposed model. Since 1995, numerous researchers have cited his prediction in order to motivate research into visual languages and other programming environments for scientists, secretaries, engineers, salespersons, and other non-professional programmers (for example, [5] [11] [14] [16] [20] and [21]).

Boehm's method depends on a 1989 survey by the US Bureau of Labor Statistics (BLS) that asked each respondent about personal computer usage at work. His prediction method begins by assuming that those usage rates would not change, and it makes the simplification that all end users would eventually perform programming-like tasks. His method then copes with these approximations by incorporating undocumented multiplicative factors to adjust for the rising usage of computers and the fact that not all end users are programmers. Based on these assumptions, Boehm's method yields the prediction of 55 million.

Incorporating more recent BLS data allows us to relax some of Boehm's assumptions and update the estimate. Empirically, BLS data shows the rate of computer usage significantly rose through the 1990's [11]. Using this information, we now estimate that the end user population at workplaces in the United States is about 80 million in 2005 and will be 90 million in 2012. In contrast, for each year, BLS estimates fewer than 3 million professional programmers.

As the structure of this burgeoning end user population becomes clearer, we anticipate that broad sub-populations will emerge, each with characteristic needs and abilities. For example, while only 15% of end users reported they "did programming" at work in 2001, 62% reported they used computers for spreadsheets or databases. Some, but not all, of these end users utilized these tools for programming. Staking out the bounds of end user programmer sub-populations will allow researchers to target studies and tool-development efforts at large or interesting sub-populations, thereby helping to unlock their potential.

Section 2 of this paper examines Boehm's original context and prediction method more closely. Section 3 improves on the method by incorporating innovation

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

diffusion theory [19] to take into account the rising rate of computer usage among American workers. Section 4 briefly describes a number of other estimates of the current end user and end user programmer populations.

2. The original "55 million" prediction

This section reviews the context and method that generated Boehm's prediction in order to highlight the weaknesses that we address later. Specifically, we note that Boehm's method depends on judgment-based factors to account for rising computer usage rates and the fact that not all end users are programmers.

2.1. Original context

In 1995, Boehm et al. introduced COCOMO version 2.0, an extension and elaboration of the earlier COCOMO cost-estimation framework, in order to address various shortcomings in the earlier version [2]. Many of these shortcomings had arisen due to changes in the way professional programmers performed their trade (such as the increasing focus on producing reusable modules). To accommodate these changes, COCOMO 2.0 incorporated a variety of new variables representing cost drivers ranging from "team cohesion" to "process maturity" in an attempt to better capture the details of the professional programming environment.

Many variables in both versions of COCOMO do not apply to end user programmers. These workers, who range from marketing specialists to business administrators, generally build small software. They often develop their software on their own rather than in organized teams (making the "team cohesion" variable less meaningful), and they make little attempt to apply a disciplined software development process (making the "process maturity" variable less meaningful). In the words of the original COCOMO 2.0 technical report, "The User Programming sector does not need a COCOMO 2.0 model. Its applications are typically developed in hours to days, so a simple activity-based estimate will generally be sufficient" [2].

Boehm not only recognized COCOMO 2.0's limited applicability, but he also took pains to estimate how many future programmers would be end users having little need for COCOMO 2.0. Through the method discussed below, he projected 55 million end user programmers by 2005. This contrasted in his report with a prediction of fewer than 3 million professional programmers (who could make use of COCOMO 2.0) in 2005. This showed that while COCOMO 2.0 represented a great leap forward, there remained a wide variety of situations in which it did not apply.

2.2. Original method

Boehm's original technical report [2] contains a footnote outlining his method:

"These figures are judgement-based extensions of the Bureau of Labor Statistics moderate-growth labor distribution scenario for the year 2005 [CSTB 1993; Silvestri and Lukaseiwicz 1991]. The 55 million End User programming figure was obtained by applying judgement based extrapolations of the 1989 Bureau of the Census data on computer usage fractions by occupation [Kominski 1991] to generate end user programming fractions by occupation category. These were then applied to the 2005 occupation-category populations (e.g., 10% of the 25M people in "Service Occupations"; 40% of the 17M people in "Marketing and Sales Occupations)."

