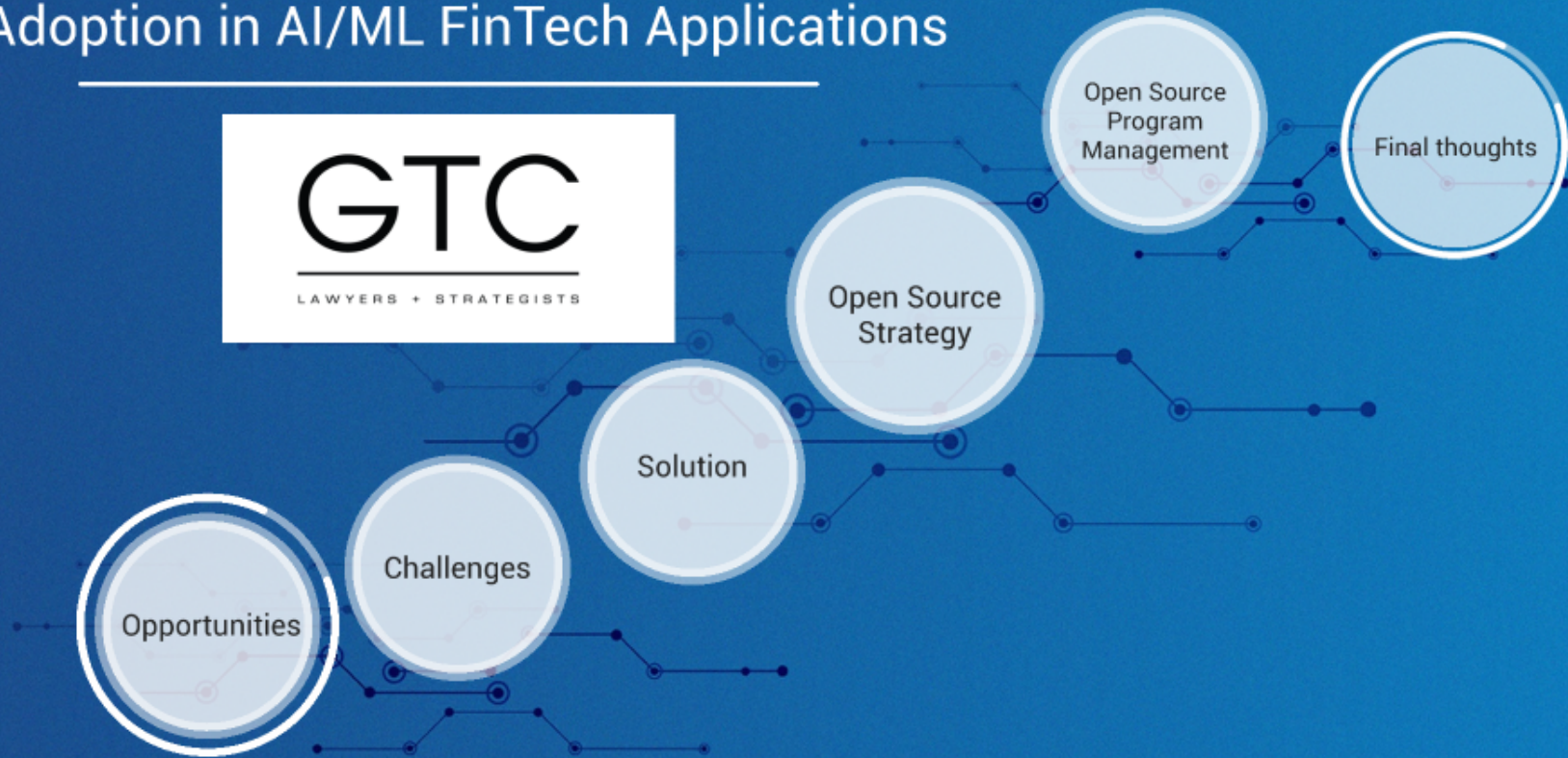


Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Opportunity

- 80% of enterprises are actively using one or more elements of AI, including ML and NLP
- 36% of senior decision makers said they would be making AI investments in the next 3 years.
- Financial services are in the top 5 sectors where AI will have a +ve impact
- - Vanson Bourne, "State of Artificial Intelligence for Enterprises"

- What does AI mean?
 - Human computer interactions
 - Personalization
 - Decision making
 - Pattern detection and recognition
 - Prediction



AI unlocks new competitive strategies for financial institutions across value chains

Opportunities: Overview

Radical
Change in
Capabilities
and
Business
Models

Cheaper and more efficient services

- "just in time" lending
- improve underwriting
- automate compliance and reporting
- automated portfolio construction

Tailored Products and Advice

- provide detailed advice realtime
- increase capabilities of advisors
- bespoke incentives
- highly personalized insights for advisors

Greater Interaction with Customers

- always-on virtual agents
- mobile first insurance offerings
- "invisible payments infrastructure"
- tap into 3rd party platforms to distribute services

Improved decision making

- predict defaults accurately
- new and unique correlations between datasets
- improve deal identification, pairing and sales activities
- improve trade speed and price using dynamic execution methods

New Value Propositions

- use proxy data to insure new risk categories
- develop unique strategies and new investment products
- develop real-time pre- and post-trade risk management solutions
- deploy new order types to protect investors from risks proactively

AI/ML Complexity

(Adapted from "The New Physics of Financial Services," <https://www.leadsun.org/insight/NFT-New-Physics-of-Financial-Services.pdf>)

