

Zpráva ze služební cesty

| | |
|---|---|
| Jméno a příjmení účastníka cesty | PhDr. Bedřich Vychodil |
| Pracoviště – dle organizační struktury | Odbor digitálních fondů |
| Pracoviště – zařazení | Vedoucí odd. Standardů |
| Důvod cesty | GSoC |
| Místo - město | Mountain View |
| Místo – země | USA |
| Datum (od – do) | 17. -22. října 2012 |
| Podrobný časový harmonogram | 17. 10. přílet, 19. 10. uvítání a zahájení, 20. 3. první den summitu, 21. 10. druhý den summitu |
| Spolucestující z NK | - |
| Finanční zajištění | Cestovné hradí Google, ostatní náklady VaV 0133 |
| Vztah k projektu | Formáty digitalizace, technická metadata pro dlouhodobé uložení, popř. bezproblémové zpřístupnění |
| Cíle cesty | Výměna zkušeností v oblasti vedení projektů v rámci GSoC. Prezentace projektu DIFFER pro odbornou veřejnost. Propagace aktivit NKP a projektu DIFFER. |
| Plnění cílů cesty | Diskuze řešení konkrétních problémů grafických formátů. Způsoby jak získat studenty programátory pro rok 2013 |
| Další podrobnější informace | DIFFER je určen k identifikaci, charakterizaci a validaci grafických formátů |
| Podpora publicity projektu | V rámci diskuze a prezentace se zainteresovanými účastníky v rámci projektu GSoC |

| Související materiály | |
|---|---|
| Materiál | Místo uložení |
| GSoF home page | http://www.google-melange.com/gci/homepage/google/gci2012 |
| Webová stránky vyvíjené aplikace | http://differ.nkp.cz/ |
| Zdrojové kódy vyvíjené aplikace | https://github.com/xrosecky/differ |

| | |
|------------------------------------|--------------------|
| Datum předložení zprávy | 26. listopadu 2012 |
| Podpis předkladatele zprávy | Bedřich Vychodil |

Datum **Podpis**

Podpis nadřízeného

Vloženo na intranet

Přijato v mezinárodním oddělení

Summit Summer of Code 2012

Summit Summer of Code 2012 (dále jen GSoC) je program s celosvětovým záběrem, zapojující studenty (starších 18 let, po maturitě/post-secondary students) vývojáře, kteří se specializují na vytváření open source kódů.

Celý program GSoC odstartoval v roce 2005 v rámci kterého vznikaly open source a free software. Každoroční projekt je vždy omezen na tři pracovní měsíce a rozdělen do dvou (1,5 měsíců) semestrů, přičemž každý projekt je vždy dedikován určité společnosti nebo organizaci (Mentor Organization), kterou zastupuje zástupce organizace (*Organization Admin*).

Myšlenkou celého projektu je výchova nových developerů a nabídnutí férových a zajímavých zkušeností v konkurenčním prostředí. Samozřejmostí je to, že společnost Google dále takto vzniklé kódy nadále sleduje a využívá pro své účely.

Do programu se může přihlásit prakticky jakákoli organizace pracující na nějaké formě smysluplného a zajímavého open source kódu a splňující dané požadavky programu, viz <http://code.google.com/soc/>. Pozor, požadavky se mění každým rokem. Společnost Google na základě interních směrnic a preferencí vybere určité projekty, které se stanou součástí programu GSoC, více viz <http://www.google-melange.com/gci/homepage/google/gci2012>

Společnost Google za celou historii projektu akceptovala jen malou část přihlášených projektů (společností a organizací). Nad rámec toho někteří neuspěli nebo nesplnily podmínky programu. I přesto tímto programem úspěšně prošlo ~6000 *studentů* a ~3000 *mentorů* z více jak 100 zemí z celého světa.

Vybrané projekty jsou zveřejněny na <http://www.google-melange.com/gsoc/projects/list/> a studenti z celého světa se mohou přihlásit k participování na daných projektech. Výběr studentů je vždy na mentorech (Mentor) daného projektu společně se zástupcem organizace (Organization Admin).