The starting point is Silvestri and Lukaseiwicz's projection of the growth of the American worker population in each of nine occupational areas to 2005 [18]. They achieved this by combining a moderate-GDP-growth model of the economy (developed by BLS) with a variety of survey data (including BLS's Current Employment Statistics, Occupational Employment Statistics, and Current Population Survey). This allowed them to predict the number of people working in each occupational category in 2005 (see the second column in Table 1).

The next step is to estimate the rate of computer usage in the workplace. This comes from Kominski's analysis of the 1989 Current Population Survey (CPS) of households (discussed in [11]). This survey, run by BLS, included items asking respondents whether they used a computer at work. As the survey also included items that facilitated occupational categorization, it was possible to determine what fraction of each occupation's workers actually used a computer at work (as shown in the third column of Table 1). For example, roughly 10% of "Service Occupation" workers used a computer at work.

The third step takes these 1989 usage fractions and multiplies them against the projected 2005 occupational population sizes, then adds up the results as shown in the fourth column of Table 1 below. The result thus predicts the number of end users in 2005.

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

Table 1. Multiplying occupational projections by computer usage yields approximately 55 Million

Occupational Category	Projected Occupational Count in 2005 (in thousands)	Actual Percentage Using Computers at Work in 1989	Projected Computer Usage at Work in 2005 (in thousands)
	Silvestri [18]	CPS [11]	Computed
Managerial and Professional	36,773	56.2%	20,666
Technical, Sales, Administration	48,078	55.1	26,491
Precision Prod., Craft, Repair	15,909	15.3	2,434
Service	24,806	10.2	2,530
Operators, Laborers, Fabricators	17,961	9.5	1,706
Farming, Forestry, Fishing	3,665	4.4	161
Totals:	**147,192**		**53,989**

Finally, each of these rows is then multiplied by a factor to adjust for rising computer usage rates in that occupation, and then multiplied by another factor to deal with varying rates of programming-like behavior among end users of that occupation[1]. It happens that these two factors generally cancelled out, yielding a net projected end user programmer population of 55 million, not far from the 54 million shown in Table 1.

In passing, we note that [2] also includes four predictions concerning subcategories of professional programmers totaling a projected 2.75 million for 2005. As discussed later, this is broadly consistent with information from the government: BLS data indicate there were fewer than 3 million professional programmers in 2001, and BLS projections suggest this number will remain under 3 million in 2012.

2.3. Discussion

Boehm's method multiplies occupational population projections for 2005 against computer usage rates from 1989. This method is organized around two approximations, each requiring a correction factor.

The first approximation is that all end users are practitioners of programming. Nardi defined "programming" as "behavior in which sequences of procedural instructions… are written in a language that is compiled or interpreted," then broadened her definition to include any attempt "to create an application that serves some function" [13]. Blackwell has noted [1] that a wide variety of software and consumer products do indeed allow end users to engage in programming-like activities, such as recording behavior or algorithms for playback later.

[1] Personal communication between Barry Boehm and Christopher Scaffidi, May 19, 2005.

Clearly, not all end users perform programming activities to the same extent. For example, claiming that browsing or emailing constitutes programming requires a fairly inclusive but esoteric view of what programming entails. One argument might be, "When you enter an email, you type a series of commands and transmitting them to a machine, which acts on those commands to decide what to display on the screen. When you browse the web, you click the mouse in a way that enters commands into the computer, which executes those instructions like any Turing machine."

There is something inherently unsatisfying about this argument. Lumping all end user programmers into one estimate misses an opportunity to add texture for guiding further research. We address this in Section 4.

The method's second approximation is that computer usage rates within each occupational category remain constant. Section 3 develops an improved approximation by extending the method to use data collected more recently than 1995.

3. Improving the prediction method

In this section, we incorporate innovation diffusion concepts and fresh BLS data to develop a model-based approach to adjusting for rising computer usage rates. We can then replace some of Boehm's judgment-based factors and provide a better estimate of the end user population size in 2005 as well as predictions for 2012.

