

UNIVERSITY OF VAASA

FACULTY OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

NORRGÅRD FILIP

TTE.177 INTERNET TECHNOLOGIES

EXTRACTING PERSONAL SCHEDULES FROM WOMPAT FOR
CONVERSION INTO vCALENDAR, iCALENDAR AND MICROSOFT
OUTLOOK CALENDAR APPOINTMENTS

VAASA 2004 - 2005

1. Introduction

Once upon a time, planning the studies for the near future was done using paper, eraser and pen. Not only was this a slow, tedious process of planning by hand but also a problematic situation when updates or schedule changes happened. The improvement was Wompat, a web-based timetable planning program.

Today all lecture details are accessible from Wompat with only an option available to let it be printed to a hardcopy. As the lecture information was available in digital form, the next natural step was getting it distributed as calendar appointments to calendar software. This would mean that the students and lecturers can maintain the lecture details on their PDAs, cellphones and PCs and avoid the mess and cost of maintaining pen-and-paper-calendars.

Thus, to fulfill the natural step, in to the picture comes Dingo. Dingo is the intermediate software between Wompat and calendar software. It parses the information from Wompat into something that most calendar software should be able to recognize namely vCalendar and iCalendar files.

2. System Requirements

Dingo was built using the C# programming language which is part of the Microsoft .NET framework, and hence requires the .NET framework version 1.1 to run. The .NET framework itself is available to be installed on: Windows XP, 2000, ME, 98, NT.

In the near future, Dingo might be able to run on open source ports of the .NET framework: Mono framework and DotGnu Portable.Net framework.

Both Mono and DotGnu support GNU, Unix, Linux, Mac OS X and Windows operating systems. As it currently stands, Dingo is not able to run on on these platforms since they have a partial support for Windows.Forms namespace and

might have to be recompiled into their respective implementations of the Microsoft intermediate language.

3. The Architecture of Dingo

In this chapter, we will look at the major components that make up Dingo.

The main components in Dingo can be divided into four parts:

- **Automatic downloader** – downloads the Wompat timetables and handles the saving of the files.
- **Parser** – parses the Wompat timetable HTML file and extracts the course details.
- **Week number to date converter** – calculates the date based on the year, weekday and week number.
- **iCalendar and vCalendar generator** – creates and writes the iCalendar or vCalendar files to the disk.

3.1. *Automatic Downloading of the Timetable*

Dingo can use an Internet connection to download, or a local stored copy of the Wompat timetable HTML file for generating calendar appointments.

If the user uses the built-in timetable downloader to download the timetable (see image 1) from the Internet, then Dingo follows the following procedure:

First it uses the user credentials (provided by the user) to log in via an HTTP POST request – HTTP POST requests are used in all of the communication with Wompat -- to the Wompat log-in page, `autentikointi.php`, and stores the session identification cookie sent by the log-in page. Immediately after getting and storing the cookie, it fetches the list of timetables from the server-generated HTML file generated by `show_lukkarit.php`. Storing that HTML file in memory, it parses that page for the stored years and seasons (e.g. spring 2005, and fall 2005) to allow the user to select which timetable should be downloaded.

After the user has decided which timetable to download and where it should be saved he/she clicks the Download button and Dingo starts downloading the timetable HTML page. Based on the selected year and season, Dingo uses the that season's timetable ID to fetch the next webpage using a request to `suunnittelu_frameset.php`. From `suunnittelu_frameset` page, it proceeds to download the course selections page, `esivalinta.php`, and saves the current selection of faculty courses the user currently has, while at the same time generating a request for all of the faculty courses. Dingo has to save the user's selection for returning the selection later in the downloading phase since the selection is stored on the server, and the user would not like to have all the faculty courses selected the next time he/she returns to Wompat using a normal web browser.