Velkou motivací pro studenty developery je finanční ohodnocení (2x2500 USD), která se vyplácí semestrálně po úspěšném dokončení předložené práce odpovědným. Mentori dostávají jednorázovou odměnu za dobře provedenou práci (500USD). Zástupce organizace Organization Admin nedostává žádnou odměnu.

Projekt DIFFER (Determinator of Image File Format propERTies) vznikající v rámci Koncepce rozvoje Národní knihovny České republiky jako výzkumné organizace od r. 2010 do r. 2015, Priorita 2: Kontrola kvality digitalizovaných obrazů (2011-2012), byl úspěšně akceptován do projektu GSoC 2012.

Po úspěšném, tříměsíčním vývoji a implementaci vytyčených částí aplikace v rámci GSoC. Implementace metrik pro porovnávání (vícevláknové) do existujícího frameworku (*Multithreaded implementation of new similarity metrics into existing framework*) a Migrace existující aplikace z Smarty (PHP) do GWT a Vaadin (*Migrating the existing web application from Smarty (PHP) to GWT or Vaadin*)¹ byly zainteresované osoby pozváni na každoroční Summit GSoF 2012 konající se v Googleplexu, Mountain View.

Celý Summit byl nejen setkáním s profesionály z celého světa, ale i díky Googlu se jednalo o velmi příjemnou a neformální akci. Celý summit je koncipován jako prostředí vhodné pro seznámení odborníků z oblasti open source a umožňuje kreativní výměnu názorů.

Aplikace DIFFER byla představena v krátké prezentaci viz *obr. 1* a byly nastíněny další postupy v rámci tohoto vývoje. Byly přivezeny a předány propagační předměty. Tímto děkuji paní L. Abessi z ÚK/1/, viz *obr. 3 a 4*.

¹ Github, Josh Mabrey: <https://github.com/jmabrey>

Organizační zajištění v rámci GSoC - Projektový tým

Projektový manažer: PhDr. Bedřich Vychodil (*Organization Admin*)

I. Programátor a mentor GSOc studenta: Mgr. Václav Rosecký (*Mentor*)

Student programátor: Dushyant Goyal (Jaipur, Indie) (*Student*)

- Implementace metrik pro porovnávání (vícevláknové) do existujícího frameworku (*Multithreaded implementation of new similarity metrics into existing framework*)²

II. Programátor a mentor GSOc studenta: Mgr. Jiří Kremser (*Mentor*)

Student programátor: Josh Mabrey (Louisville, Kentucky, USA) (*Student*)

- Migrace existující aplikace ze Smarty (PHP) do GWT a Vaadin (*Migrating the existing web application from Smarty (PHP) to GWT or Vaadin*)³

Tester: Ing. Jana Boučková⁴ a PhDr. Bedřich Vychodil

- Testování funkčnosti aplikace, grafického rozhraní GUI, reportování chyb a požadavků v rámci programu Redmine⁵

Odborný konzultant: Mgr. Petr Štolcpart⁶

- Podpora v oblasti vedení projektu a tvorby projektové dokumentace. Webové rozhraní a grafický návrh, reportování chyb a požadavků v rámci programu Redmine.

Obrázek 1: Prezentace vyvíjené aplikace DIFFER, přednášející Bedřich Vychodil



Zdroj: Bedřich Vychodil, 2012

²OpenCV Implementation of Image Similarity Metrics, Dushyant Goyal <https://github.com/dushyantgoyal/differ/wiki/OpenCV-Implementation-of-Image-Similarity-Metrics>

³ Github, Josh Mabrey: <https://github.com/jmabrey>

⁴ Ing. Boučková se od května 2012 se na projektu nepodílí

⁵ Redmine je open source program využívaný v rámci řízení projektu, tzv. bug tracking system.