3.1. Empirically, the estimate is too low

Boehm's original prediction was derived from 1989 CPS data (plus other data). In fact, BLS worked with the Bureau of the Census to collect these same CPS computer usage items on surveys in 1984, 1989, 1993, and 1997. (The 2001 CPS also asked about computer usage, but with a slightly different wording. We will

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

return to this issue below.) These have been tabulated to show computer usage trends for each occupational category [11], as shown in Figure 1.

BLS's data reveal that a larger fraction of workers use a computer now than in 1989. For example, technical-sales-administration-support occupational workers reported a usage rate of almost 70% in 1997, compared to only 55% in 1989. As a result, using computer usage rates from 1989 under-estimates the total number of computer users in 2005. In fact, the CPS showed that in 1997, over 64 million used the computer at American workplaces.

3.2. Extending the prediction method

The key to extending Boehm's method is to model the salient S-shaped curve of some occupations' computer usage, which is particularly apparent for the lower four curves of Figure 1. (The top two curves resemble the right halves of S-curves, with the left halves occurring prior to the first survey in 1984.) This "logistic curve" typifies many diffusion phenomena—ranging from the propagation of a virus through a population to the adoption of technological innovations [19]. In each case, a phenomenon of interest (such as infection) initially affects only a small fraction of the population. However, as those people interact with the population, they share the phenomenon, causing the incidence to increase until it affects virtually everybody. The rate of increase starts flat (since there are few people promulgating it) and ends flat (since the population is nearly saturated).

Although we used only four CPS data points, it is interesting to fit each curve to the theoretical functional form of innovation diffusion phenomena in order to determine where the curves might be headed in the future. (More complex functions exist for certain contexts, but it makes sense to use one of the simpler but successful forms when only four data points are available. See [19] for a discussion of other possible functional forms.) The S-shaped logistic form is

$$A(t) = \frac{K}{1 + e^{(-m \cdot t + b)}}$$

A(t) = adoption (usage) at time t
K = maximum adoption
m, b = fitting parameters (related to rate of adoption, and the "zero-point" of t)

A least-squares fit for each occupation's computer usage rates yields a function approximately describing how usage will develop in the future.

As with Boehm's original method, we can multiply these now-improved usage estimates against occupational head counts to estimate the total number of end users in the future. For example, inserting t=2005 into the fitted functions and multiplying by actual January 2005 CPS occupational head counts [4] reveals that approximately 81 million people currently use computers at American workplaces, as shown in Table 2. (Note that the January 2005 CPS does not include computer usage questions, only occupational and other "basic" items.)

Predicting the end user population's future size requires projections of future occupational head counts. Unfortunately, BLS only issues projections for one year per decade; hence, the only projection available is for the year 2012 [9]. We insert t=2012 into the fitted functions to predict computer usage rates for 2012 and multiply each rate by the corresponding projected head count. As shown in Table 3, this yields the prediction that approximately 90 million American workers will use a computer at work in 2012.