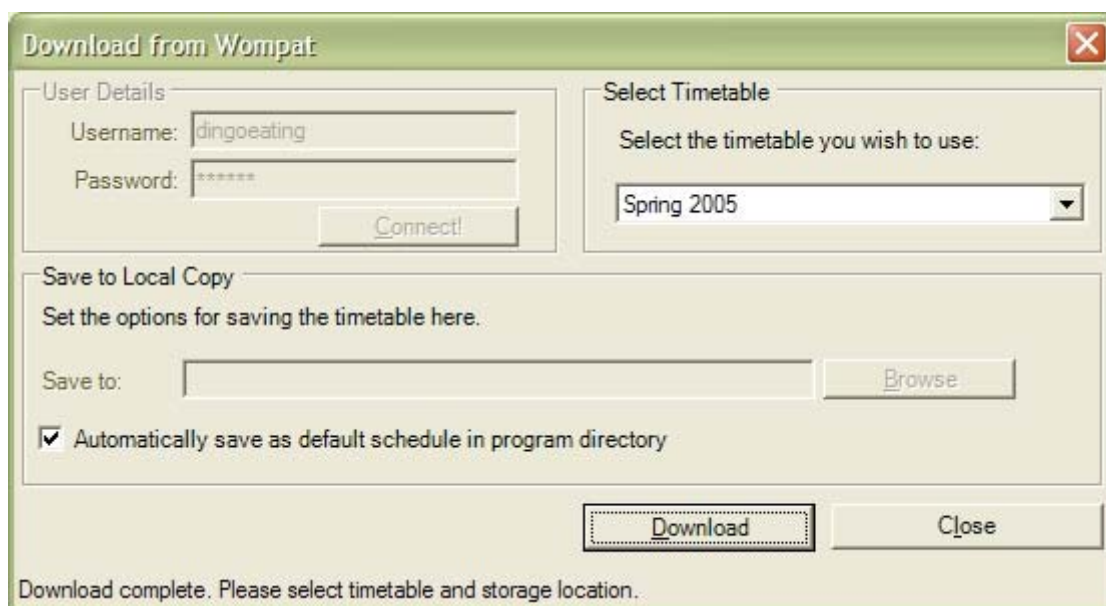


Image 1 – Screenshot of the Wompat schedule downloader

Next, Dingo posts a new HTTP request that contains a selection for all of the faculty courses to `kurssilista_frameset.php`. This request basically tells the server to show all of the courses that all of the faculties have and generate a list of all the available courses. The list of all available courses is retrieved by a new request to `kurssilista.php`. The `kurssilista.php` HTML response contains all of the

course selections that the user has made, in the form of checked checkboxes in the HTML code. Dingo parses that HTML response for the checked checkboxes and based on the checked boxes' attribute *name*, Dingo generates a POST request for all of the courses previously chosen by the user. The request is made to `kurssilista_action.php` that responds with an HTML file to Dingo that contains the timetable file that we want to parse.

The timetable that is retrieved by using the steps above, is the same timetable that is viewed in a normal web browser after 1) logging in to Wompat, 2) selecting a timetable, 3) selecting a faculty course, 4) pressing the "*Näytä kurssit*" button, 5) selecting a course (or more), 6) pressing the "*Päivitä*" button, 7) the generated timetable in the right frame (or through the "*Näytä tulostettava muoto*" link).

3.2. Opening Local Stored Copy

In cases where the user might not have Internet access at the computer where Dingo is installed, the user can choose to open a locally stored copy (i.e. on a USB memory stick, or some other portable media) of the Wompat's schedule HTML file. The user must in such cases manually save the schedule print page ("*Näytä tulostettava muoto*" link) in an HTML file and take that file to the computer with Dingo installed. In Dingo, the file can then be opened by clicking File > Open local copy and Dingo will parse the file for the course information. Thereafter the user can export the required course details into calendar appointments.

As of current version 0.5, Dingo will assume current year for the HTML file and not ask the user which year it is. This will be improved upon in future releases.

3.3. *Parsing the HTML File*

As the Wompat timetable HTML file does unfortunately not use XHTML, the parsing method had to be written from scratch. If the timetable would have used XHTML, then it could have been possible to parse for information using the .NET framework's XML parsing features.

The parser first analyzes the first row of the HTML file to see if there is a comment from Dingo automatic downloader. The Dingo downloader adds the year for which the timetable is valid. So, if someone downloads the timetable for spring 2003, it adds to the top of the file the following HTML comment:

```
<!-- Dingo Year;Spring 2003;-->
```

Information about the year is essential when calculating the dates for the calendar appointments, since most of the course schedules run according to week numbers.

Next, the parser start extracting the course data from the HTML file. The first thing that the parser looks for, is the tag for the first course which is the same image HTML tag in all scenarios:

```
<img src=kuvat/K0.GIF>
```

The parser starts extracting from that HTML tag to the bottom of the file and extracts the following information for each course that it can find:

- Course name and code
- The lecturer's name
- Halls for the lectures
- Lecture's start and end times
- The lecture's start and end week numbers

