

Bern, 19. April, 2006

A Critique of the swissUp Ranking 2006 of Computer Science

This document presents an analysis and critique of the swissUp Ranking of Computer Science in Universities and Federal Polytechnics in Switzerland, released on April 6, 2006.

The ranking documents referred to in this report can be found on the www.swissupranking.ch web site. See in particular the pages [uni ranking»Informatik»swissUp Ranking](#), which ranks Computer Science departments in Switzerland, [uni ranking»Informatik»myRanking](#) to see other indicators, [uni/eth»Universität Bern»Informatik](#) to see the average scores achieved, and [info/faq](#) for details concerning methodology, the questionnaires and the numbers of responses received.

Additional information was graciously and promptly provided by swissUp by email when requested, in particular methodological clarifications and the raw data of the answers given by CS students at Bern to the questionnaires.

Summary

In the swissUp Ranking study, 21 out of 22 third semester students stated that they would recommend Computer Science at the University of Bern to beginners starting their studies, and 80% of Computer Science Alumni positively evaluate the impact of their studies on their career.

The swissUp Ranking of Computer Science, however, presents the Institute of Computer Science at the University of Bern (IAM) as obtaining four red scores out of [five criteria selected for comparison](#). Bern clearly comes out worst out of the nine Computer Science departments evaluated, thus leading *Der Bund* (April 7, 2006) to report *Mathematik top, Informatik flop*, and the *Berner Zeitung* *Human- und Zahnmedizin sowie Mathematik top — Informatik flop*.

There are essentially three key problems with the way that the swissUp Ranking has been carried out that have led to this curious result:

- *Selection of swissUp indicators is misleading.* By selecting other criteria for ranking, such as Career Assessment of Graduates, Research Reputation, Study Program, and PhDs per professor¹, Bern comes out as one of the best Computer Science departments in Switzerland.
- *Reduction of raw data to ranked indicators is misleading.* Regardless of how well schools are evaluated, they are classified as “top” (green), “middle” (yellow) or “bottom” (red). If all schools obtain good scores for selected criteria, then schools at the bottom appear to be doing badly, despite their good scores. This creates a false impression for the reader.
- *Questions biased towards large schools and departments.* Many indicators are inherently biased towards large schools and departments, and against tiny Computer Science departments. CS departments at the two Federal Polytechnics are each about 6 times the size of CS at Bern. It can therefore be difficult to compete in areas such as IT Infrastructure and Course Offerings. The ranking is rigged so that the strengths of smaller departments are not in evidence. For example, the fact that Bern, Fribourg and Neuchâtel are launching a joint Graduate School in order to better compete is not reflected in the swissUp Ranking.

Although many of the data collected in the swissUp study are certainly of interest, the way

¹Praxisurteil von Absolventen, Forschungsreputation, Studienorganisation, Lehreputation, Doktorate pro Professor

in which these data are summarized, processed and presented is extremely misleading and potentially highly damaging to the reputation of Computer Science at the University of Bern. At a period when the demand in industry for qualified Computer Scientists and IT specialists is constantly rising, yet inexplicably the numbers of new students starting studies in Computer Science are falling, rankings of this form can only have a catastrophic effect.

In addition to the three main issues listed above, there is a host of further problems with the swissUp Ranking as it has been carried out and presented:

- *Ranking criteria are unclear and flawed.* The criteria for ranking schools into green, yellow and red groups is not transparent. In particular, the way in which the ranking is defined can lead to the paradox that one school may obtain a better score than another school for the same indicator, yet be ranked worse. (e.g., in Biology under the indicator for General opinion (Studiensituation), [Fribourg obtains a slightly better score than Zürich](#) — 2.1 vs. 2.2 — but [Fribourg is ranked worse than Zürich](#) — yellow vs. green!)
- *Ranking presentation is misleading.* The classification of schools into the categories green (go), yellow (proceed with caution) and red (stop!) is particularly unfortunate, as this gives the impression that one should stop certain schools, even though the actual evaluations are generally favourable (e.g., CS at Bern was systematically evaluated “rather good” to “good” by the students, but was painted red). Furthermore, the same score for the same indicator but in different departments can lead to different rankings. (Rooms score 1.9 for both Math and CS, but are ranked green for Math and yellow for CS; IT infrastructure scores 2.4 for both, but is ranked yellow for Math and red for CS.)
- *Statistical sample is unrepresentative.* The statistical basis for the evaluation is extremely low in many cases (e.g., 12% of CS students — 22 out of 187 — were polled in Bern). Furthermore, some key groups, such as final year Masters students and Doctoral students, were not polled. Finally, not all schools in Switzerland were evaluated for all indicators. In the swissUp Ranking of Compute Science, in particular, only 5 of 9 schools were evaluated for all five of the selected indicators.
- *Questionable definitions of indicators.* Definitions of various indicators are highly questionable, or worse, do not represent what they purport to measure (e.g., Teaching Reputation).
- *Selected indicators unsuitable for ranking.* Numerous indicators are unsuitable for ranking since they do not represent an absolute scale (e.g., Research Funds and Attractiveness).
- *Choice of swissUp ranking indicators is arbitrary and misleading.* The [swissUp Ranking page for Computer Science](#) presents only five of 23 indicators without any clear explanation of the choice of indicators, however it is only possible to see the other values by a tedious process of selecting different indicators to display using the [myRanking page](#). One cannot simply page through all possible indicators. Furthermore, the breakdown of the detailed responses students gave to 39 different questions is unfortunately not available on the web site, but only on request.

The net result of these problems is that a completely inaccurate and unrepresentative picture may be painted of the actual situation at a given school. Certainly we acknowledge and recognize that there is room for improvement and appreciate constructive criticism. There are surely many areas where CS at Bern could be improved, given the availability of appropriate resources. However it is most unfortunate that the first impression the ranking gives is at odds with the facts that, for example, 21 of 22 third semester students who responded to the swissUp questionnaire agreed that they would recommend CS at Bern to other beginners ([question 21b](#)), and [nearly 80% of CS Alumni](#) positively evaluate the impact of their studies on their career.

Let us first consider the question of how ranking is carried out, and then we will look at the individual indicators and see how we should interpret the results for CS in Bern.

How the Schools were Ranked

For each selected indicator for a given subject, each school was grouped into green, yellow or red categories. Since the rankings are based on the numerical values of the indicators, and most of these indicators are computed from students' answers to questions, ranging from "very poor" to "very good", one might naïvely suppose that green corresponds to "very good" and red to "very poor". This is far from the truth, which requires some effort to comprehend. Since this is a critical point, it is worth stepping through the details.

Now, if we look at the [numerical scores for CS at Bern](#), we note that for the three indicators based on student evaluations, Guidance, IT Infrastructure and General Opinion, the numerical scores are all 2.4 or 2.5. These scores may not be stellar, but they are certainly positive rather than negative. So why is Bern ranked red in nearly every case? Well, there are two parts to the answer. First of all, the ranking does *not* reflect only a comparison of CS departments within Switzerland, but rather within all of the "German-speaking" region (Switzerland, Germany and Austria), even though we do not see the other schools on this ranking. Second, and more important, is exactly how the classification works:

1. An overall average value is computed for each indicator and subject *across the three countries*.
2. A school is classified yellow if its *confidence interval* (*i.e.*, a representative range of values above and below its average score) overlaps the overall average. Otherwise, if the interval is fully above, or fully below, it is classified respectively green or red.

Note that this scheme implies that two schools may obtain the same average score, yet be classified differently. In unusual situations, a school with a lower score may be ranked more highly (as is the case with the [General Opinion of students in Biology at Fribourg and Zürich](#)).

The key issue, however, is that the traffic signals red/yellow/green do *not* indicate absolute assessments of goodness or badness, but only *relative rankings* within a field within the German-speaking academic world. Should all the schools be considered fantastic for some indicator, or even execrable, there would still be red or green schools.

swissUp Ranking Indicators for Computer Science

The [swissUp Ranking for CS](#) presents five of 23 possible ranking indicators, culled from data that has been collected from (i) questionnaires filled out by the various institutes and faculties in the Fall of 2005, (ii) [questionnaires filled out by Professors](#) at the same institutes, (iii) [questionnaires filled out by selected students](#) taking studies at those institutes.

The five indicators selected for the swissUp Ranking of Computer Science are:

1. Guidance,
2. IT Infrastructure,
3. Research Funding,
4. Teaching Reputation, and
5. General Opinion of the Students.

At first glance this may appear to be a good basis for ranking and comparing Computer Science departments. Let us take a closer look at how these indicators have been defined and computed.