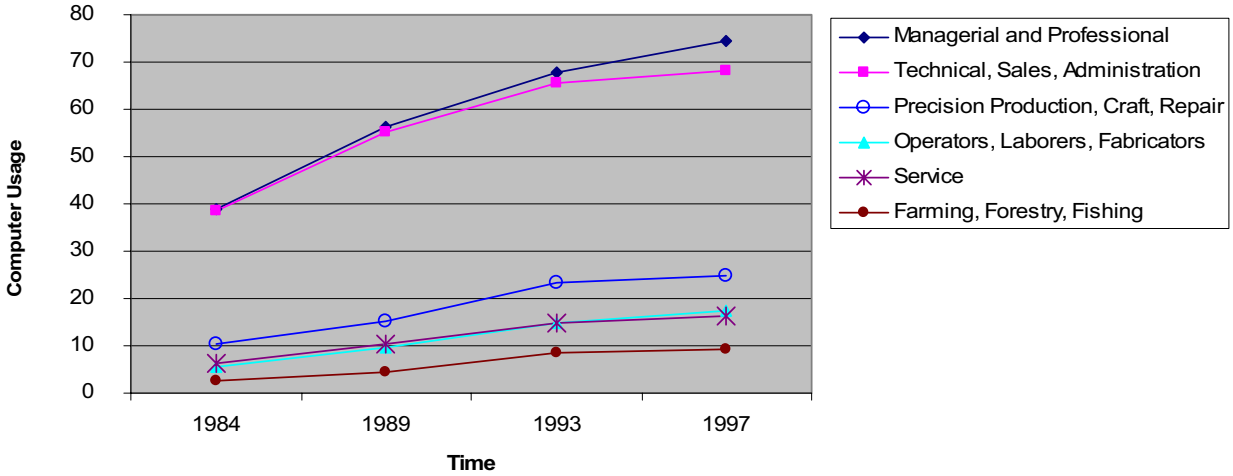

Figure 1. Rising computer usage rates among American workers

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

These estimates illustrate an enhanced method for use in generating estimates of end user populations. It has removed the judgment-based factors that adjusted for rising computer usage rates; instead, this method assumes that computer usage rates can be modeled by a simple innovation diffusion curve. Such an assumption is somewhat suspect, in part because it presumes that the innovation under discussion (here, the computer) does not change substantively during the course of diffusion. This is, of course, not true: computers continually increase in power and utility. Hence, future computer usage rates will likely exceed those indicated by the foregoing model. Therefore, 90 million represents a lower bound for the number of American end users in workplaces in 2012, any of whom might perform programming-like activities.

4. Comparisons to other counts

The BLS always asks about occupation in the CPS. Moreover, in addition to computer usage at work, home and school, the 1989 and subsequent CPS's asked about software usage. These data reveal that the number of end users potentially programming vastly exceeds the number of professional programmers, and indeed exceeds the number of self-reported end user programmers. End users evidence a diversity of software usage practices at work and outside work.

4.1. Other counts of end users

Other counts of end users provide a way to check the method that we have developed. In the past, researchers produced at least two alternative counts of the end user population, each of which agrees well with comparable estimates generated by our method.

The first alternative count is the 2001 CPS. As briefly mentioned earlier, we did not include the 2001 CPS computer usage data in the fit discussed above. Unfortunately, the question's wording changed from "Do you directly use a computer at work?" in 1997 and prior years to "Do you use a computer at your main job?" in 2001 [4]. Under this new wording, the computer usage rate shot up sharply between 1997 and

Table 2: We estimate that 80 million people use computers at American workplaces in 2005

Occupational Category	Actual Occupational Count in 2005 (in thousands)	Estimated Percentage Using Computer at Work in 2005	Estimated Computer Usage at Work in 2005 (in thousands)
	CPS [4]	Fit to CPS ([10] data)	Computed
Managerial and Professional	50,584	81.3%	41,136
Technical, Sales, Administration	37,943	71.6	27,185
Precision Prod., Craft, Repair	15,010	29.0	4,358
Service	24,731	18.7	4,617
Operators, Laborers, Fabricators	19,648	20.4	4,010
Farming, Forestry, Fishing	1,022	10.9	112
Totals:	**148,938**		**81,418**

Table 3: We estimate that 90 million people will use computers at American workplaces in 2012

Occupational Category	Projected Occupational Count in 2012 (in thousands)	Projected Percentage Using Computer at Work in 2012	Projected Computer Usage at Work in 2012 (in thousands)
	Hecker [9]	Fit to CPS ([10] data)	computed
Managerial and Professional	52,030	83.0%	43,209
Technical, Sales, Administration	42,695	72.1	30,804
Precision Prod., Craft, Repair	14,860	29.9	4,442
Service	31,905	19.1	6,098
Operators, Laborers, Fabricators	22,723	21.0	4,782
Farming, Forestry, Fishing	1,107	11.1	123
Totals:	**165,320**		**89,459**

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

2001 for two occupational groups (service and farming), and in fact no longer appears S-shaped in one case (farming). Nonetheless, if we set t=2001 and multiply by occupational population sizes, the model would predict that 71.9 million people used computers at American workplaces in 2001; this compares extremely well with actual 2001 responses to the new question wording, which indicated that 72.3 million people used a computer at American workplaces. If we redo our model, using the 2001 survey as another data point and fitting the farming data to an exponential rather than logistic curve, the resulting model would cause us to revise our 2012 prediction slightly upward to 96 million end users in American workplaces, rather than 90 million.