- Any additional comments or information

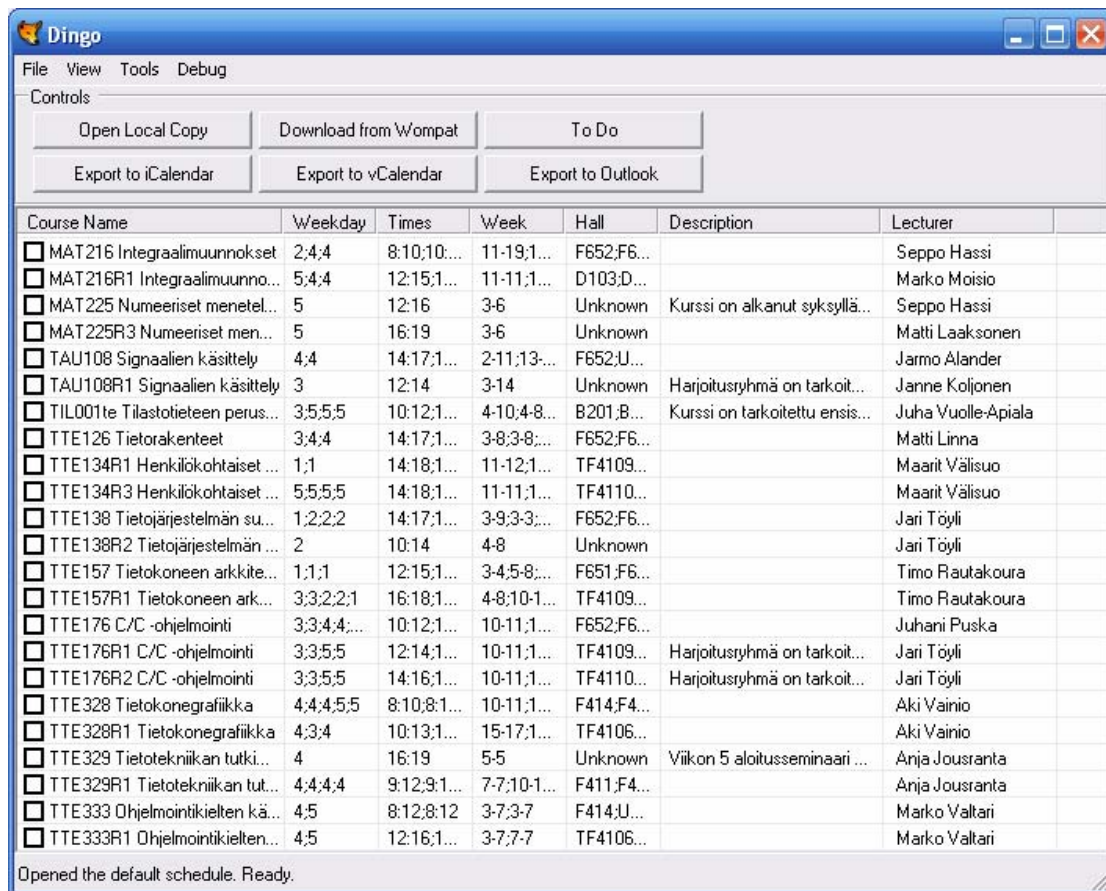


Image 2 – the Dingo main window with sample of extracted course information

All of the extracted information is then added to the ListView control in the main window of Dingo and currently only stored in that control (and in memory).

3.4. Creating the Appointments

When a user has made checkmarked some (or all) of the courses in the main list in Dingo, the user can choose to export the selection to iCalendar or vCalendar files, or to have Dingo insert the data directly into Microsoft Outlook. Due to noted compatibility problems between scheduling applications and their ability to handle vCalendar and iCalendar files, the user has to decide whether to have the appointments exported to a single file or multiple files.