Guidance

The first indicator for which CS at Bern receives a “red traffic light” is Guidance (Betreuung). The numerical value is 2.5, which is halfway between “good” and “rather good”.

Turning to the questionnaire and the raw data we presume that the numerical value is based on the average score of thirteen questions grouped under 12a² and 12b³. A minor source of confusion arises from the fact that the scale used on the questionnaires is reversed from that used in the ranking. For consistency with the swissUp Ranking web site, we will only refer to the reversed scale, where 1 = “very good” and 6 = “very poor”. The average score over all these questions is 2.24 (not 2.5), which is definitely “good” rather than “poor”.

The question with the poorest score (3.33) concerns Guidance with regard to Student Mobility⁴, which is in fact a weakness within the institute. The other twelve questions receive an average score of “rather good” (3) or better. We see immediately that the way in which schools are ranked can make a positive evaluation appear very bad.

IT Infrastructure

Here we must presume that the numerical value is based on question 18 (Bitte beurteilen Sie die studentischen Computerarbeitsplätze). The score obtained here is 2.3 (not 2.4), or once again between “rather good” and “good”. This is certainly not as bad as the red traffic light suggests.

But let us look more seriously at the seven points which contribute to this score. The worst score is 3.35 (between “rather poor” and “rather good”) for Guidance (Beratung). This is a legitimate weakness, since the student Computer pool rooms are in a different physical location than the institute itself. Next worst is Computer Hardware (2.64), which is a traditional sore point in computer labs, though this score can certainly not be considered “poor”! All remaining scores are better, up to a score of 1.42 (good to very good) for availability.

This is all very well, but still there are two important lessons to draw from this exercise:

First, it is clear that the question is strongly biased towards large (and financially stronger) schools. The ranking page clearly demonstrates that the two Federal Polytechnics, ETH and EPFL, shine here. This cannot possibly be a surprise.

Second, we see the danger of amalgamating answers to very different questions, and using them to rank schools relative to each other. We not only lose the overview of what precisely are the advantages and disadvantages at a given place (e.g., Availability vs. Guidance) but criteria that students actually evaluate as being “rather good to good” appear to be ranked as “poor” (red).

Research Funding

This indicator is particularly grotesque for two key reasons. First, it is based on an absurd computation of “funding per researcher”, and second, it glosses over the fact that different kinds of research have very different kinds of budgets. This indicator is completely unsuitable for ranking schools.

This indicator was computed from data provided by the institutes participating in the swissUp evaluation. However interesting it may be to rank schools according to their ability to attract research funding, there are several fundamental problems with this specific indicator and how it

²“Bitte beurteilen Sie die Qualität der Beratung oder Betreuung durch die Dozierenden”

³“Bitte beurteilen Sie ob die Professoren im Durchschnitt ihre Vorlesungen selbst durchführen”

⁴Hilfestellung bei der Vermittlung von Mobilitätsaufenthalten

was computed.

- The numerical value for this indicator does not reflect the overall budget per professor or per research group, but per “researcher”. This is a meaningless basis for ranking since not every researcher is able to apply for funding. It also strangely penalizes institutions with larger numbers of researchers per professor.
- The “number of researchers” has evidently been estimated in different ways by different institutions. IAM interpreted this question as a head count, ending up with a rather high number of researchers (*i.e.*, about 50 people physically carrying out research). With a more conservative estimate that counts the actual number of full research positions paid for by University funds rather than the total number of physical people who carry out research, we arrive at 5 Professors + 2 Assistant Professors + 2×5 Assistants or 17 research positions. This clearly leads to a research funding figure that is 3 times higher than that used for the ranking.
- The amount of research funding received does not take into account fundamental differences between different kinds of research budgets. Some research is relatively inexpensive and the budgets may consist mainly of salaries for PhD students. Other research may require budget for technical infrastructure, salaries for full-time researchers, or possibly for technicians or programmers. As a consequence, research budgets *cannot* be used as a basis for ranking.
- This indicator is partially biased towards large schools which have the facilities to host “big” research with more complex needs and thus bigger budgets. Smaller schools tend to focus on research that entails less administrative and infrastructural overhead.

Teaching Reputation

This indicator (together with Research Reputation — see below) is perhaps one of the strangest used for ranking schools. It is not only severely biased towards larger schools, but the way it has been computed from the responses to questions is fundamentally flawed and misleading.