⁶ Mrg. Petr Štolcpart se od roku 2012 se na projektu aktivně nepodílí

Harmonogram summitu GSoC 2012

Friday, October 19, 2012

- 6:30pm - 10:30pm Shuttles available to transport attendees between hotels
- 6:45pm Thai Dinner at Wild Palms Hotel

Saturday, October 20, 2012

- 8:30am-9:30am Shuttles transport attendees from Wild Palms and Domain Hotel to Googleplex
- 8:30am Breakfast at Googleplex
- 10am Summit starts
- 11:00am-5pm Sessions
- 12:30pm-1:30pm Lunch
- 5:30pm-7pm Shuttles transport attendees from Googleplex to Domain Hotel
- 6:45pm Dinner at Wild Palms Hotel
- 6:30pm - 12:30am Shuttles available to transport attendees between hotels

Sunday, October 21, 2012

- 8:30am-9:30am Shuttles transport attendees from Wild Palms and Domain Hotel to Googleplex
- 8:30am Breakfast at Googleplex
- 10am Sessions starts
- 12:30pm-1:30pm Lunch
- 5pm Summit Ends
- 6pm-7pm Shuttles transport attendees to Wild Palms and Domain Hotel
- 6pm One shuttle departs for SFO and one shuttle departs for SJC

Detailní rozpis prezentovaných témat

Tabulka 1: Rozpis na sobotu 20. 10. 2012

| W | Ouagadougou (43-2) (20) | Timbuktu (43-2) (10) | Tripoli (43-2) (16) | Tunis (43-2) (100) | Abidjan (43-1) (10) | Accra (43-1) (16) | Banfora (43-1) (8) | Fes (43-1) (14) |
|-------------|---|---|--|---|--|--|--|---|
| 10:00-10:45 | Open Source Projects and the cloud (Gary B) | | Preparing and encouraging more women to apply for GSoC (Mamina Z) | Open Games Part II | FOSS in Core Engineering. Though FOSS is expanding, core engineering disciplines find it difficult to do all jobs with FOSS (H.S. Rai) | FOSS Operating System Summit (Dr. Joel) (3 hr block) | | Umbrella Org Meetup: discuss strategies and tools successful and fairly manage GSoC for those orgs. (Hamish B) |
| 10:45-11:30 | Building GSoC within your organization (Dustin) | Project Hosting * What works? * What would you like to have? * Share your experience. (Lance Aberison) | Techniques to prevent traffic loss when systems go awry (Adriano Marques) | Give us some feedback on GSoC for the last 8 years & let's talk about the next 8 (Carol S) | Design for hackability: making tools to broaden the benefits of open source (and others). (Forrest Oiphrant) | FOSS Operating System Summit: Minority open source operating systems: sharing knowledge on hardware and drivers across projects. | Advances in biological network visualization. (Thomas Handorf) | In some countries GSoC overlaps entirely with local systems. Is this a problem? And ideas to accommodate this. (Dieter Adriaenssens) |
| 11:30-12:30 | All about free software licenses (Deb) | | Does the unix "many minimal tools working together" philosophy still matter? (Micah Cowan) | GSoC Lightning Talks: Showcase an awesome student project in less than 60 seconds (Sean) | Software/Hardware testing (Amar) | FOSS Operating System Summit: Minority open source operating systems and porting beyond Linux/Windows/OS X: how to make it better for app developers | Bio Geeks: * Project overlays? * New Collaborations? * Brainstorming for next years GSoC projects (B. Good) | Documenting software tools and practices. (Stefan Seefeld) |
| 12:30-1:30 | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch |
| 1:30-2:15 | Boot me, baby! Network-booting as many obscure OSes as we can find. * (linux is too easy) (Michael Brown, iPXE) (1.5 hrs) | Machine Learning and AI | Postgres | Free software non-profits and umbrella organizations demystified. (Bradley Kuhn & Rob Kaye) | JavaScript Sci. Comp. Library? (SOOCR) | Community Metrics: Can we create a common platform to measure contributions & engagement (Michael Downey) | PowerPoint Karaoke | Open Source in Education: Do you use it? How do you use it? What new tools will you need? (Virgilio) |
| 2:15-3:00 | cont. | FOSS machine translation/language processing meetup (Jac Obidofalk) | Continuous integration: Are you using it? No? We'll tell you what it can give your project! (Pierre) | Ideas and techniques to improve student selection phase. (Adriano Marques) | Defensive Publications and prior art: How can they help the community? (dreasillas) | CACERT Part II | FLOSS Audio Project Meetup | |
| 3:00-3:30 | Break | Break | Break | Break | Break | Break | Break | Break |
| 3:30-4:15 | Are students working full time? How do we measure their effort against our expectations that GSoC is "full time"? (globalist) | PHP | Government and open source | What other programs should the Google Open Source Programs office have to support FLOSS? (Cat Allman) | | Moved to Benghazi: Wiki, Etherpad, Code review - how to use them for GSoC/GCI and your project in general (Thomas Waldmann) | Bug report management (Issue tracking): Problems and common practices | Image and Graphics Meetup * Can we inter-operate between? |
| 4:15-5:00 | Mobile open source Android, iOS, and beyond. | FOSS 4 Geo Meetup - All things Maps, Geospatial Projects, etc. (Hamish B) | Beyond-Pass/Fail: canceled - GSoC will not change this - so no point in discussing it. | | | Network Time Foundation/protocol(hatman) | (Mostly) open firmware for the 21st century. Tiano core UEFI discussion/Q/A | FLOSS funding for mums and dads? Suddenly you're a parent. You need some funding. How do you achieve this? Let's collect ideas. (bzg) |