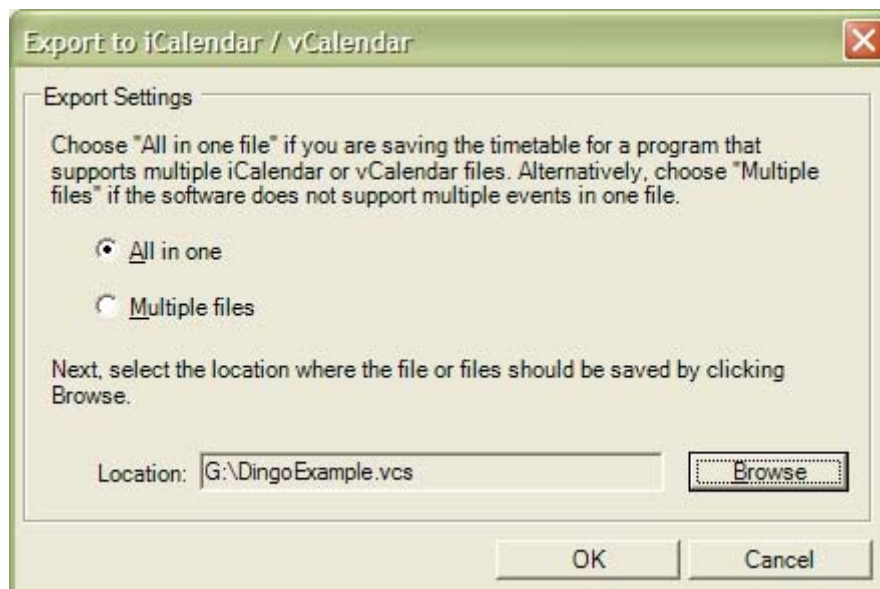


Image 3 – Screenshot of exporting course lectures into a vCalendar file

If the user chooses the option “single file” (see image 3), all selected courses will be exported to one single file. However, upon choosing the “multiple files” option, the user has to select (or create) a folder to which the files should be saved. In that event, a new file will be created for each date that there is a course lecture — the result naturally being many files.

Dingo does not generate vCalendar/iCalendar files that utilize vCalendar’s recurring events property. This is due to the fact that many programs (like Microsoft’s Outlook and Apple’s iCal) have a poor implementation of the vCalendar and iCalendar specification. According to some websites, some calendaring programs might even miss a few of the repeating occurrences in vCalendar files! Microsoft’s Outlook, for example, does not import vCalendar files that have got recurring events.

All of the iCalendar and vCalendar files generated by Dingo use the *categories* property, which is automatically set to the "Dingo". The purpose of this is help the user delete appointments when or if the user decides to update their calendar

appointments using Dingo. The deletion can be done by searching for all appointments with the category name "Dingo" and deleting those appointments.

It is possible to create a "cancelled" vCalendar or iCalendar appointment for deleting/removing previous vCalendar or iCalendar appointments, but that assumes the use of GUID (global unique identifier number) for each appointment. Dingo does not yet implement GUIDs into the vCalendar and iCalendar files and the reason for this is that Dingo does not keep track of the exported appointments. This feature might be used sometime in the future when Dingo might start to support synchronization from Wompat to calendar appointments.

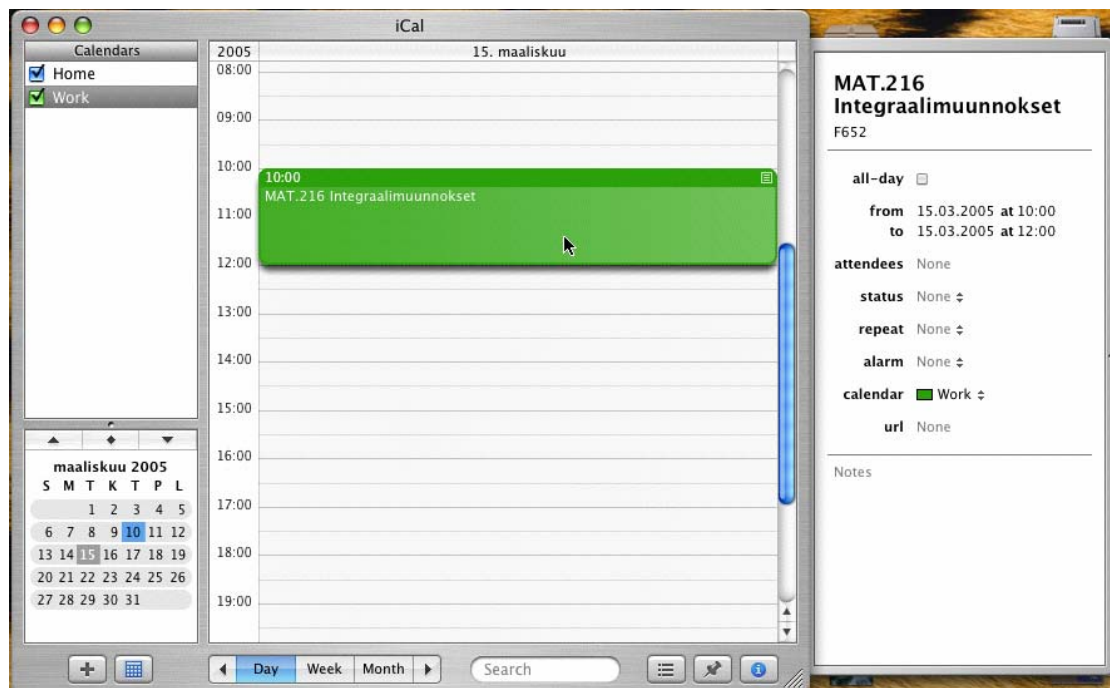


Image 4 – Apple iCal using imported Dingo iCalendar data

During testing it was noted that not all scheduling programs respect the category property of vCalendar and iCalendar files, and that application was Apple's iCal. It still remains up to the user to test and see whether his/her calendaring program supports this property or not and find a workaround in the event that the program does not support it. As for Apple iCal, it is possible to create another calendar

“category” and then, when using File > Import, choose to import the vCalendar/iCalendar files to that new category.