On the surface, this should be one of the most important criteria to be used for comparing schools. Teaching Reputation, however, is not based on a poll of students but of other professors. This is slightly strange, since Professors are not necessarily in the position to evaluate the teaching reputation of their own colleagues, let alone those of colleagues at other Universities. Since they have no direct exposure or first-hand experience of the teaching abilities of their colleagues, their responses to questions of teaching reputation can only be based on hearsay.

Ignoring that fact, the description of this indicator states, “Percent of Professors who recommend this school for its teaching quality in their area of study (excluding their own school) (compared within Switzerland).”⁵

The actual question asked of Professors, however, was rather different. [The questionnaire](#) asks, “Suppose your daughter or son is planning to start studies in your area of expertise. Which five Universities in German, Austria and Switzerland, excluding your own institution, would you recommend purely on the basis of their teaching reputation?”⁶

⁵Lehrreputation (Prof.) — der maximale Wert ist der beste. Prozentualer Anteil der Professoren, die die Hochschule aufgrund ihrer Lehrqualität für ein Studium in ihrem Fach empfehlen (ausser ihrer eigenen Institution) (interner Vergleich Schweiz)

⁶“Gesetzt den Fall, Sie hätten eine Tochter oder einen Sohn, die oder der Ihr Fach (nach Einstufung in Frage 2) studieren möchte. Welche fünf Hochschulen in Deutschland, Österreich und der Schweiz, ausser Ihrer eigenen Institution, würden Sie ihr/ihm empfehlen, wenn allein die Qualität der Ausbildung ausschlaggebend wäre?”

Several problems now emerge. First of all, those answering the questionnaire are not necessarily qualified to answer the question, so the answers are likely based on hearsay or on other factors that are not necessarily relevant. The question is therefore biased towards large schools. In fact, [the numerical scores](#) seem to reflect nothing other than the size of the corresponding CS departments, from ETH Zürich down to Neuchâtel.

Second, the question most definitely did not ask one to assess each individual school concerned, but only to name the top five. Again this biases the results towards the largest schools.

Third, the entire exercise ignores the fact that each school, and especially small schools, will typically have strengths in specialized domains. The nature of the question is biased against these strengths.

One can only conclude that the numerical value of this indicator cannot be used to establish a ranking that reflects anything except the size of the CS department.

General Opinion

This indicator presumably is based on the answer to [question 21a](#), “How do you generally assess the study situation in your subject at your university?”⁷.

Here CS at Bern obtains the [worst score](#) in Switzerland, a 2.5. A quick look at the raw data reveals, of 23 responses, 4 “rather poor”, 5 “rather good”, 12 “good” and 2 “very good”, for an average of 2.48. In this case, an average of “rather good” to “good” translates to “worst”.

Clearly there are reasons for the poor marks obtained from some students polled, which should be addressed, but we must not forget that the negative marks came from just 4 students in the third semester. Here we see that a very small number of negative responses can change the colour of the ranking.

Question 21b, on the other hand, asks, “Would you recommend studies in your subject at you university to other freshmen?”⁸ Here 21 out of 22 responded “yes”. This paints a rather different picture than the first impression given by the ranking for this indicator.

Let us now consider some of the other indicators, which can be displayed and ranked by means of the “[myRanking](#)” page.

Career planning and Job market

There are two indicators in this group: Career Assessment of Graduates and Job Market.⁹

Career Assessment of Graduates measures the percentage of CS graduates who are of the opinion that their studies have qualified them well to very well for their career¹⁰. Here CS at Bern earns 78.9%, slightly behind Fribourg (82%) and leaving behind both Federal Polytechnics.

The Job Market indicator is computed from the responses to five questions concerning the support offered to students for finding jobs and planning their career (see question group 11 on the [student questionnaire](#)). Here CS at Bern does not do too well, with an overall score of 3.5.¹¹

⁷“Wie beurteilen Sie insgesamt die Studiensituation in Ihrem Studiengang an Ihrer derzeitigen Hochschule?”

⁸“Würden Sie Ihren Studiengang an Ihrer Hochschule anderen Studienanfängern empfehlen?”