The second alternative count of the end user population is a Forrester survey commissioned by Microsoft in 2003, which indicated that 129 million people between ages 18 and 64 used a computer in American workplaces or homes [7]. If we plug t=2003 into our model and multiply by actual occupational population sizes reported by the 2003 CPS, our model estimates that 80 million people used a computer at American workplaces in 2003; since the 2001 CPS indicated that 97% of all workplace end users were between 18 and 64, we can safely approximate that nearly all of our estimated 80 million workplace end users of 2003 were between ages 18 and 64. In order to compare this number to the Forrester number, we must adjust by the number of people age 18 through 64 who used a computer at home but not at work. As this number was 43 million in 2001, according to the 2001 CPS, we can add at least 43 million to our 2003 estimate of 80 million adult workplace users to yield an estimate of 123 American home and workplace end users in 2003 between ages 18 and 64, in rough agreement with the Forrester number of 129 million.

In short, our model generates estimates for 2001 and 2003 that are broadly consistent with other data.

4.2. Counts of professional programmers

Boehm predicted that professional programmers would number around 2.75 million by 2005 [2], and that prediction has proven roughly correct. According to CPS data, the total number of American "computer scientists and systems analysts," "computer programmers," and "computer software engineers" ranged from 2.0 to 2.5 million between 2000 and 2005 [4]. Moreover, BLS projects that the total number of programmers in these categories will remain under 3 million through 2012 [9]. Clearly, professional programmers currently constitute, and will continue to constitute, a tiny minority of all computer users.

4.3. Counts of end user programmers

CPS interviews since 1989 have included a number of questions revealing software usage trends in the workplace end user programming population:

"Do you do programming?" (1989, 1993, 1997, 2001)
"Do you use spreadsheets?" (1989, 1993, 1997)
"Do you use databases?" (1989, 1993, 1997)
"Do you use spreadsheets or databases?" (2001)

The data reveal that usage of end user programming environments grew steeply throughout the 1990's. Specifically, around 10% of American end user workers reported "using spreadsheets" in 1989, and by 1997 this had risen to over 30%. Likewise, around 10% of American end user workers reported "using databases" in 1989, and by 1997 this had also grown to over 30%. Over the next four years, usage of end user programming environments continued to explode, with over 60% of American end user workers reporting that they "used spreadsheets or databases" in 2001. This amounted to over 45 million end users of spreadsheets or databases in American workplaces.

In contrast, the proportion of American end user workers reporting they "do programming" has remained relatively constant, rising from around 10% in 1989 to only around 15% in 2001 (about 11 million end users). While this number still greatly exceeds the number of professional programmers, it does not approach the number of spreadsheet and database users, any of whom might be said to "write programs." In fact, in 2001, only 20% of workers who "used spreadsheets or databases" indicated that they also "do programming." (While the foregoing discussion has focused on aggregate rates, averaged over all occupational groups, the same striking trends and contrasts appear upon consideration of each occupational group individually.)

Two hypotheses could account for the growing divergence between the number of self-reported end user programmers and the number of spreadsheet/ database users.

One possibility is that many end users *do* create programs using spreadsheets and databases, but that they prefer not to use the term "programming" to describe their own activities. In such a case, simply asking end users whether they "program" will provide little guidance on whether or how to improve end user software development environments. Instead, we must identify the key stumbling blocks and key opportunities *within* programming and then specifically study those aspects of end user practice. We have argued in a separate paper [17] that the use of

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

abstraction can facilitate or impede achieving key software engineering goals and that therefore this aspect of end user programming deserves additional study in order to guide further tool development.