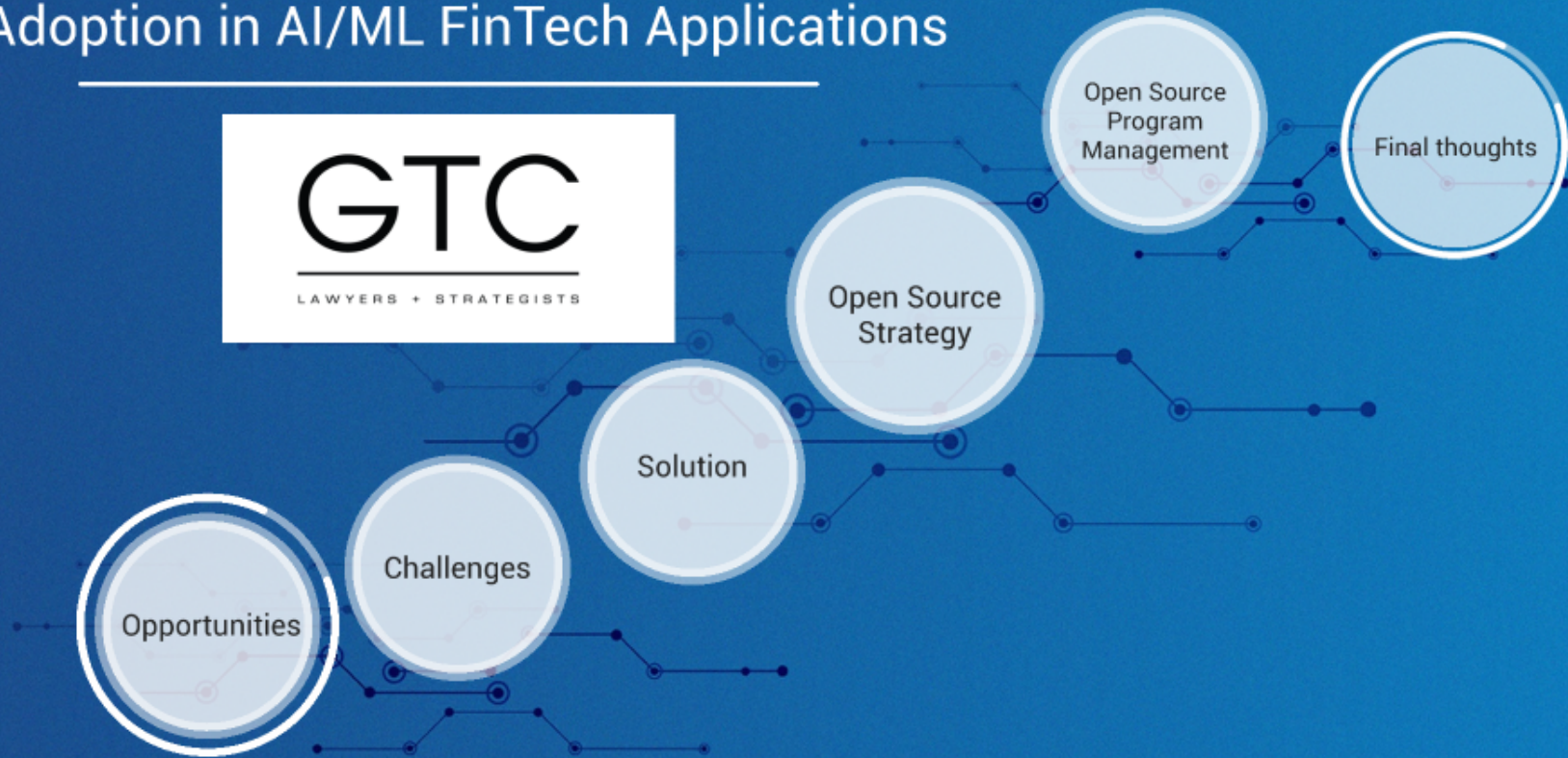
IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Challenges

Overcoming execution challenges and meeting timelines ...

(see "The New Physics of Financial Services," http://www3.weforum.org/docs/WEF_New_Physics_of_Financial_Services.pdf)



Infrastructure

Talent

Data

Regulation

Challenges: Infrastructure

AI/ML solutions consist of resource intensive algorithms ingesting vast quantities of data.

Efficient data management requires: sufficient processing power, flexibility, scalability and capacity to accommodate different types and volumes of data.

Challenges:

- Legacy systems and technical debt -> need major restructuring (e.g. API configuration and handling real-time data flows)
- Extent of process re-engineering (re-map data flows to support increased interaction between machines and humans; privacy/security/regulatory impact)
- Multiparty effort to transform processes across value chains (intermediaries are resistant to change)
- Lack of agile, cloud-based architecture (limits storage, processing power and slows updates)