Zdroj: GSoC 2012, http://gsoc-wiki.osuosl.org/index.php/Main_Page

Tabulka 2: Rozpis na neděli 21. 10. 2012

| | Ouagadougou (43-2) (20) | Timbuktu (43-2) (10) | Tipoli (43-2) (16) | Tunis (43-2) (100) | Abidjan (43-1) (10) | Accra (43-1) (16) | Banfora (43-1) (8) | Fes (43-1) (14) |
|-------------|---|---|--|--|---|---|---|---|
| 11:00-11:45 | Implementing advanced scientific algorithms - challenges of recruiting & mentoring students (Bronstengel) | | X | | | | | X |
| 11:45-12:30 | Computational Physics Simulations in JavaScript and WebGL: Finding students who have the necessary skills + interests: science, javascript, WebGL (OpenGL), Visualization | | Engaging Students with the Community during GSoC (D.G. Siegel) | Mentoring lessons learned, and how to apply them to future GSoCs (Micah Cowan) | Memmapping on Android: Fitting 50MB data into 10 MB RAM (Jacob Nordfalk) | Everything you always wanted to know about intellectual property & some stuff you didn't (Deb) | | Building, supporting and funding academic-industry collaborations - ideas, partners, challenges and prospects (Socr) / Code Sprints in the projects (Andreas Fabri) |
| 12:30-1:30 | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch | Lunch |
| 1:30-2:15 | Programming Languages | Open Source Community Events: Mentors from Drupal will discuss the successes of our community based on user groups/events (Sturpee) | Student Application Templates What works and what doesn't? (asmeurer) | Ensuring the continued success of a project after the completion of the GSoC program (Sargu Blumtriu) | 200 contributors and counting: strategies of LibreOffice to grow its community (Cedric Bostdomat) | Google Code-in feedback and ideas session (S Taylor + Carol S) | Global Mentoring: how to mentor effectively across time-zone, language and cultural barriers (Matt) | Recruiting mentors - how to recruit mentors and how to set their expectations (Dustin) |
| 2:15-3:00 | LLVM/Clang: Disruptive change in the distribution world (Silvestre) | Open Source & Government civic engagement / public organizations (Rick Dietz) | Semantic Data: Collaborations between organizations and students to tame the wild wild web (Andrew Webb) | Collaboration tools for your org: What works what doesn't, what's missing? (Michael Downey) | Increasing US participation in GSoC (Eric Schultz) | Brainstorm ideas to successfully run an open source project with participants across time zones (Antis) | | FLOSS 4 Humanity Software for a developing world (Hamish B) |
| 3:00-3:30 | AppSec 101: An introduction to application security - testing & securing your code the osway way! (Kostaak Fabio) | Open Source & Government civic engagement / public organizations (Rick Dietz) | Open Source Gaining Projects Meetup (Kai) | Open Source Gaining Projects Meetup (Kai) | | | Get together of projects dealing with geometric entities (Andreas Fahn) | CA cert, PGP/GPG Key Signing (harian Stenn) |
| 3:30-4:00 | Break | Break | Break | Break | Break | Break | Break | Break |
| 4:00-4:45 | Open Source & Science : making the most of recent cultural shifts (Nick Barnus) | JSON-RPC, Next-gen-RPC (Mo Morsi) | Source Code hardening essential tools and best practices for improving code quality, verification and validation, etc. what works, what doesn't (Sean) | Applying right project management principal in open source projects e.g. project tracking, scheduling, monitoring (rahu) | FLOSS audio project meetup (RJ Ryan) | | Student Plagiarism detecting and dealing with it. Let's develop official guidelines for GSoC (Robbin Haw) | |
| 5:00-5:30 | Group photo | Group photo | Group photo | Group photo | Group photo | Group photo | Group photo | Group photo |

