

Manual of BilderKatalog

Martin Sprengel

February 27, 2012

Contents

1	Motivation: Why another image database	1
2	Introduction	2
2.1	General data structures	2
2.2	General usage	2
2.3	Interaction	3
3	Installation	3
3.1	Overall process	3
3.2	The configuration form (installation mode)	3
4	Using BilderKatalog	4
5	Import/Export	4
5.1	Overview	4
5.2	Working with variables	4
5.3	How it works in practice	4
5.4	Special notes	5
5.4.1	IPTC encoding	5
5.4.2	Line breaks in metadata	5
6	SQL	5
7	Regular expressions	6
A	GNU GENERAL PUBLIC LICENSE - Version 2, June 1991	7

1 Motivation: Why another image database

Once you have a lot of images especially from digital cameras you get easily lost through your thousands of images. You can organise your images in the directory tree of your hard drive, e.g. according to the time the pictures were taken. Then you get a structure like Images/2010/04/Eastern. But such a tree can not group the pictures according to different categories (e.g. date, people, place,...). Therefore you need a specialized program to organise your pictures.

This program should be able to

1. read data from the files (e.g. date, size)
2. have an input form where the user can enter additional information (e.g. comment, people)
3. have some capabilities to work with this data (e.g. edit, replace)
4. search for certain properties
5. display pictures which satisfy certain criteria

There are already a lot of programs providing these features; but none of the ones I tried satisfied me in every respect.

For me, it is especially important to be able to enter the data very efficiently as I have a lot of pictures to care about. Hence, this program is optimized on fast and easy input while still being very flexible and satisfying all the other criteria.

2 Introduction

2.1 General data structures

This program distinguishes between three different kinds of data:

1. the image file
2. auto data: data which are automatically extracted from the files like filename, date and image size
3. manual data: data entered by the user like comment, people and place

In all processes the image file is left unchanged.

Both auto and manual data are stored in a separate database, a so called SQL database. Currently only a MySQL [2] backend is supported. More information on SQL in section 6.

Apart from adding new entries to the database or comparing database and file system, the files are ignored and only the database is taken into account. For example when search for all files in a directory, the program finds all the file which have database records with this path name regardless whether they still exist or not. Newly created files are not found until they are explicitly added to the database.

In all appropriate situations the set of file to work on can be selected via a search in the database using the central search form.

2.2 General usage

In short, the general usage is as follows

1. Add files to the database via *Add file* or *Add directory*
2. Enter manual data
3. Use data to

- (a) display a selection
- (b) edit data
- (c) do statistics

2.3 Interaction

This program is only an interface to manage image files stored in folders and meta data stored in a database.

The data storage and some management functions are done by a database server: MySQL. Therefore this server has to be installed before installation of this program and must always run when executing BilderKatalog .

3 Installation

3.1 Overall process

First, you have to install a MySQL server. The MySQL server is a free program that can be downloaded from <http://dev.mysql.com/downloads/>.

Then copy all the files from the BilderKatalog archive to the folder you like to install it to.

Finally, all you have to do is to start the program for the first time to get the configuration form. After having filled in everything simply restart and enjoy!

3.2 The configuration form (installation mode)

License You have to accept the GPLv2 license [3] of this program in order to use it. In short the license tells you that you are free to use the program as you like and free of charge.

Furthermore, you can read and modify the source code, but you have to publish any changes you make to the source code.

The whole license can be found in the file license.txt and in the Appendix A.

Database Here you have to enter the database name you want to use and a user name/password combination for the MySQL server. Standard values are

database jdbc:mysql://localhost:3306/bilder

user root

password root

All values can be changed later.

The rest All the rest is not mandatory at installation time. All entries including the database ones can be changed later for the configuration form (see section ??).

4 Using BilderKatalog

5 Import/Export

5.1 Overview

The metadata from the database can be exported to and imported from the corresponding JPEG files. This is done with the *ExifTool* by Phil Harvey. ExifTool is an open source Perl library with an executable command line interface and can be found at <http://owl.phy.queensu.ca/~phil/exiftool/index.html>.

Please install ExifTool if you want to use this feature. In case of problems it can be helpful to consult the ExifTool homepage or try ExifTool directly (on the command line or with the Windows GUI).

Writing can be undone 100%. For example: At the beginning the image files contain no comment. Then the comment is exported from the database to the file. Later the comment can be erased by exporting an empty field to the files. These new files are binary identical to the original files. To achieve this you have to select the same variables for deleting that you selected for writing, otherwise you will either not remove all written data or remove entries that the image contained from the beginning.

All tags supported by ExifTool can be used; these include the EXIF header and XMP and IPTC standards.

5.2 Working with variables

For the export and import, first variables are filled from the existing data. A variable is always a % character followed by a single letter. Each variable must be unique. Then in the second step the data is filled into templates defined for the output. Let's have an example:

Input data	Berlin, Germany
Input pattern	%t, %c
Values of variables	%t=Berlin %c=Germany
Output template	My town is %t which is in %c.
Resulting output	My town is Berlin which is in Germany.

In the simplest case, if you just want to copy the data, every input field and every output template consists of exactly one variable. But you could copy the data from all fields to a single output field (for example EXIF comment field).

5.3 How it works in practice

First define how the data in the database is structured in the left half of the window. If there is now special structure just let every field be one variable. Empty fields are ignored.

Then in the right half you can set up the fields in the file. For a complete list see the ExifTool manual. Rows can be (de)activated by the checkbox in the first column. The second column gives the name of the data field as called by ExifTool. The third column contains the template that is filled with the data

(or in the import the pattern that is expected for reading). After setting up everything just click import or export.

A click on default values closes the window and deletes all safed changes. Upon reopening the default values are restored. For debugging purposes the output of ExifTool can be displayed.