Challenges: Data

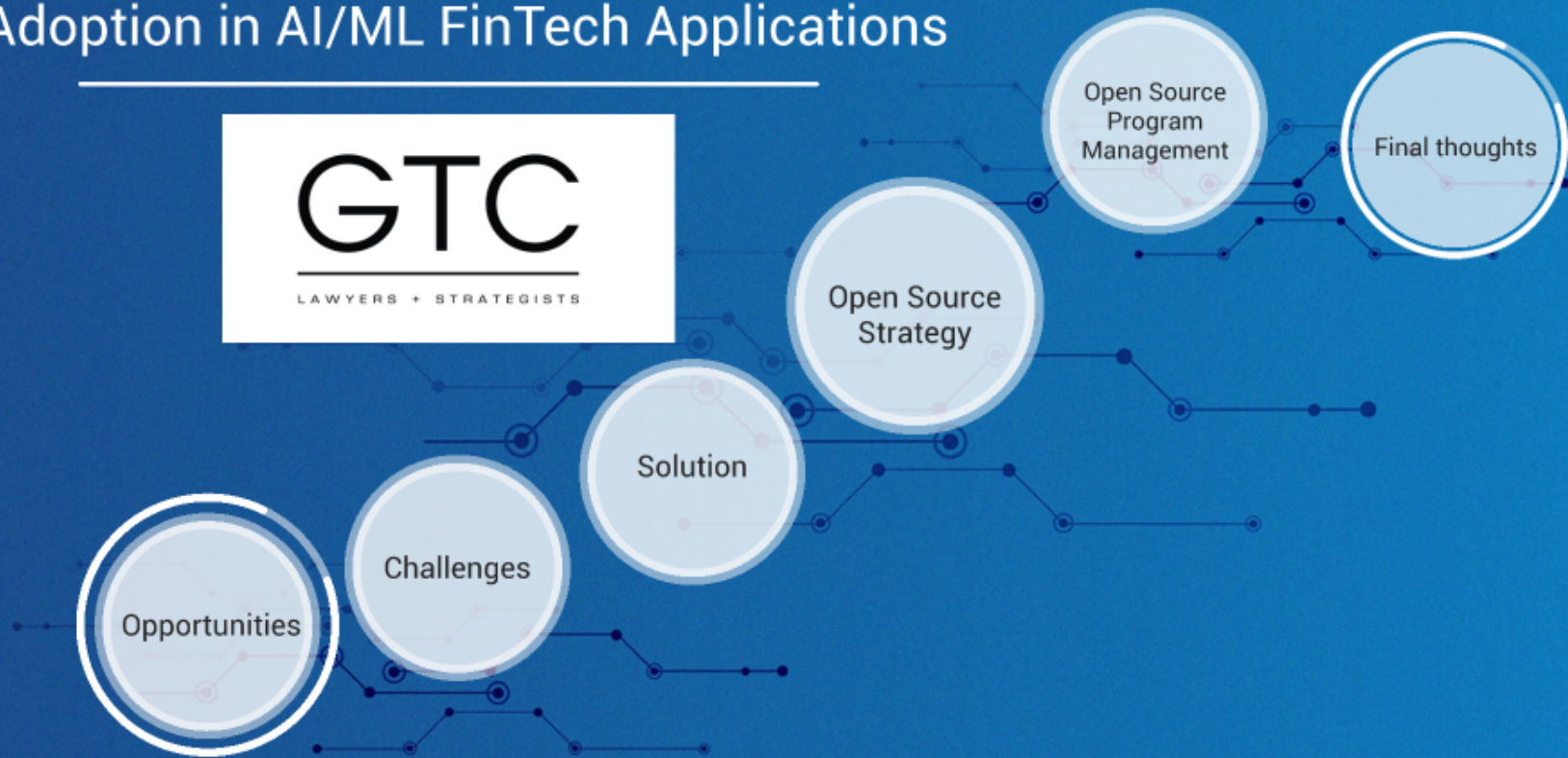
A large pool of data is not sufficient...

- Dealing with formatting inconsistencies, missing data and errors
- Fragmented internal data storage (need significant re-engineering to create data lakes)
- Insufficient breadth and depth of data (require data partnerships)
 - Requires creation of standards to share across institutions (e.g. for improving trade speed and price across market infrastructure)
- Lack of data digitalization

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Solutions

Companies
are open
sourcing AI
projects

Case Study (1)

Case Study (2)

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Open Source AI Makes Business Sense.....

- Technology companies are open-sourcing their AI systems
 - Amazon's Alexa
 - Google's TensorFlow
 - Even DARPA!
- Why?
 - Complexity of AI systems
 - Less value in the actual code than in the way data is utilized
 - Improving functionality; configure code with data sources
- Open-source ML automation tools typically address a subset of the core processes associated with modeling, training, and refinement.
 - Most support ML programming in Python or R and come integrated with common ML libraries (e.g. TPOT, Scikit, Machine JS, etc.)
 - On the other hand, commercial ML automation products also provide ML-pipeline automation features including front end collaboration, exploration, processing, and back-end management (e.g. H2O.ai Driverless AI, TellMePlus, etc.)

Open Source FinTech Tools: Case Study (1)

- **Data storage**

- Man AHL open source licensed its tick store (specialized type of data tool for storing high frequency financial services market data), Arctic.
- Arctic is the primary market data store for the hedge fund manager's quantitative researchers ("quants"); impressive scalability (~hundreds of millions of rows per second for each MongoDB instance)
- Licensed under LGPL v.2.1
- By building Arctic on MongoDB, Man AHL realized a 40x cost savings when compared to its legacy time series data store and processing performance improved by 25x.
- Use cases for write-intensive workloads outside financial service e.g. IoT applications and social media feeds