Another hypothesis is that many end users rely on spreadsheets and databases as a place to store information but generally *do not* create programs in those environments. There is some support for this hypothesis: Hall's study of spreadsheets created by well-educated Australian workers found that only 47% used the "if" function [7], while Fisher and Rothermal's survey of spreadsheets on the web revealed that only 44% contained any formulas at all [6]. These numbers suggest that approximately half of spreadsheet users use conditionals or formulas to create programs, which is consistent with a very broad definition of "programming" by end users.

If spreadsheet/database usage rates stay constant or rise, then at least 60% of our projected 90 million American end user workers in 2012 will utilize spreadsheets or databases (as reported in the 2001 CPS mentioned earlier). If half of these write programs, then we expect that in 2012, around 25 million spreadsheet and database end users could be programming. In contrast, if only around 15% of end users continue to report they "do programming," then the number of self-reported end users will number around 13 million.

Additional data collection concerning this large population, particularly studies with a tightly focused definition of "programming," should help to clarify what end users are creating with their programming environments and how to improve those tools.

4.4. Counts of end users outside workplaces

In the 1997 and 2001 questionnaires, BLS asked about computer usage by people at home and school [4]. They also asked about software usage at home. The data reveal a large population of home and school end users whose needs differ from the population of end users at work. In particular, relatively few people use spreadsheets or databases outside work.

Although the total number of computer users outside work is growing, a decreasing fraction of them utilize spreadsheets or databases. Specifically, in 1997, approximately 37 million people did not use a computer at American workplaces but did use one at home or school. Of these, approximately 19% (7 million) used spreadsheets or databases at home. By 2001, this had shrunk to approximately 13% (13 million out of approximately 101 million people) [4].

In contrast, during the same time period, the fraction of workplace end users who utilized spreadsheets rose sharply from 32% to 62% (21 of 64 million in 1997, to 45 of 73 million in 2001).

In short the large population of end users at home and school utilizes spreadsheets and databases at a lower rate than end users at work, and this divergence is rapidly growing. We expect that end users of different populations have differing needs, interests, and capabilities. Ideally, research will yield a diversity of environments to help these populations attain the most value possible with respect to personal needs.

5. Conclusion

We extended the "55 million" estimation method by modeling rising computer usage rates. We predict that 90 million American workers will use computers in 2012, significantly exceeding the 3 million anticipated professional programmers. We project that in 2012, over 13 million workers will "do programming" in a self-reporting sense; however, based on spreadsheet and database usage, it seems likely that the number of end user programmers will be much higher. Consequently, end user programming environments will remain a vital research area for many years.

In the future, we will perform additional analyses of the CPS data, as well as collect and analyze new survey data, in order to better understand the software usage of end users. We anticipate that more precise estimates and characterizations of end user practices will help researchers target further work in developing languages and tools to assist end users in programming tasks.

6. Acknowledgements

The authors thank Barry Boehm for helpful discussions about his original method. This work has been funded in part by the EUSES Consortium via the National Science Foundation (ITR-0325273), by the National Science Foundation under Grant CCF-0438929, by the Sloan Software Industry Center at Carnegie Mellon, and by the High Dependability Computing Program from NASA Ames cooperative agreement NCC-2-1298. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**
IEEE COMPUTER SOCIETY

7. References

[1] Blackwell, A. First Steps in Programming: A Rationale for Attention Investment Models. In *Proceedings of the IEEE 2002 Symposia on Human Centric Computing Languages and Environments*, 2002, 2-10.

[2] Boehm, B., et al. Cost Models for Future Software Life Cycle Processes: COCOMO 2.0. *Annals of Software Engineering Special Volume on Software Process and Product Measurement*, J.C. Baltzer AG Science Publishers, Amsterdam, The Netherlands, 1995.

[3] Boehm, B., et al. *Software Cost Estimation with COCOMO II*. Prentice-Hall, 2000.