Zdroj: Zdroj: GSoC 2012, http://gsoc-wiki.osuosl.org/index.php/Main_Page

Prezentace

Autor/Přednášející: PhDr. Bedřich Vychodil

Téma: DIFFER (Determinator of Image File Format propERTies)

Tabulka 3: Prezentace DIFFER

DIFFER
Determinator of Image File Format propERTies

Lecture: GSoC 2012
Lecturer: Bedřich Vychodil
Web: <http://differ.nkp.cz>
Contact: bedrich@gmail.com

Bedřich Vychodil, Digital Preservation Standards Department, The National Library of the Czech Republic

DIFFER – Finds Differences

Data processing

Two images are compared. The difference image shows a red box highlighting a region. A zoomed-in view of this region shows a color difference. A text box says "HASH IS NOT EQUAL". A dB scale at the bottom right shows values 26, 14, and 0 dB.

Bedřich Vychodil, Digital Preservation Standards Department, The National Library of the Czech Republic

Digitization, Migration, QA

Difference between layers
DEVIATION: Black - Min White - Max

Two newspaper pages are shown side-by-side. A red box highlights a difference in the text area. A zoomed-in view shows the difference between the two layers. A dB scale at the bottom right shows values 26, 14, and 0 dB.

Bedřich Vychodil, Digital Preservation Standards Department, The National Library of the Czech Republic

DIFFER – Glitches Detection

Data processing

Two images are compared. A red box highlights a glitch in the image. A zoomed-in view shows the glitch. A dB scale at the bottom right shows a value of 250.

Bedřich Vychodil, Digital Preservation Standards Department, The National Library of the Czech Republic

Zdroj: Bedřich Vychodil, 2012

DIFFER - GUI

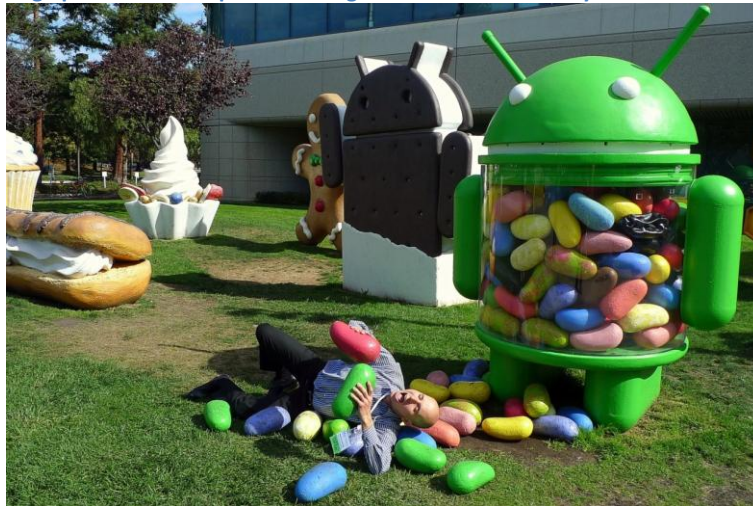
Data processing

Three images are shown side-by-side. Each image has a corresponding graph below it. The graphs show the difference between the images. A dB scale at the bottom right shows values 26, 14, and 0 dB.