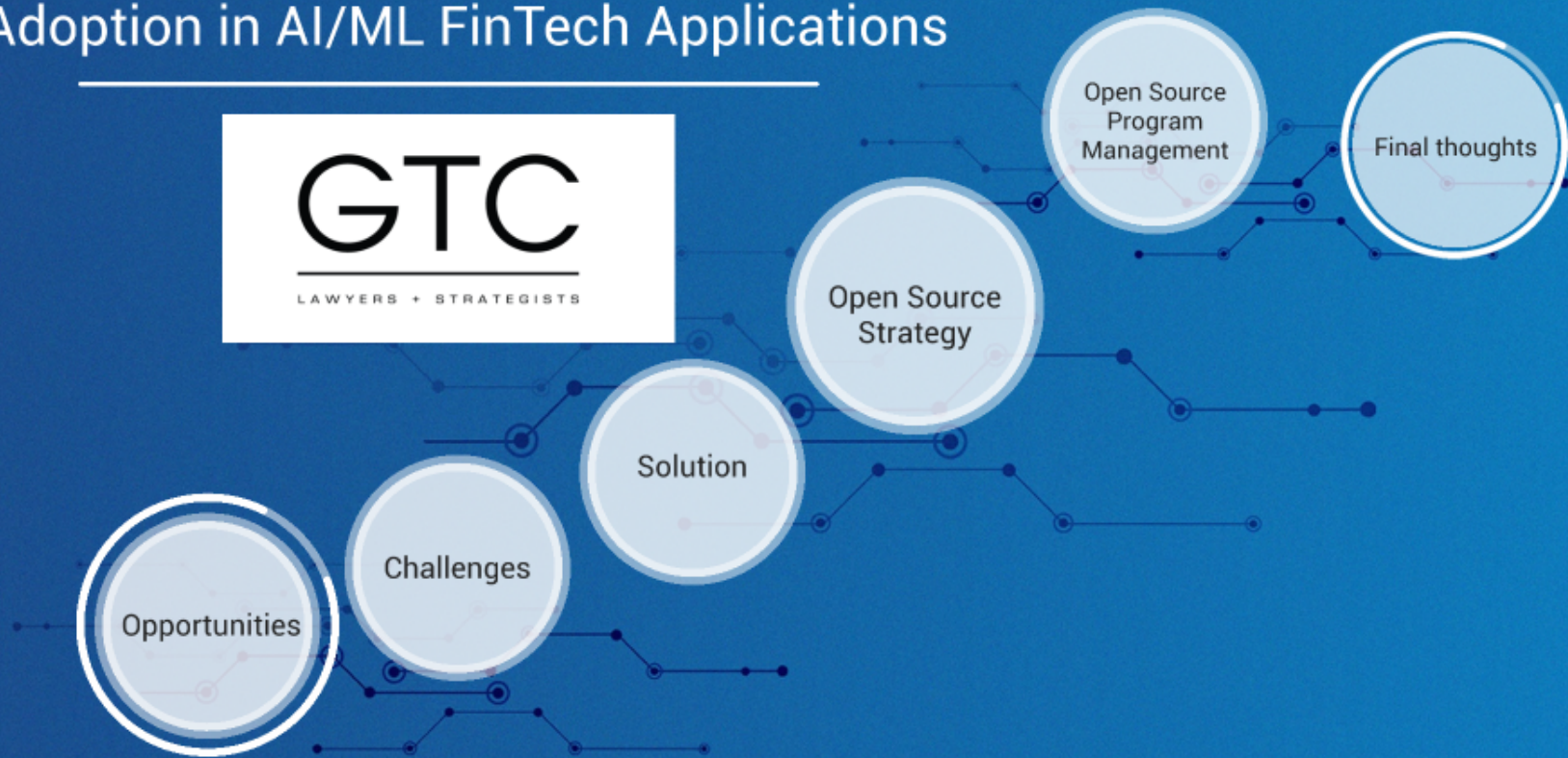
Open Source FinTech Tools: Case Study (2)

- **Data analysis tools**
- Popular data analysis tool Pandas was developed at AQR Capital Management
- BSD licensed
- Pandas address key features missing in Python (handling CSV files, dealing with spreadsheets and other data manipulation); Pandas instrumental for popularity in Python
- Benefits to AQR more nuanced:
 - Indirect benefit to popularizing underlying programming tool Python
- Pandas creator now works at another fund, Two Sigma, and continues to work full time on Pandas
 - Mandate includes making sure data science tools for Python continue to develop
- Open source project longevity is an important goal
 - Without community support, no bug fixes or patches

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Open Source as a Business Strategy

1. Analyzing Goals for the Open Source Project

- To what extent does your product implement a new platform e.g data storage or automation?
- Is it in your interest to maintain ownership of that platform
 - Lower development costs
 - Better product -> more users/customers
 - Competitive advantage?
 - Counter: Infrastructure costs

Adapted from "Open Source as a Business Strategy," Brian Behlendorf

2. Evaluating market need for your product

- Conduct a competitive analysis of the space
- Separate offerings into chunks that could be bundled/licensed/sold/open-sourced separately

3. Marketing your vision

- Determine minimum resource set

4. Which license to use?

- Permissive v. viral

Example

ML Product = Data Pre-processing + Feature Engineering/algorithm selection/model training + Model Evaluation

Conclusions:

- Company's core data processing tools provide only a little more functionality than Pandas; this feature is not a competitive advantage
- the model evaluation has no freeware competition
- the developer libraries are surpassed by Python but has little Java or R competition

Option 1:

- replace core data processing tools with open source Pandas; bundle up with the model evaluation solution, and sell the Java and R libraries while providing and supporting the free Python library

Option 2:

- focus on the model evaluation solution to be sold as a separate product

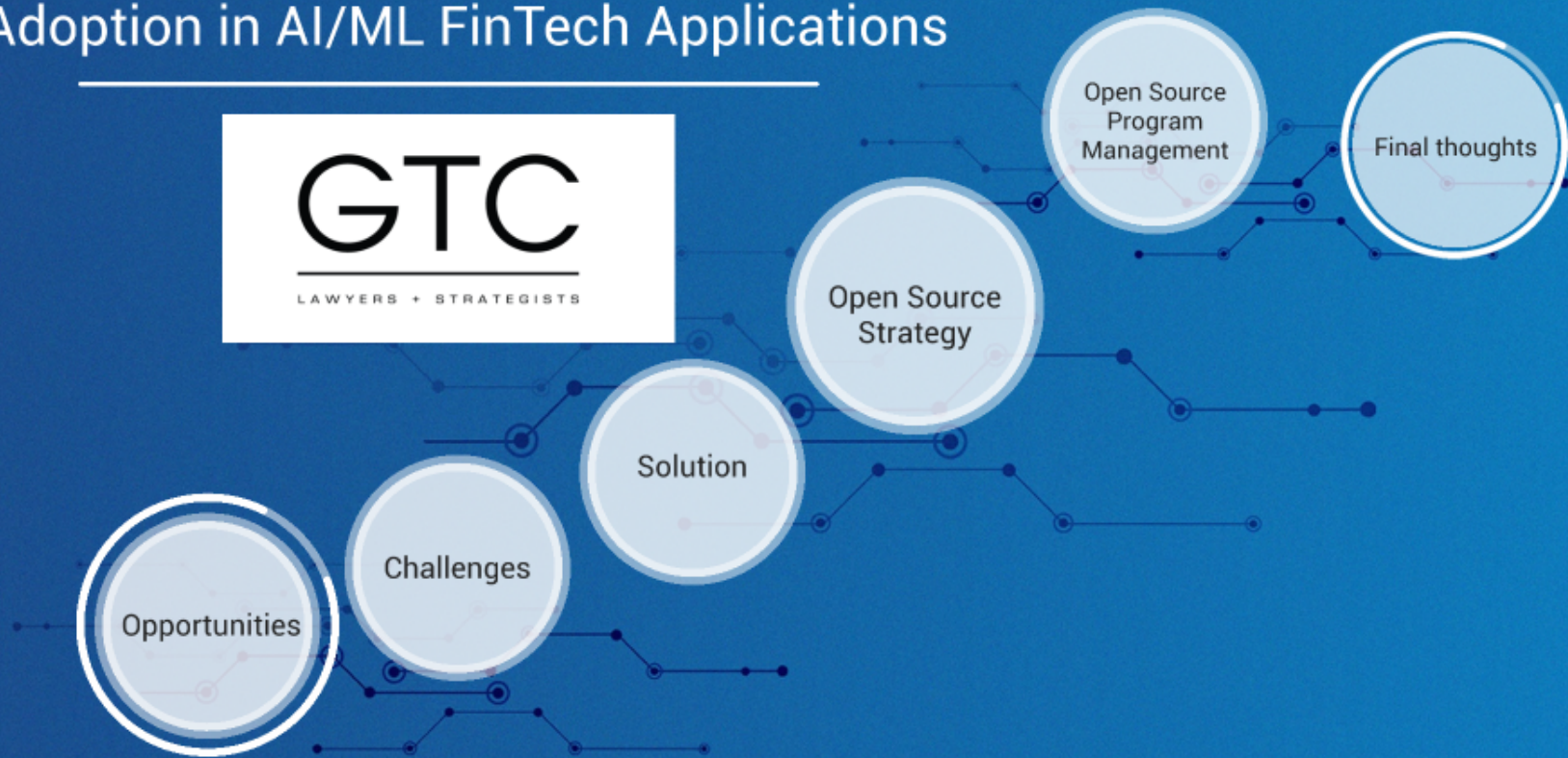
Option 3:

- open source the entire product

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Open Source Management Program: Compliance

1: Administration

2: Compliance Management

Disclosure,
Review and
Approval

License
Compliance

Remediation

3. Audits and Verification

Scope of Review: General

Define and describe the approval process for Third Party Software in connection with Company's products and services

Third Party Software means third party commercial and open-source software (including object code, binary code, source code, firmware, microcode, drivers, libraries, routines, subroutines or other code, application programming interfaces and software development kits, all whether commercial, open source or freeware) and any documentation or other material related to such software provided by a third party

Includes software that is combined with Company products and software that is used to develop, test, maintain, offer, and support Company products, even if such software is not included with any distribution or hosted offering

Does not include non-development related internal-use only software, such as back-office accounting software, CRM software not integrated with Company's products or services, and office productivity software

Key Drivers for Review

Inability to cleanly make required representations and warranties in a transaction/IPO

Use, modification, distribution and hosting of third party components beyond the scope of the applicable licenses, possibly in breach of the agreements and resulting in copyright infringement -> Potential: injunctions; lawsuits; significant re-engineering; PR nightmare

'Viral' infection of proprietary code such that the proprietary code must be made available in source code form at low or no cost for the purpose of allowing others to make derivative works and re-distribute the code at no cost (e.g. copy/paste errors, linking)

Undesirable dependence on code from a competitor or other hostile party

Possible automatic grant of licenses to certain of your patents upon distribution and hosting of certain third party components

May cause license to the third party code to terminate if a patent suit is brought against any party that contributed to the third party code

