

ソフトウェアライセンスと その周辺の話題

村脇 有吾

1 March 2012; Updated 17 April 2012

DISCLAIMER

- 法律の専門家でも何でもないので、内容の正確性は保証できません
- この発表を信じた結果、何らかの損害が生じたとしても、一切責任は負いません

念頭においている疑問

- 自分が作ったソフトウェアをどのようなライセンスで配布すればよいか?
- 他人が作ったプログラムをどういった条件なら流用して良いか?
- 自然言語処理ではプログラムだけでなく、データも重要だが、ライセンス的にはどういった問題があるか?

本日のお題

- ライセンスとは何か
- ソフトウェアライセンス
- ドキュメントなどのライセンス
- プログラムとデータの関係
- プログラムを介したデータ同士の関係
- 他の権利との関係

そもそも著作権とは何か

- 著作者が著作物について持っている財産権
 - － 譲渡可能
- 著作物を作ったら自動的に発生
 - － 届け出などは不要
- 著作権者は、他人に対し、著作物の利用を許諾することができる
 - － 利用の条件を定めたのがライセンス

ライセンスとは何か

- ライセンサからライセンシへの利用許可
- 利用にあたっての様々な条件
 - ライセンシによる再配布、改変の条件
 - ライセンサの免責
 - ライセンサが一定の権利を行使しないことを保証

ソフトウェアライセンス

- オープンソース業界では既存のライセンスを利用するのが普通
- オープンソースライセンスだけでも山のようにある
 - Open Source Initiative認定済だけで60ぐらい
- 代表的なオープンソースライセンス
 - 修正BSD
 - GPL v2 (GNU General Public License)
 - LGPL v2 (Lesser GPL)

ソフトウェアライセンス：修正BSD

- ライセンサは無保証（免責）
- ライセンシは再頒布時に著作者を明記し、ライセンスを頒布する義務がある
- ライセンシが派生物を同じライセンスで頒布する義務は**ない**
 - ライセンシが派生物のソースコードを公開する義務はない
 - 商業利用との親和性が高い

ソフトウェアライセンス: GPL v2

- ライセンサは無保証 (免責)
- ライセンシは派生物を同一ライセンスで頒布する義務がある (copyleft)
 - 自分のプログラムにGPL由来のソースコードを混ぜると、全部GPL (互換) で頒布しないといけない (ウイルス性)
 - GPL互換でないライセンスのソースコードをGPLのソースコードと混ぜて頒布できない

ソフトウェアライセンス: LGPL v2

- ライブラリ向けに設計されたGPL互換 (LGPL → GPL) のライセンス
- LGPLのライブラリAとプログラムBをリンクするとき、Bを (L)GPLで配布する義務がない
- LGPLのライブラリそのものを改変した場合は、それを (L)GPLで頒布する義務がある

GPLの影響範囲 1/2

影響範囲は基本的には各国の著作権法に依存する問題だが、GPLが比較的詳細に解説しているので紹介

- 静的, 動的リンクするプログラム: 影響が及ぶ
- プログラムが読み書きするデータ: 影響が及ばない
- 集積物 (aggregate): 頒布用に別々のプログラムをまとめたもの。相互に影響が及ばない

GPLの影響範囲 2/2

- どこまでが動的リンクか?
 - fork-exec は通常は動的リンクではない
 - 相互に関数呼び出しと構造化データの受け渡しを行うなら動的リンクと見なしえるかも (FSF見解)
 - pipe, socket通信なども同様 (FSF見解)
 - RESTful API は影響が及ばない?

GPLにおける頒布の罣

- GPL上の諸々の義務が発生するのは頒布時
- 個人/会社が内部利用するだけ (二次的ライセンス不在) の場合は義務が発生しない
- ソフトウェア自体を頒布せず、サービスとしてのみ使う場合も義務が発生しない
 - この抜け穴をふさいだのがAGPL v3

複数ライセンス

- MeCabは修正BSD, GPL, LGPLのトリプルライセンス
- 派生物は複数ライセンスをそのまま引き継げる
- いずれかのライセンスのみ (e.g. 修正BSD) を選択することも可能

ドキュメントなどのライセンス

- ソフトウェア用ライセンスには、「オブジェクト」など、ドキュメントにおける対応が不明な概念がある
- ドキュメントやその他のデータ向けのライセンス
 - GFDL (GNU Free Documentation License)
 - Creative Commons licenses

GFDL

- GPLの文書版
 - ウィルス性をもつ
- 表表紙、背表紙、変更不可部分など、「本」形式を想定した煩雑な規定がある
- GPLと双方向に互換性がない!
 - GPLにない制限を課すため

Creative Commons (CC) licenses

- 6通りの組み合わせからなるライセンス群
 - 表示 (BY): 著作権者の表示
 - 非営利 (NC): 非営利目的のみの利用
 - 改変禁止 (ND): 派生物の作成禁止
 - 継承 (SA): ライセンス継承義務 (ウィルス性)
- CC-BY-SAはGFDL v1.3と互換性がある
- WikipediaはCC-BY-SA 3.0