3.4.1. Exporting to Microsoft Outlook

Exporting to Microsoft Outlook was selected as a separate feature since it is possible to create, manage and delete appointments in Outlook via its object model. The object model is accessed by opening a reference to Microsoft Office’s Interop DLL (dynamic linked library) included with all the (latest) Outlook versions. Dingo’s export to Outlook feature was tested and concluded to work with Microsoft Outlook 2002 (XP) and 2003.

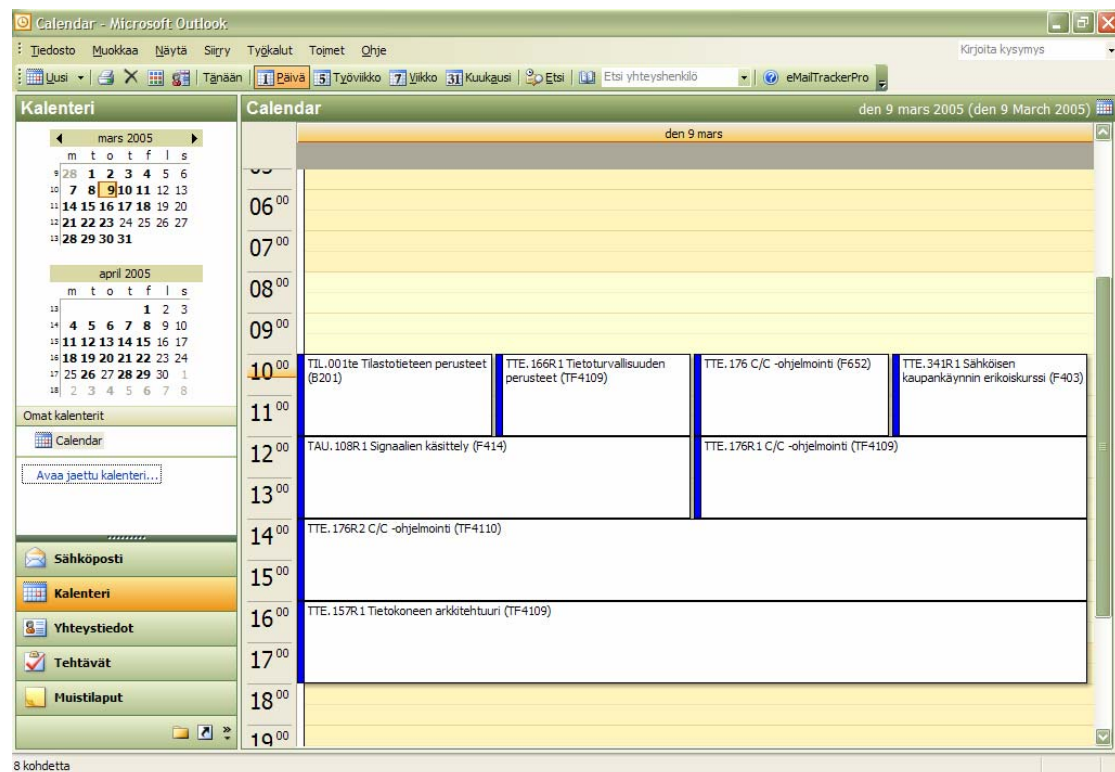


Image 5 - Example of exported calendar appointments in Outlook performed by Dingo

The pros with this feature are that it requires virtually no interaction by the users -- as with the iCalendar and vCalendar files -- and it is possible for Dingo to remove all

the appointments that it inserted into Outlook (although, this is not yet implemented). The cons is that Outlook is one single program that is not free for all to use and limited to one platform — Microsoft Windows.

However, since most data synchronization programs support syncing with Outlook, this was more feasible to be included as a bonus feature. Furthermore, the process for creating appointments in Outlook requires minimal effort from the programmer's perspective, when compared to the process of generating vCalendar and iCalendar files.

3.5. Getting the Date Based On the Week Number

The hardest part to code in Dingo was the week number to date converter. In most European countries and especially in Finland, the week numbers are determined by the ISO 8601 standard. Accordingly, if the year starts or ends on a Thursday, the year will have 53 weeks instead of the "normal" 52 weeks.