Bedřich Vychodil, Digital Preservation Standards Department, The National Library of the Czech Republic

Fotodokumentace

Obrázek 2: Prohlídka Googleplexu a vesměs povinné fotografie u maskotů verzí systémů Android



Zdroj: Bedřich Vychodil, 2012

Obrázek 3: Stůl samolepek a propagačních materiálů (vpravo dole materiály dovezené z NK ČR, díky p. L. Abessi z ÚK/1/)



Zdroj: Bedřich Vychodil, 2012

Obrázek 4: Stůl samolepek a propagačních materiálů (vpravo dole materiály dovezené z NK ČR, díky p. L. Abessi z ÚK/1/)



Zdroj: Bedřich Vychodil, 2012

Obrázek 5: Původní kavárna v rámci Googleplexu, kde se odehrávala velká část GSoC summitu



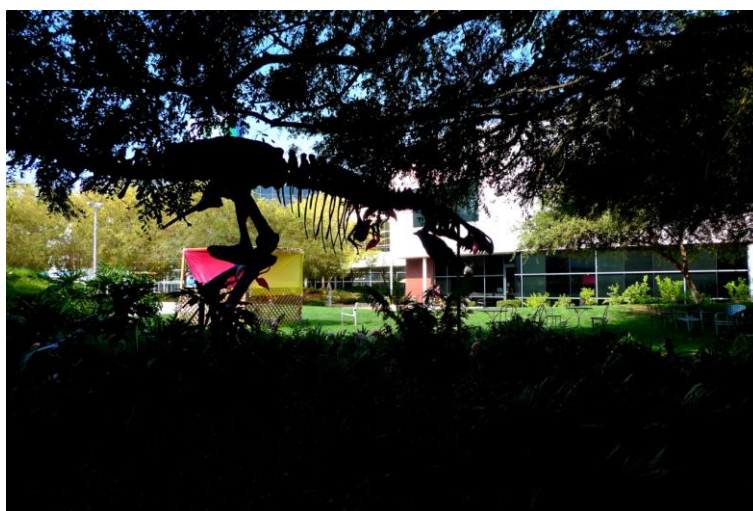
Zdroj: Bedřich Vychodil, 2012

Obrázek 6: Google zahrádka, ze které se plody využívají v rámci stravování Googleplexu. Tzv. sezónní strava.



Zdroj: Bedřich Vychodil, 2012

Obrázek 7: V rámci Googleplexu je možné narazit na velké množství zajímavých kuriozit



Zdroj: Bedřich Vychodil, 2012

Obrázek 8: Googleplex, budova č. 41 do které jsme též neměli přístup



Zdroj: Bedřich Vychodil, 2012

Obrázek 9: V Googleplexu zaměstnanci mohou trávit dokonce i svůj volný čas a také se tak děje.



Zdroj: Bedřich Vychodil, 2012

Obrázek 10: Pro zaměstnance a návštěvy Google často pořádá venkovní BBQ. Kdo nezažije, ten neuvěří!



Zdroj: Bedřich Vychodil, 2012

Obrázek 11: Každoročně pořadatelé GSoC požádali účastníky se o přivezení čokolády. Letos se kladl důraz na kvalitu a ne na kvantitu jako rok minulý. Sešlo se mnoho zajímavých příchutí. Za Českou republiku jsem přivezl dvě Studentské pečete a byl to úspěch!



Zdroj: Bedřich Vychodil, 2012

Obrázek 12: Zaměstnanci a návštěvníci Googleplexu mohou zdarma využívat Google kola pro osobní dopravu



Zdroj: Bedřich Vychodil, 2012

Obrázek 13: Areál Googleplexu z kopce (37.425979,-122.083973) ne nadarmo se této oblasti přezdívá říká Mountain View



Zdroj: Bedřich Vychodil, 2012