⁹Berufsbezug und Arbeitsmarkt

¹⁰“Praxisurteil von Absolventen — der maximale Wert ist der beste. Prozentualer Anteil der Absolventen, die der Meinung sind, dass die im Studium erworbenen Qualifikation gut oder sehr gut den Anforderungen im Berufsleben entsprechen (BFS Daten für 2004, Berechnungen swissUp, Uni: Niveau Fachbereichsgruppe, FH: Niveau Studiengang, interner Vergleich Schweiz)”

¹¹Unfortunately the design of the swissUp ranking web site does not support unique URLs for the various myranking

If we consult the raw data, we see that four questions earn a between “rather good” and “rather poor” and one question concerning support in the transition to a job earns between “rather poor” and “poor” (4.4). In fact, IAM does not have many resources to support such activities, and an improvement here would certainly be welcome.

However, if we look more closely at the raw data, something rather remarkable emerges. The students polled were mostly in the third semester. As such, one might expect that they may not be in a very good position to answer questions concerning the job market and the support they receive. In fact, over all the questions, fully 49% of the responses (53 of 108) were “I have no opinion” (“kann ich nicht beurteilen”). In the case of the question with the exceptionally low score of 4.4, only 24% of the students filling the form had an opinion, *i.e.*, 5 students of the 21 who answered.

Although Bern could certainly contemplate actions to better support the students in establishing a career, the basis for ranking schools here is questionable at best. The negative Job Market indicator is fully at odds with the positive Career Assessment of Graduates.

Infrastructure

The student responses to questions of infrastructure were generally positive, but were ranked negatively. In addition to IT Infrastructure, there are indicators to rank schools with respect to their Audiovisual media (question 13), Lecture rooms (question 14), Student labs and study rooms (question 15), and Libraries (question 19). In the ranking Bern comes out poorly with four red marks and one yellow.

Question 13 concerns audiovisual media (“CD-ROM, Video etc.”). The question is not particularly relevant for CS, as the multimedia needs are mainly supported through computer hardware. Here the students gave the full range of responses, perhaps because of the difficulty in interpreting the intent of the question.

Question 14 consists of six points related to lecture and seminar rooms. The overall score is 1.9 (“good”). Only 9% of the responses were “rather poor” or “poor”.

Question 15 consist of six points concerning student work areas and labs. The average responses to the 6 question range from 2.3 to 3.4, with an overall average of 2.9 (“rather good”). Although this score is the lowest in Switzerland, this indicator was evaluated for only 5 of 9 schools, and the actual range of values is from 1.8 to 2.9. This is clearly another area where bigger schools can be expected to shine, but even ETHZ and EPFL only obtain scores of 2.2. Nevertheless, this is clearly an area where more resources would be welcome to improve the quality of infrastructure available to students.

IT infrastructure (question 18) has been discussed earlier.

Library (question 19) consists of eleven points concerning opening times, literature etc. Here too, we can expect that the larger (*i.e.*, richer) schools will fare better, which in fact they do. The overall score for Bern is 2.6 (“rather good” to “good”). The lowest score is for user help, which is not surprising given the limited resources available. Interestingly, however, fully 42% of the responses expressed “no opinion”. This is consistent with the observation that the main clients of the Library are not the Bachelors students but rather the Masters and PhD students (who were not polled).

Research

There are three indicators for Research. We have already discussed Research Funding above

views.

and noted that this indicator, as it is defined by swissUp, is not suitable for use in ranking schools. The remaining two indicators are the number of doctorates per Professor and the Research Reputation of a school. Both are biased towards large schools, and neither makes sense as a ranking criterion. Furthermore, no graduate students actually involved in research were polled.

With regard to 0.9 doctorates per year per Professor, Bern comes out quite well in the green (for once). However we must ask ourselves, is this a sensible criterion for ranking schools? Is more always better? Would five doctorates a year (*i.e.*, 20 PhD students per professor) represent stellar performance? On the other hand, is someone who focuses his research activity and only produces 0.5 PhDs per annum (*i.e.*, 2 PhD students supervised over 4 years) necessarily a bad Professor? The numbers may be interesting, but they are not necessarily meaningful for *ranking* schools.

The final indicator is Research Reputation. This indicator was, just like Teaching Reputation, discussed earlier, computed from the responses to questionnaires filled out by Professors, and suffers from precisely the same shortcomings. This indicator is described as: “the percentage of Professors who recommend this school on the basis of its research quality, independent of research specialty”¹²

The [actual question posed](#), on the other hand, was: “Which five universities in Germany, Austria and Switzerland, excluding your own institution, are, in your opinion, overall research leaders in your domain?”¹³

Here too, the question posed does not correspond to the indicator as described. The question is focused on identifying the leaders, *i.e.*, the big schools, not on evaluating each individual school. The fact that the two Federal Polytechnics come out on top comes as no surprise. Individual research areas are not taken into account, so the strength of small schools are discounted. A serious evaluation would take into the account the opinion of experts in the research domains covered by the individual schools. The indicator represents nothing but a general impression. Although Bern comes in third, and in fact does have a good research reputation internationally, the basis for the ranking is highly questionable.