The .NET framework does unfortunately not implement the ISO 8601 standard fully, and errors are known to happen with the week number calculation system used in .NET. Hence, this feature had to be written from scratch and the implementation is currently incomplete for most years that are leap years.

The problem with the current way the code is built is that with certain leap years, the dates will be wrongly calculated from the week numbers. Due to the time restrictions, the errors in the calculations were not yet corrected before this release.

Getting a date from the week numbers requires two other pieces of information: the year and the weekday. With these three pieces of information at hand, it is mathematically possible to calculate the date for an event that occurs on, e.g. a Wednesday in week 23 in the year 2007.

4. vCalendar and iCalendar

In this chapter, we will more closely examine the properties of vCalendar and iCalendar files created by Dingo.

4.1. *The Differences Between iCalendar and vCalendar*

The differences between vCalendar and iCalendar files are minor when comparing their contents, with the major difference being the support for these two formats.

vCalendar was a format created by the Internet Mail Consortium and is open file standard for distributing and sharing calendar appointments. The whitepaper for version 1.0 of vCalendar was released in 1997, with version 2.0 of vCalendar, later re-branded as iCalendar, released as an RFC paper in 1998. The support for iCalendar varies greatly from application to application, whereas vCalendar is more widely supported.

Since both formats are closely related to each other, the most noticeable difference between vCalendar and iCalendar is the version number property in the files and the file extension on the files: iCalendar: *.ics* – vCalendar: *.vcs*

A vCalendar file can look as the following example:

```
BEGIN:VCALENDAR
VERSION:1.0
PRODID:-//University of Vaasa//Dingo//FI
BEGIN:VEVENT
SUMMARY:MAT.216 Integraalimuunnokset
DESCRIPTION;ENCODING=QUOTED-PRINTABLE:
DTSTART:20050322T080000
DTEND:20050322T100000
LOCATION:F652
CATEGORIES: Dingo;
END:VEVENT
END:VCALENDAR
```

While an iCalendar file of the same appointment can have the following appearance:

```
BEGIN:VCALENDAR
VERSION:2.0
PRODID:-//University of Vaasa//Dingo//FI
BEGIN:VEVENT
SUMMARY:MAT.216 Integraalimuunnokset
DESCRIPTION;ENCODING=QUOTED-PRINTABLE:
DTSTART:20050322T080000
DTEND:20050322T100000
LOCATION:F652
CATEGORIES:Dingo
END:VEVENT
END:VCALENDAR
```

While the differences in the generated iCalendar and vCalendar files are little (look at the *version* property in the above examples), different calendaring software comply and require different degrees of information to be present in the files. For example, Outlook would not accept importing Dingo's generated iCalendar files while Dingo's vCalendar files were accepted. See also chapter 4.3 for more on the compatibility of the generated iCalendar and vCalendar files.

4.2. Multiple Events and Single Event in vCalendar and iCalendar

vCalendar and iCalendar specifications both have support for multiple event exports, i.e. a single file can contain more than one event.

This is useful for situations where you want to transfer one file only to another computer or device and add new appointments there. However, the support for multiple event vCalendar and iCalendar files varies strongly between calendaring software and as such, Dingo had to be written to generate either single event files, or a multiple event file for all selected courses. To more clearly show the difference

between a single event file and multiple event file, here follows some examples. A single event vCalendar file could look like this:

```
BEGIN:VCALENDAR
PRODID:-//University of Vaasa//Dingo//FI
VERSION:1.0
BEGIN:VEVENT
SUMMARY:FYS.112R2 Fysiikka II
DESCRIPTION;ENCODING=QUOTED-PRINTABLE:
DTSTART:20050119T120000
DTEND:20050119T140000
LOCATION:F411
CATEGORIES:Dingo
END:VEVENT
END:VCALENDAR
```

While a multiple event vCalendar file might look like this:

```
BEGIN:VCALENDAR
PRODID:-//University of Vaasa//Dingo//FI
VERSION:1.0
BEGIN:VEVENT
SUMMARY:FYS.112R2 Fysiikka II
DESCRIPTION;ENCODING=QUOTED-PRINTABLE:
DTSTART:20050119T120000
DTEND:20050119T140000
LOCATION:F411
CATEGORIES:Dingo
END:VEVENT
BEGIN:VEVENT
SUMMARY:FYS.112R2 Fysiikka II
DESCRIPTION;ENCODING=QUOTED-PRINTABLE:
DTSTART:20050202T120000
DTEND:20050202T140000
LOCATION:F411
CATEGORIES:Dingo
END:VEVENT
```

END : VCALENDAR

4.3. Testing Dingo Generated iCalendar and vCalendar Support in Calendaring Software

As the support for vCalendar and iCalendar with one or more events differs from calendar software to calendar software, testing was done to see how well the Dingo generated iCalendar and vCalendar files were supported in various calendaring software.