プログラムとデータの関係

- ライセンスの影響がどこからどこまで及ぶか?
- OSS界隈で比較的話題になる疑問:
 - プログラムのライセンスがデータに及ぶか?
- 派生物 (derivative work) の定義: 各国の著作権法に従う
 - 日本の場合は「二次的著作物」

例1: Bison

- パーサ生成器 (GPL頒布)
 - 入力: 文法規則と対応する処理 (ソースコード)
 - 出力: パーサのソースコード
- Bison自身のソースコードを出力パーサに埋め込む
- ソースコードがGPLに汚染される?
- 例外条項を設けて出力にGPLが及ばないようにしている
 - 原著作者による権利の放棄だからGPLに抵触しない。同じ例外条項を設けるように派生物に強制できない。また、原著作者の許可なしに、派生物にこうした例外条項を新たに追加できない。

例2: WordPressのテーマ

- WordPress: ブログ生成ソフトウェア (**GPL**頒布)
 - 言語: PHP (HTML, CSS, 画像含む)
- テーマ: 見た目などを変更する追加パッケージ
 - 言語: PHP (HTML, CSS, 画像含む)
- テーマのライセンス問題 (Software Freedom Law Centerの見解)
 - テーマのうち**PHP**はWordPressの派生物
 - CSS, 画像は別の作品とみなしてもよい

プログラムを介した データ同士の関係

- プログラムがデータを加工したとき、
データ同士のライセンスの影響関係は?
- 自然言語処理 (特に解析まわり) で問題になる
- OSS界隈ではあまり議論されていないので
何が正解かよくわからない

例1: GNU FreeFont

- (プログラムとして) GPLで頒布
- 例外条項: このフォントを埋め込んだ文書にGPLの影響は及ばない
 - Bisonと同様、原著作者による権利の放棄だからGPLに抵触しない。同じ例外条項を設けるように派生物に強制できない。また、原著作者の許可なしに、派生物にこうした例外条項を新たに追加できない。
- 未解決?の問題: 文書から抽出して再構成したフォントのライセンスは?

例2: 辞書を用いたテキスト解析

- 解析プログラム (データへのライセンス影響なし)
 - 読み込みデータ: 辞書 (修正BSD?)
 - 入力: テキスト (任意のライセンス?)
 - 出力: 解析結果 (ライセンス???)
- 解析結果には辞書 (の一部) が埋め込まれる
 - 解析結果は入力テキストと辞書の派生物
 - JUMANの辞書は意味情報のアノテーションが豊富なので、創造性を有すると考えられる
 - 辞書が修正BSDライセンスなら、著作者表示とライセンスの頒布が必要
- 解析結果を制限なく頒布できるようにするには、辞書をフォントと同様の扱いにすべき?

例3: Wikipediaからの 解析用辞書の自動構築

- 変換プログラム (データへのライセンス影響なし)
 - 入力: Wikipedia (CC-BY-SA)
 - 出力: 解析用辞書 (ライセンス???)
- 立場1: Wikipediaの著作権が及ばない
 - 創造性に欠けるから
 - 反論1: 編集著作物として創造性を有する?
 - 反論2: 本文冒頭から上位語までも抽出しているので、創造性がないとは言えない?
- 立場2: CC-BY-SA
 - 著作権があるので、ライセンス継承の義務がある

例4: ウィスル性の辞書を含んだ 解析ソフトウェアの頒布

- パッケージ構成
 - プログラム (修正BSD)
 - 普通の辞書 (修正BSD?)
 - Wikipediaから抽出した辞書 (CC-BY-SAと仮定)
- 集積物とみなせ、個別の作品が別の作品のライセンスに影響を及ぼさないはず
 - 作品ごとにライセンスを明記する必要

例5: ウィルス性の辞書を用いた テキスト解析

- 解析プログラム (データへのライセンス影響なし)
 - 読み込みデータ: 辞書 (CC-BY-SAと仮定)
 - 入力: テキスト (ライセンス任意?)
 - 出力: 解析結果 (ライセンス?)
- 解析結果がCC-BY-SAに汚染される?
 - 入力テキストもCC-BY-SA互換でないといけない?
- 解析結果自体を頒布しなければどうということはない?
 - 解析結果から得られる統計情報だけを頒布するという方針ならOK?

その他の権利との関係: 特許

- ライセンスにとっての罠: 「ソフトウェア自体は無償だけど、**特許料は別だよ**」と突然ライセンサに請求される
- GPL v2: ライセンサが暗黙のうちに特許利用を認めているとみなせる
- Creative Commons: 特許に触れていないのでライセンスにとって危険
- 修正BSD: 同様に実はヤバい?

蛇足：その他の権利(?) との関係： 場所の占有に意味がある場合

- 例1: Wikipediaはwikipedia.orgに置かれているからこそ持続的な発展が見込める
 - 誰にでもクローンは作れるけど発展しない
 - wikipedia.orgに投稿するための条件として、投稿者に追加制約を課せる
 - GFDL (バージョン指定、変更不可部分、表表紙、裏表紙なし) でもライセンスすることを投稿者に認めさせている
- 例2: FSFにソースコードを寄贈するとき、GPLなのに著作権を譲渡しなければならない
 - 著作権者が多いと訴訟がやりにくいから