International Orientation and Mobility

The three criteria considered here (Percentage of Foreign Students, Percentage of Students from other Cantons, and Attractiveness) are certainly interesting, but it is questionable whether they constitute a good basis for ranking schools. Other influencing factors such as geographical location can play a deciding role and therefore renders the indicator as inherently questionable and flawed. Surely Basel, Geneva, Zürich, Bern and Lugano attract very different percentages of foreign students and students from other Cantons.

Clearly Geneva and Lugano are more attractive to foreign students than to students from other cantons. Obviously the reverse is true for Bern.

A minor point is that it is far from clear how the indicator “Attractiveness” is computed. It is merely described as: “Capacity of the school to attract students from outside the canton”.¹⁴

¹²“Prozentualer Anteil der Professoren, die die Hochschule aufgrund ihrer Forschungsqualität empfehlen (ohne Berücksichtigung der eigenen Einrichtung) (interner Vergleich Schweiz)”

¹³Welche fünf Hochschulen in Deutschland, Österreich und der Schweiz sind Ihrer Einschätzung nach in Ihrem Fach (lt. Frage 2) insgesamt gesehen führend in der Forschung, ausser Ihrer eigenen Institution?”

¹⁴“Kapazität der Universität, ausserkantonale Studierende anzuziehen (Gewichtung: 40% Studierenden aus dem Ausland und 60% Studierenden aus anderen Kantonen) (BFS Daten für 2004, Berechnungen swissUp, Uni: Niveau Studienrichtung, FH: Niveau Studiengang, interner Vergleich Schweiz)”

Studies and Teaching

The remaining eight indicators concern studies and teaching. We have already discussed Guidance (Betreuung) earlier.

E-Learning is covered by question 17 (three points). Bern obtains a 2.14 overall for on-line teaching material, downloadable material and on-line interaction, which can be considered “good”, though it is ranked red. Scores of other schools are rather similar (ranging from 1.8 to 2.2), so this does not appear to be a very critical indicator for ranking.

Practical Relevance (Praxisbezug) was based on six points under question 7. The average score at Bern for these questions was 3.4 (given as 3.5 on the web site), which is between “rather poor” and “rather good”. However a closer look reveals that 31% of the responses were “no opinion”. This is natural given that many of the questions are not particularly relevant in the first three semesters when the focus is more on establishing the foundations of computer science than on practical experience, which essentially starts in the fourth semester. The worst score obtained was for Course Offerings by Practitioners, which is not at all a theme in Bern. This pulls down the overall score considerably. As mentioned earlier, the evaluation by graduates of the practical relevance of the study program paints a rather different picture.

Course Offering (Lehrangebot) is based on seven points of question 4. The overall average for Bern is 3.15. This appears to be an area where there is room for improvement, though not all criteria are necessarily relevant.

1. Size of study groups (2.3) is “good”
2. Diversity of the curriculum (2.87) is “rather good”
3. Possibility to participate early in research (2.95) is “rather good”. In fact, this is already possible during one’s bachelor’s thesis (sixth semester)
4. International orientation of the course offering (3.23) is still “rather good” — though it is hardly clear what is intended by this question
5. Research relevance of course offerings (3.27) is barely “rather good”. Here one must be aware that Bachelors courses focus on fundamentals, whereas only Masters courses cover advanced topic close to research. (Graduate students were not polled.)
6. Didactic quality of the courses (3.58) is “rather poor”. Strangely the university’s internal evaluations in recent years have consistently rated most CS courses quite highly. Here is clearly a need to examine more closely the source of this discrepancy.
7. Interdisciplinary relevance of the course offering (3.83) is “rather poor”. This is a case in point where CS could improve its attractiveness to a broader range of students. The current course offering is very focussed on hard-core computer science rather than on the interplay between CS and other fields, as is the case in most university CS departments.

This particular indicator is a good example where the raw data are essential to make any sense of the numbers used to establish a ranking.