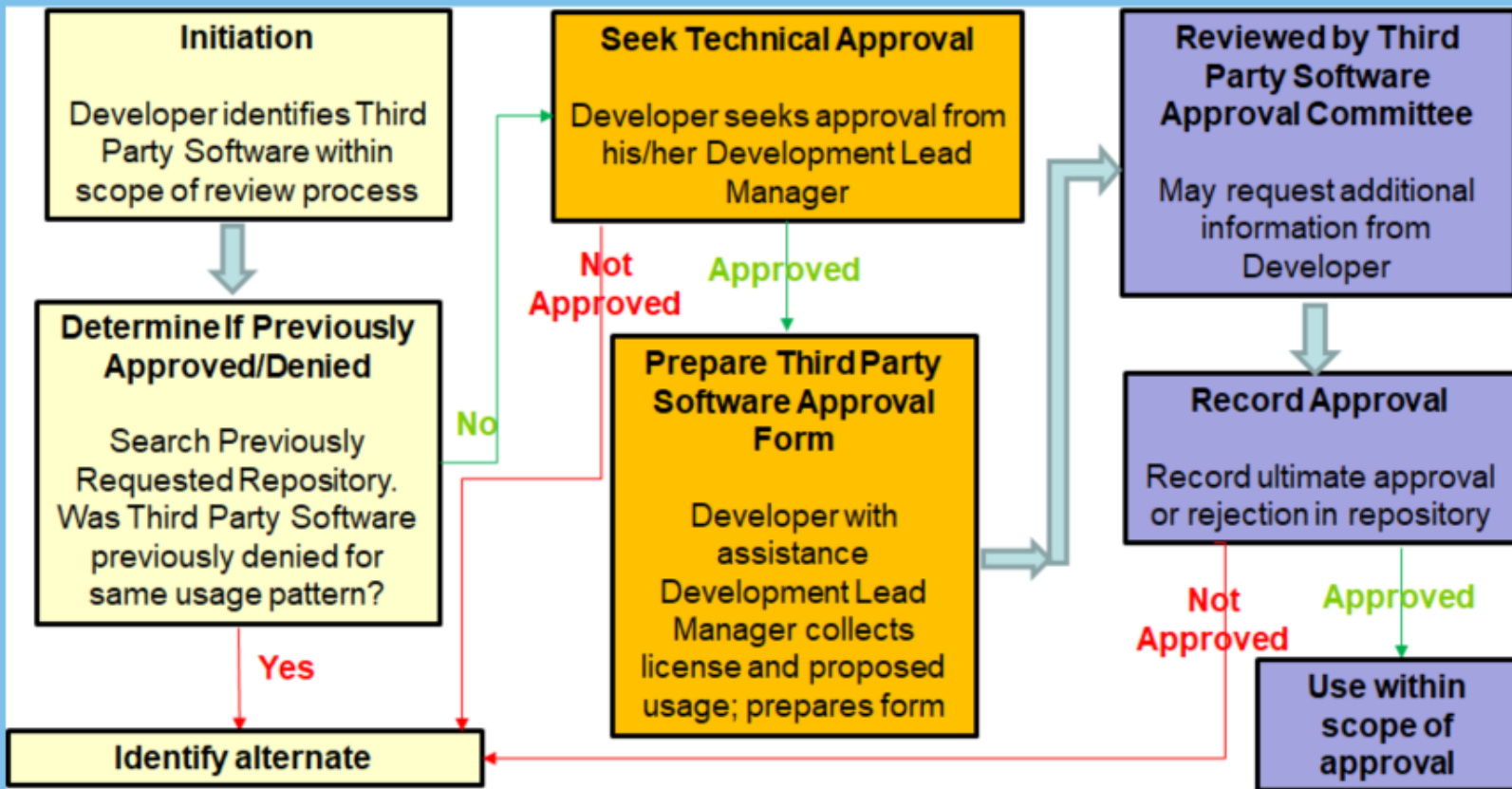
Objectives of Review

Identify, quantify and mitigate inbound licensing-related risks

Understand all dimensions of legal and operational risk associated with use of open source, freeware and commercial third party software in connection with Company's offerings, whether the components are used only internally, distributed or hosted

At a high level, it is a compatibility assessment between the terms of the agreements and the actual use of the components

Review Process: Developer Initiated



Review Process: Periodic Audit of Company Code Bases

Identify



- Aim to identify all of the third party software (both commercial and non-commercial) embedded in or used in the development, maintenance, support and offering of Company's products, along with the applicable licenses and usage facts
- Third party software is identified through self-disclosure, string searching and code scans

Analyze



- Understand incompatibilities between the described or proposed use of a given third party component and the license terms for that component
- Analyze license terms which may be incompatible with current or proposed business practices

Plan/Remediate



- Create a remediation plan to address identified issues
- Remediation may include:
 - Code remediation: removing, rewriting or replacing code
 - Legal remediation: amending agreements, seeking waivers of past liability, re-licensing components and obtaining new licenses
 - Notice and attribution: revising notices and documentation to be in compliance
 - Risk mitigation: allocating risk through the use of representations and warranties, conditions, indemnities and escrow

Review Process: Miscellaneous

Each developer has a duty to report any unauthorized use of Third Party Software of which he or she becomes aware to the Third Party Software Approval Committee

Process for contributing to open source projects

Open Source License Spectrum (common licenses)

| Permissive | Less Permissive | Less Restrictive | Restrictive |
|---|--|--|---|
| <ul style="list-style-type: none"> • MIT License • BSD License • Apache v1.1 • W3C Software Notice and License • Open Symphony License | <ul style="list-style-type: none"> • Apache v 2.0 | <ul style="list-style-type: none"> • Mozilla Public License (MPL) • Common Public License (CPL) • Eclipse Public License (EPL) • Common Development and Distribution License (CDDL) • IBM Public License (IPL) • Microsoft Public License (MS-PL) • Artistic License • Perl Artistic License • Open Software License v2.0 • Creative Commons Attribution v2.0 and v2.5 • Academic Free License v2.1 • Code Project Open License (CPOL) | <ul style="list-style-type: none"> • GNU General Public License (GPL) • GNU Lesser General Public License (LGPL) • GNU Affero General Public License (AGPL) • Creative Commons Attribution NonCommercial ShareAlike License |

Less Restrictive License Examples (MPL-1.1,2.0)

Mozilla Public License 2.0 (MPL-2.0)

1.10. "Modifications"

means any of the following:

- a. any file in Source Code Form that results from an addition to, deletion from, or modification of the contents of Covered Software; or
- b. any new file in Source Code Form that you create that contains or is derived from the contents of Covered Software.

notice of non-compliance with this License from such Contributor, and You become compliant prior to 30 days after Your receipt of the notice.

5.2. If You initiate litigation against any entity by asserting a patent infringement claim (excluding declaratory judgment actions, counter-claims, and cross-claims) alleging that a Contributor Version directly or indirectly infringes any patent, then the rights granted to You by any and all Contributors for the Covered Software under Section 2.1 of this License shall terminate.