4.3.1. Multiple Event vCalendar Compatibility

vCalendar files containing single or multiple events from Dingo works with:

- Microsoft Outlook 2002 (XP) - when using File > Import and Export
- Microsoft Outlook 2003 - when using File > Import and Export
- Apple iCal 1.5.5 - when using File > Import
- Mozilla Sunbird version 0.2 - when using File > Import

4.3.2. Single Event vCalendar Compatibility

vCalendar files with only a single event works in:

- Microsoft Windows Mobile Pocket PC 2003 Second Edition

4.3.3. Multiple Event iCalendar Compatibility

iCalendar files with one or more events works in:

- Apple iCal 1.5.5 - when using File > Import
- Mozilla Sunbird version 0.2 - when using File > Import
- Rainlendar 0.21 - when using iCalendar import plug-in

4.3.4. Single Event iCalendar Compatibility

iCalendar files with only a single event works in:

- Nokia Series 60
- Microsoft Windows Mobile Pocket PC 2003 Second Edition

5. Conclusion

As it has been seen in this document and the final product Dingo, the downloading and processing of Wompat schedules into iCalendar, vCalendar and Microsoft Outlook appointment is possible using .NET framework technologies. This also opens up a whole new world of scheduling ease and possible uses for the students and employees at University of Vaasa.

5.1. Information Flow with Dingo

Probably the best part with Dingo is that it enhances information flow from Wompat to other devices such as PDAs and cellphones. This introduces a new level of ease to the users since they can more easily check which courses they have today, and which lectures are overlapping without having to login and read through Wompat and worrying which course starts when.