Study Program (Studienorganisation) is computed from eight points of question 5, assessing the formal aspects of the study regulations. Bern is the highest ranked of the Swiss schools evaluated, though all scores are very similar, ranging from 2.1 to 2.3.

Student/Teacher Contact (Kontakt Studierende-Lehrende¹⁵) corresponds to point four of question 9. Bern has an average score of 2.67 (“rather good”), similar to the scores of the other

¹⁵“Urteil der Studierenden zum Verhältnis von Studierenden zu Lehrenden”

schools, with only Lugano and Fribourg obtaining significantly better scores. One should note that only three students out of 22 gave a score less than “rather good”, so the red mark Bern obtains does not in fact seem to signal a serious problem with student-teacher relationships at Bern.

Internal Evaluation of Teaching (Interne Evaluation der Lehre) assesses the students' opinion of how well the University internally evaluates the quality of teaching¹⁶. This is based on the response to three points of question 6. Bern obtains here a score of 2.9 (“rather good”), which is of course ranked red. In fact most schools received similar scores, with only Lugano doing significantly better (2.4). Looking at the raw scores, one may note that Bern did particularly poorly (3.57 — “rather poor”) in the third point concerning realisation of measures arising from the evaluations¹⁷. Clearly there is a need to dig deeper to understand this result (10 out of 21 students polled gave a negative score to this question).

Contact between Students (Kontakt zwischen Studierenden) is covered by the first three points of question 9. Bern earns a 2.5 (between “rather good” and “good”). Interestingly, as one heads further south, the figures improve, with Lugano scoring 1.7.

Concluding remarks

There are numerous lessons that one may draw from the swissUp Ranking 2006 of Computer Science. First and foremost is that the “traffic light” scheme of ranking schools is highly misleading since it can turn “good” into “bad” and thus severely damage the reputation of an institution that is actually doing quite well.

Second, we may note that many of the selected indicators offer totally inappropriate criteria for ranking schools. Whether a given criterion, such as number of doctorates per Professor per year, is better to be high or low is entirely context-dependent. Painting these as green or red does a disservice to the community. Many definitions of other indicators raise serious doubt as to whether they truly measure what they are supposed to measure.

Third, we can see in many of the indicators a serious bias towards large schools (in Switzerland, the Federally-funded Polytechnics). Many indicators therefore monotonously repeat the answer to the question “big or small?” This systematic bias ignores fundamental differences in orientation between the Federal Polytechnics and the more interdisciplinary Cantonal Universities. Not are the smaller schools hard put to compete with the larger ones in the areas used for the ranking, but criteria which would might show them in a better light are not considered. Small schools, for example, tend to focus on certain research specialties rather than attempting to cover the entire field of a given discipline, but these local strengths cannot be seen from the ranking. Far more valuable than a ranking would be a survey indicating the individual strengths and weaknesses of each school in different areas.

Fourth is the observation that ranking obscures the interesting data. The raw data that were used to generate the ranking are only available upon request. Considerable effort is needed to tease some insights from these data, but there is no inherent reason why a straightforward processing of these raw data could not be publicized instead of the artificially construed rankings. Although Computer Science at Bern scored miserably in the rankings, it was only through a detailed examination of the raw data that one could learn that Bern was actually evaluated quite positively in almost all areas. Furthermore, the more detailed numbers not available on the web site exposed the specific strengths and weaknesses. (Student guidance, support for career planning, support for student mobility, and more interdisciplinary course offerings are perhaps some

¹⁶“Urteil der Studierenden bezüglich Evaluation des Studiums und der Lehre durch die Universität.”

¹⁷“Umsetzung von Ergebnissen von Evaluationen der Lehre”

of the areas where more resources could be invested.)

In conclusion, ranking seems not to be very useful as it can paint a distorted image of reality. (For further evidence, see as well the articles [Rankings — mit Vorsicht zu genießen](#) and [“Hochschul-Rankings dürfen nicht überinterpretiert werden”](#) in the NZZ Online, April 11, 2006.) As such, rankings can do far more damage than good. We must now consider seriously what kind of evaluations would truly serve not only the purpose of informing students of the strengths and weaknesses of schools they are considering as places of study, but also the purpose of informing those schools where they should consider investing resources in improving their offerings.

Prof. Dr. O. Nierstrasz
Director, Institute of Computer Science
and Applied Mathematics,
University of Bern