5.4 Technical notes

5.4.1 IPTC encoding

As there is no standard encoding for IPTC, BilderKatalog always writes in UTF-8 and sets the value `codedcharacterst` accordingly. To completely remove all comments to have the images binary identical to the original this header has to be removed manually by calling `exiftool -codedcharacterst=` on the file. Maybe this will be implemented later.

5.4.2 Line breaks in metadata

Without any special procedure line breaks are correctly exported from the database to the files. The other way round is not that easy because ExifTool replaces every control character like a line feed or carriage return by a dot.

So far there is one way around: There is an option to replace an arbitray sequence of the comment field (the only true multiline field in the database) with a line break. If you did not use a dot in your comments, you can simply replace every dot with a new line and you are done.

But what if you have used dots and you want to distinguish them from new lines? You have to define a virtual field for ExifTool – say `mycomment` – which is the exif comment but all line breaks replaced with `^^` because you did not use that. Then you can read in the field `mycomment` just as any other field and safely replace `^^` with new lines.

So how to define virtual fields for ExifTool? This can be done in the config file (on *nix this is `~/ExifTool_config`). The following entry will do the replacement described above:

```
%Image::ExifTool::UserDefined = (  
'Image::ExifTool::Composite' => {  
MyComment => {  
Require => 'Comment',  
ValueConv => '$val =~ s/[\n]/^^/g ? $val : undef',  
},  
},  
);
```

6 SQL

SQL means *Structured Query Language* and is a formal way of requesting something from a database.

For the user of BilderKatalog the most important query is a `SELECT` query which gets some data form the database. This `SELECT` query can have a

WHERE clause to select only these rows (i.e. the information on these pictures) that match certain rules.

The WHERE clause typically asks a *field* to have a certain *value*. For example, the path of the file should be the given one (say `/home/martin/pictures/`). Then the WHERE clause would be: `WHERE 'pfad' LIKE '/home/martin/pictures/'`.

Some rules:

- The field name must be enclosed in backticks (`'='`) and text (string) values and dates in apostrophes (`'='`). Numbers must not be enclosed by anything.
- You need a relation to compare them. For strings there are:
 - LIKE
 - NOT LIKE
 - REGEXP
 - NOT REGEXP

and for numbers and dates the usual ones (`=`, `>`, `<`, `<=`, `>=`, `!=` (not equal)).

- You can combine conditions by AND and OR.

Some examples:

- `'pfad' LIKE '/home/martin/pictures/'`
- `'name' LIKE 'p10794847.jpg'`
- `'iso' > 400`
- `'datum' > '2009-05-27 22:40:49'`
- `'pfad' LIKE '/home/martin/pictures/' OR 'pfad' LIKE '/home/sprenkel/pictures/'`
- `'iso' = 200 AND 'name' REGEXP 'p107.*jpg'` (see section 7)

Wherever you are asked for a WHERE clause you should enter something like those above.

A general SQL statement needs some more information and can be rather complex. The complete SELECT could be `SELECT * FROM 'bilder' WHERE 'iso' > 400 ORDER BY 'datum'`

You have the possibility to execute pure SQL statements from the edit menu but before you should learn more about SQL on the internet. See [1].

7 Regular expressions

A regular expressions is a pattern describing a set of strings (text values). There are a lot of rules what they can do but there are some simple basics:

- Every letter or number is just itself
- The period (`.`) is any single character

- Square brackets are a set of possibilities (`[abc]` means any one of a, b or c). They can be given by a range: `[a-z]` is any lower case letter.
- A star (*) repeats the previous any number of times including never (`d*` means zero to infinity d's)

With this we can construct already a lot of examples:

- `p107.*jpg` matches any string that starts with `p107` and ends with `jpg`, e.g. `p107jpg`, `p107852.jpg`, `p107LK5jpg`
- `p107[0-9]*jpg` matches any string that starts with `p107`, ends with `jpg` and has only digits or nothing inbetween, e.g. the first two of the above examples but not the last
- `Martin` matches only this single string.

For more detailed information please refer to [4].

References

- [1] <http://en.wikipedia.org/wiki/SQL>
- [2] <http://mysql.com>
- [3] <http://www.gnu.org/licenses/gpl-2.0.html>
- [4] <http://perldoc.perl.org/perlre.html>

A GNU GENERAL PUBLIC LICENSE - Version 2, June 1991

Copyright © 1989, 1991 Free Software Foundation, Inc.

51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software—to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish),

that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this

License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:
 - (a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.
 - (b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.
 - (c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

- (a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,
- (b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,
- (c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.
5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.
6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms

and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.
9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and “any later version”, you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.
12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS

Appendix: How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively convey the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

one line to give the program's name and a brief idea of what it does.
Copyright (C) yyyy name of author

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA.

Also add information on how to contact you by electronic and paper mail. If the program is interactive, make it output a short notice like this when it starts in an interactive mode:

```
Gnomovision version 69, Copyright (C) yyyy name of author
Gnomovision comes with ABSOLUTELY NO WARRANTY; for de-
tails type 'show w'.
This is free software, and you are welcome to redistribute it under
certain conditions; type 'show c' for details.
```

The hypothetical commands `show w` and `show c` should show the appropriate parts of the General Public License. Of course, the commands you use may be called something other than `show w` and `show c`; they could even be mouse-clicks or menu items—whatever suits your program.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a “copyright disclaimer” for the program, if necessary. Here is a sample; alter the names:

```
Yoyodyne, Inc., hereby disclaims all copyright interest in the pro-
gram
'Gnomovision' (which makes passes at compilers) written by James
Hacker.
```

```
signature of Ty Coon, 1 April 1989
Ty Coon, President of Vice
```

This General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Library General Public License instead of this License.