3. Responsibilities

3.1. Distribution of Source Form

All distribution of Covered Software in Source Code Form must be made under the terms of this License. You must ensure that the Covered Software is governed by the terms of this License and that recipients can obtain a copy of this License. You may not restrict the recipients' rights in

5.1 or 5.2 above, all end user (resellers) which have been validly

Mozilla Public License 1.1 (MPL-1.1)

8.2. If You initiate litigation by asserting a patent infringement claim (excluding declaratory judgment actions) against Initial Developer or a Contributor (the Initial Developer or Contributor against whom You file such action is referred to as "Participant") alleging that:

(a) such Participant's Contributor Version directly or indirectly infringes any patent, then any and all rights granted by such Participant to You under Sections 2.1 and/or 2.2 of this License shall, upon 60 days notice from Participant terminate prospectively, unless if within 60 days after receipt of notice You either: (i) agree in writing to pay Participant a mutually agreeable reasonable royalty for Your past and future use of Modifications made by such Participant, or (ii) withdraw Your litigation claim with respect to the Contributor Version against such Participant. If within 60 days of notice, a reasonable royalty and payment arrangement are not mutually agreed upon in writing by the parties or the litigation claim is not withdrawn, the rights granted by Participant to You under Sections 2.1 and/or 2.2 automatically terminate at the expiration of the 60 day notice period specified above.

(b) any software, hardware, or device, other than such Participant's Contributor Version, directly or indirectly infringes any patent, then any rights granted to You by such Participant under Sections 2.1(b) and 2.2(b) are revoked effective as of the date You first made, used, sold, distributed, or had made, Modifications made by that Participant.

Restrictions:

Potentially viral

Modifications are covered by third party license

Defensive patent-related termination clauses, broad (as in MPL-1.1) or narrow (as in MPL-2.0)

Automatic termination upon breach

Restrictive License Examples

Very viral, with broad definition of modifications; modifications are covered by third party license
Automatic grant of certain patent licenses
Automatic termination upon breach (i.e. no cure period)

GNU General Public License



To "modify" a work means to copy from or adapt all or part of the work in a fashion requiring copyright permission, other than the making of an exact copy. The resulting work is called a "modified version" of the earlier work or a work "based on" the earlier work.

A "covered work" means either the unmodified Program or a work based on the Program.

To "propagate" a work means to do anything with it that, without permission, would make you directly or secondarily liable for infringement under applicable copyright law, except executing it on a computer or modifying a private copy. Propagation includes copying, distribution (with or without modification), making available to the public, and in some countries other activities as well.

To "convey" a work means any kind of propagation that enables other parties to make or receive copies. Mere interaction with a user through a computer network, with no transfer of a copy, is not conveying.

5. Conveying Modified Source Versions.

You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:

- a) The work must carry prominent notices stating that you modified it, and giving a relevant date.
- b) The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to "keep intact all notices".
- c) You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not invalidate such permission if you have separately received it.
- d) If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.

A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an "aggregate" if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation's users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate.

6. Conveying Non-Source Forms.

You may convey a covered work in object code form under the terms of sections 4 and 5, provided that you also convey the machine-readable Corresponding Source under the terms of this License, in one of these ways:

Restrictive License Examples (AGPL, LGPL)

Based on the GPL
AGPL is more restrictive; explicitly
covers hosting use

LGPL is less restrictive, providing
exceptions to the viral aspects of the
GPL

GNU Affero General Public License



13. Remote Network Interaction; Use with the GNU General Public License.

Notwithstanding any other provision of this License, if you modify the Program, your modified version must prominently offer all users interacting with it remotely through a computer network (if your version supports such interaction) an opportunity to receive the Corresponding Source of your version by providing access to the Corresponding Source from a network server at no charge, through some standard or customary means of facilitating copying of software. This Corresponding Source shall include the Corresponding Source for any work covered by version 3 of the GNU General Public License that is incorporated pursuant to the following paragraph.

GNU Lesser General Public License



A "Combined Work" is a work produced by combining or linking an Application with the Library. The particular version of the Library with which the Combined Work was made is also called the "Linked Version".

3. Object Code Incorporating Material from Library Header Files.

The object code form of an Application may incorporate material from a header file that is part of the Library. You may convey such object code under terms of your choice, provided that, if the incorporated material is not limited to numerical parameters, data structure layouts and accessors, or small macros, inline functions and templates (ten or fewer lines in length), you do both of the following:

4. Combined Works.

You may convey a Combined Work under terms of your choice that, taken together, effectively do not restrict modification of the portions of the Library contained in the Combined Work and reverse engineering for debugging such modifications, if you also do each of the following:

...

- 1) Use a suitable shared library mechanism for linking with the Library. A suitable mechanism is one that (a) uses at run time a copy of the Library already present on the user's computer system, and (b) will operate properly with a modified version of the Library that is interface-compatible with the Linked Version.

Sample: Company's General Guidelines for Third Party Software

Company typically approves components that have been dedicated to the public domain or licensed pursuant to permissive licenses which do not restrict use, modification, distribution, and hosting of the components (such as Apache, BSD and MIT-style licenses).

Company typically approves, after case-by-case consideration, unmodified components that have been licensed pursuant to Public Licenses (such as Mozilla Public License (MPL) v2.0, Common Development and Distribution License (CDDL) v1.1, Eclipse Public License (EPL) v1.1, Common Public License (CPL) v1.1, and Microsoft Public License (Ms-PL)), with the exception of the Mozilla Public License (MPL) v1.1.

Company does not typically approve Third Party Software that has been licensed pursuant to the GNU General Public License (GPL) or Affero GNU General Public License (AGPL) except strictly for internal use.

Company reviews components licensed pursuant to other open source and commercial licenses, including the GNU Lesser General Public License (LGPL), on a case-by-case basis.