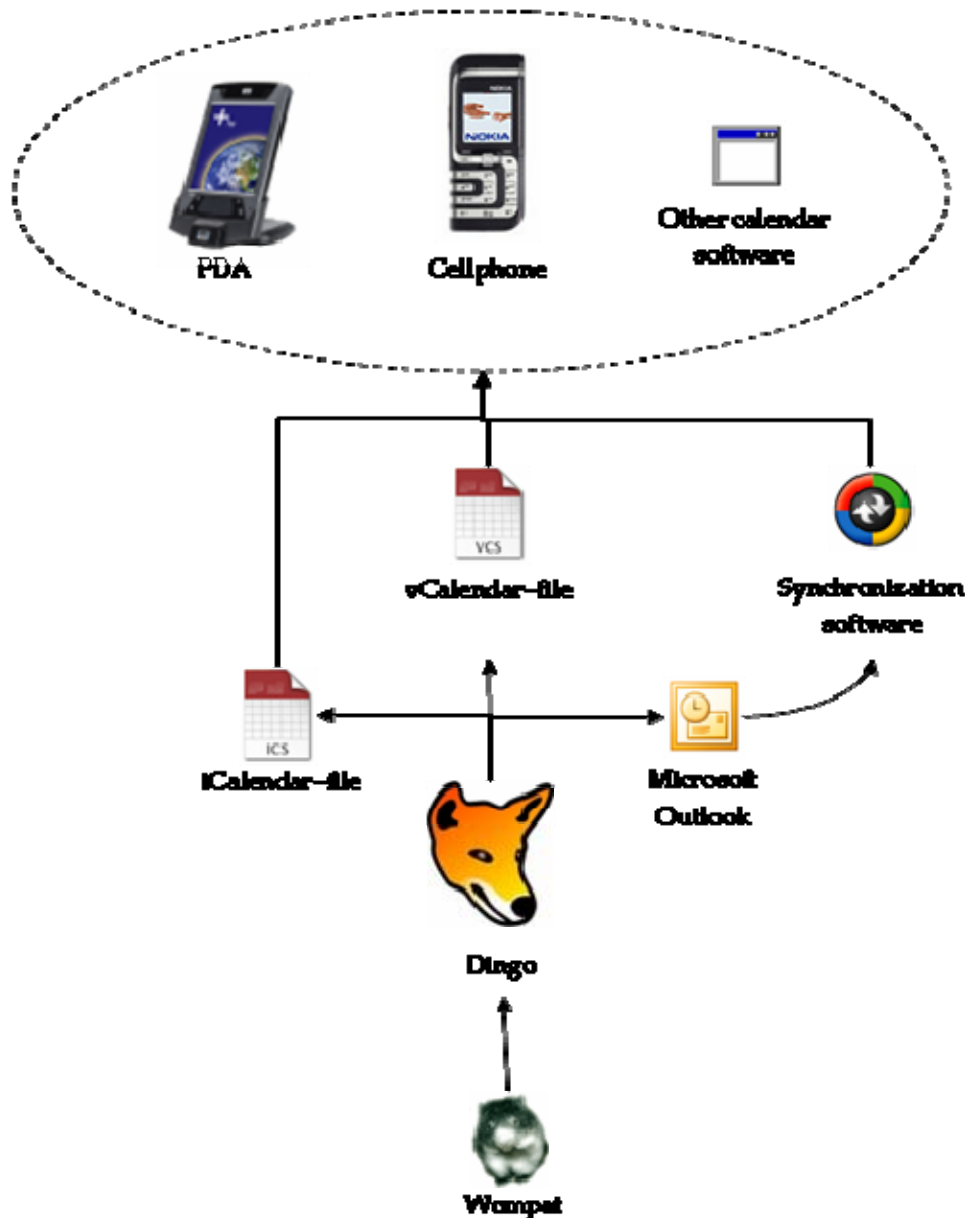


Image 6 – the information flow with Dingo as intermediate software

5.2. For The Future Versions of Dingo

There are some ideas already for the future versions of Dingo:

- Special versions for shell usage (Unix, Linux, Mac OS X), and possibly porting to the open source versions of .NET frameworks
- Synchronizing from Wompat to calendaring software
 - With support for removing previous lecture data using vCalendar and iCalendar

- Updating of calendar appointments according to changes in Wompat
- Automatic connecting to Wompat and other web resources to check for schedule changes (according to user preferences, naturally)
- Information hub feature
 - Automatically checks upcoming exams from Hessu or the WebOodi system and reports them to the user as necessary
 - Automated course enrollment to the courses selected in Dingo by a single click
 - Checking newsgroups for new course related posts

It might be that not all of these features can be or will be implemented, depending on the complexities of making them work properly.

However, using screen scraping techniques to extract information from Wompat is not a good idea in the long run as changing the layout of Wompat would require changes in the parser of Dingo. As a further development, would be to use PHP XML RPC classes to design a web service to allow Dingo to more easily (and more precisely) access the schedules and timetables stores in Wompat. Parts of Dingo can as well be used in Wompat: e.g. the week number to date calculator combined with the vCalendar and iCalendar generator can be ported to the server and made available for the rest of the users to utilize. This can either be a direct port – by installing and running Mono on the server and copying over the code to the server – or by indirectly porting the code to PHP and running it in conjunction with Wompat on the server.

6. Acknowledgements

The HTTP debugging proxy, Fiddler, helped the software developer design and structure the Dingo project. Throughout the project, it was mainly used to analyze the HTTP POST request variables sent, and the received cookies and responses when using Wompat in a “normal” web browser. Fiddler was also essential for debugging

initial problems with Dingo timetable downloader. This tool was utilized on Wompat when all information was sent in the clear over the Internet, which is before it became SSL encrypted.

7. Resources

7.1. *More Information on Week Numbers*

- “The Mathematics of the ISO 8601 Calendar”,
<http://www.phys.uu.nl/~vgent/calendar/isocalendar.htm>
- “Using C# to Calculate Week number for a date using calendar and culture”,
http://konsulent.sandelien.no/VB_help/Week/
- “Java Forums - How to get ISO week number?” ,
<http://forum.java.sun.com/thread.jspa?threadID=585484&messageID=3023960>

7.2. *Fiddler HTTP Debugging Proxy*

- The Fiddler application,
<http://www.fiddlertool.com>

7.3. *iCalendar and vCalendar*

- “RFC 2445 - Internet Calendaring and Scheduling Core Object Specification (iCalendar)”, <http://www.faqs.org/rfcs/rfc2445.html>
- “vCalendar Specification”, <http://www.imc.org/pdi/pdiproddev.html>
- “Using PHP to Make Basic vCalendar/iCalendar Events”,
<http://www.phpbuilder.com/columns/chow20021007.php3>

- Discussion on vCalendar repeating events problems with Outlook,
[http://mail-archives.apache.org/mod_mbox/jakarta-tapestry-user/200409 mbox/%3C03a501c49b63\\$ec664750\\$8b688f80@arden.virginia.edu%3E](http://mail-archives.apache.org/mod_mbox/jakarta-tapestry-user/200409 mbox/%3C03a501c49b63$ec664750$8b688f80@arden.virginia.edu%3E)