[4] Bureau of Labor Statistics. Current Population Survey (CPS) Computer Ownership/Internet Usage Supplement. Reports for all years may be retrieved from…

www.census.gov/population/socdemo/computer/

while raw data for 1994 and later may be downloaded through the DataFerrett program available at…

dataferrett.census.gov/TheDataWeb/index.html

[5] M. Erwig. Update your Program? Program Your Update! Technical Report TR02-60-01, Department of Computer Science, Oregon State University, 2002.

[6] Fisher II, M., and Rothermel, G. The EUSES Spreadsheet Corpus: A Shared Resource for Supporting Experimentation with Spreadsheet Dependability Mechanisms. Technical Report 04-12-03, University of Nebraska--Lincoln, Lincoln, NE, Dec. 2004.

www.cse.unl.edu/~mfisher/corpus/

[7] Forrester Research, Inc. The Market for Accessible Technology—The Wide Range of Abilities and Its Impact on Computer Use. Survey commissioned by Microsoft Corporation, 2003.

www.microsoft.com/enable/research/phase1.aspx

[8] Hall, J. A Risk and Control-Oriented Study of the Practices of Spreadsheet Application Developers. In *Proceedings of the 29th Hawaii International Conference on System Sciences*, 1996, 364-373.

[9] Hecker, D. Occupational Employment Projections to 2012. *Monthly Labor Review*, *127*, 2, 2004, 80-105.

[10] Hipple, S., and Kosanovich, K. Computer and Internet Use at Work in 2001. *Monthly Labor Review*, *126*, 2, Feb. 2003, 26-35.

[11] Jones, S., Blackwell, A., and Burnett, M. A User-Centred Approach to Functions in Excel. In *Proceedings of the 8th ACM SIGPLAN International Conference on Functional Programming*, 2003, 165-176.

[12] Kominski, R., and Newburger, E., Population Division, U.S. Census Bureau. Access Denied: Changes in Computer Ownership and Use: 1984-1997. Presented At the Annual Meeting of the American Sociological Association, Aug. 1999.

[13] Nardi, B, *A Small Matter of Programming*, MIT Press, Cambridge, MA, 1993.

[14] Paternò, F. From Model-based to Natural Development. In *Proceedings HCI International 2003*, 592-596.

[15] Rogers, E. *Diffusion of Innovation*, 4th edition, Simon & Schuster, Inc., New York, NY, 1995.

[16] Ruthruff, J., et al. Debugging and Finding Faults: End User Software Visualizations for Fault Localization. In *Proceedings of the 2003 ACM Symposium on Software Visualization*, 2003, 123-132.

[17] Scaffidi, C., Shaw, M., and Myers, B. An Approach for Categorizing End User Programmers to Guide Software Engineering Research. In the *First Workshop on End User Software Engineering*, 2005.

www.cs.cmu.edu/~cscaffid/papers/eu_20050221_weuse1.pdf

[18] Silvestri, G., and Lukasiewicz, J. Occupational Employment Projections. *Monthly Labor Review*, *114*, 11, Nov. 1991.

www.bls.gov/opub/mlr/1991/11/art5exc.htm

[19] Teng, J., Grover, V., and Güttler, W. Information Technology Innovations: General Diffusion Patterns and Its Relationships to Innovation Characteristics. *Transactions on Engineering Management*, *49*, 1, Feb. 2002, 13-27.

[20] Wallace, C., Cook, C., Summet, J., and Burnett, M. Assertions in End User Software Engineering: A Think-Aloud Study. In *Proceedings of the IEEE 2002 Symposia on Human Centric Computing Languages and Environments (HCC'02)*, IEEE Computer Society, 2002, 63-65.

[21] Wilson, A., Burnett, M., Beckwith, L., Granatir, O., Casburn, L., Cook, C., Durham, M., and Rothermel, G. Harnessing Curiosity To Increase Correctness in End User Programming. In *Proceedings of the Conference on Human Factors in Computing Systems*, 2003, 305-312.

Proceedings of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)
0-7695-2443-5/05 $20.00 © 2005 **IEEE**

IEEE COMPUTER SOCIETY