Company typically does not approve any open source license that requires Company to:

- distribute any of its source code

- allow reverse engineering of its products
- charge no or low fees for its software, or
- that contains other viral provisions that apply to Company's code as a whole, and not just to the third party software.

When a component is hosted, and the open source license does not expressly address hosting, Company may imply hosting rights as part of the general distribution rights in the absence of explicit prohibition related to hosting, and in some circumstances, may consider hosting as an internal use permitted by certain licenses.

License Compliance: Other issues

1. Linking

Some open source licenses (such as the AGPL, GPL, and LGPL) may infect proprietary code that is linked with the open source code

So-called “Dynamic link exception” for LGPL not always helpful due to need to allow reverse-engineering

Proper analysis requires knowledge of the type of linking

2. Most open source and freeware licenses, as well as many off-the-shelf commercial licenses, do not expressly address hosting rights

Resolving issues

Easy fixes...

- Failure to publish source code or modifications; source code versioning issues:
Make available and update compliance process as necessary to prevent happening
- Third party notice files

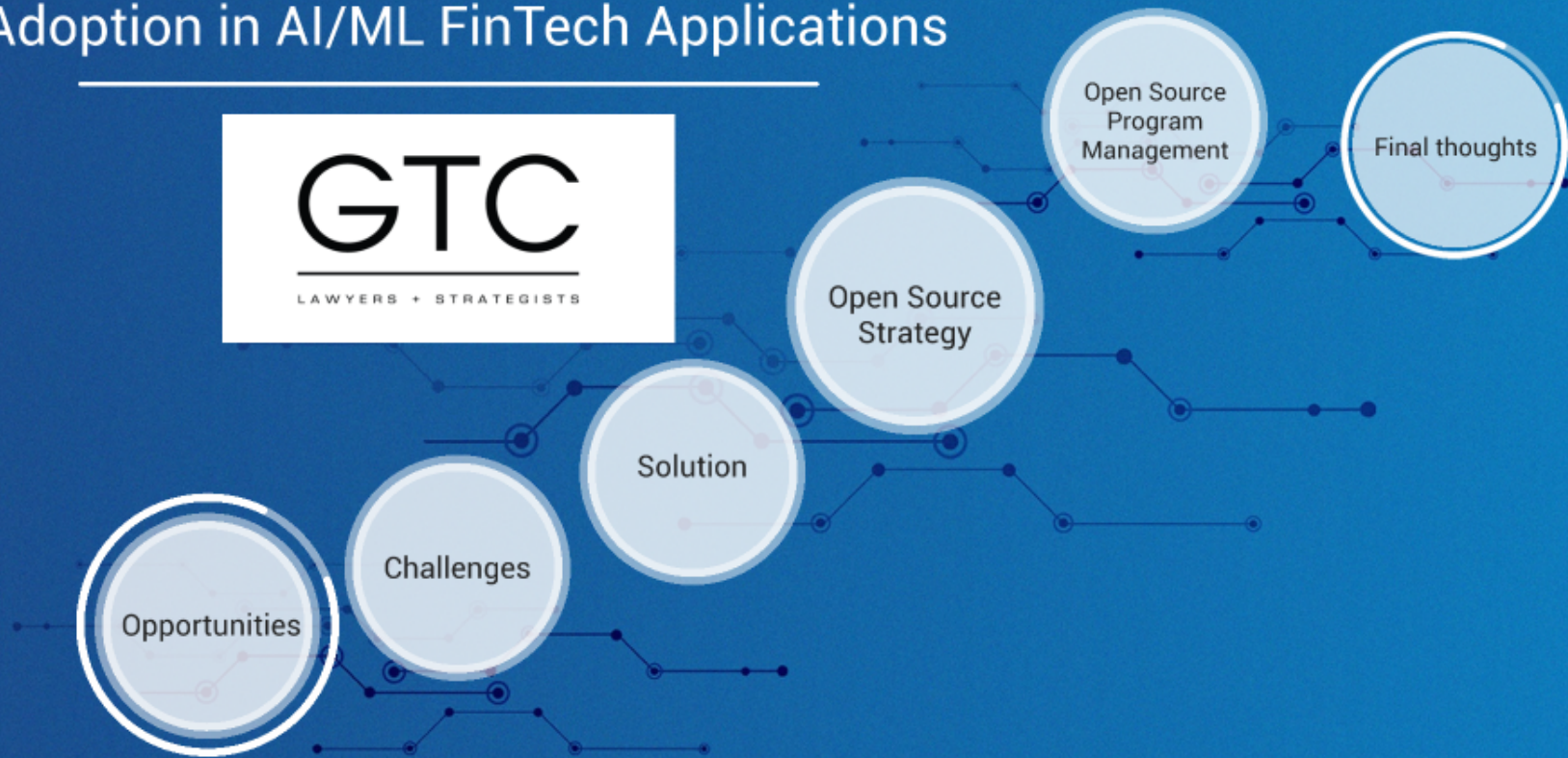
Require more effort...

- Remove or replace
- Re-write
- Re-engineer
- Versioning
- Re-license

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Final thoughts:

1. FinTech generally has been spurred by open source projects
2. AI FinTech applications can similarly leverage open source
3. Conduct a detailed evaluation of your product and business to determine optimum strategy for open source utilization: license-in and/or release
4. Sound compliance practices can mitigate risk associated with open source

GTC

LAWYERS + STRATEGISTS

Laila Paszti
Attorney
GTC Law Group PC & Associates
416.707.2818
lpaszti@gtclawgroup.com
www.gtclawgroup.com

This material is provided for your convenience and does not constitute legal advice or create an attorney-client relationship. Prior results do not guarantee similar outcomes. Attorney Advertising

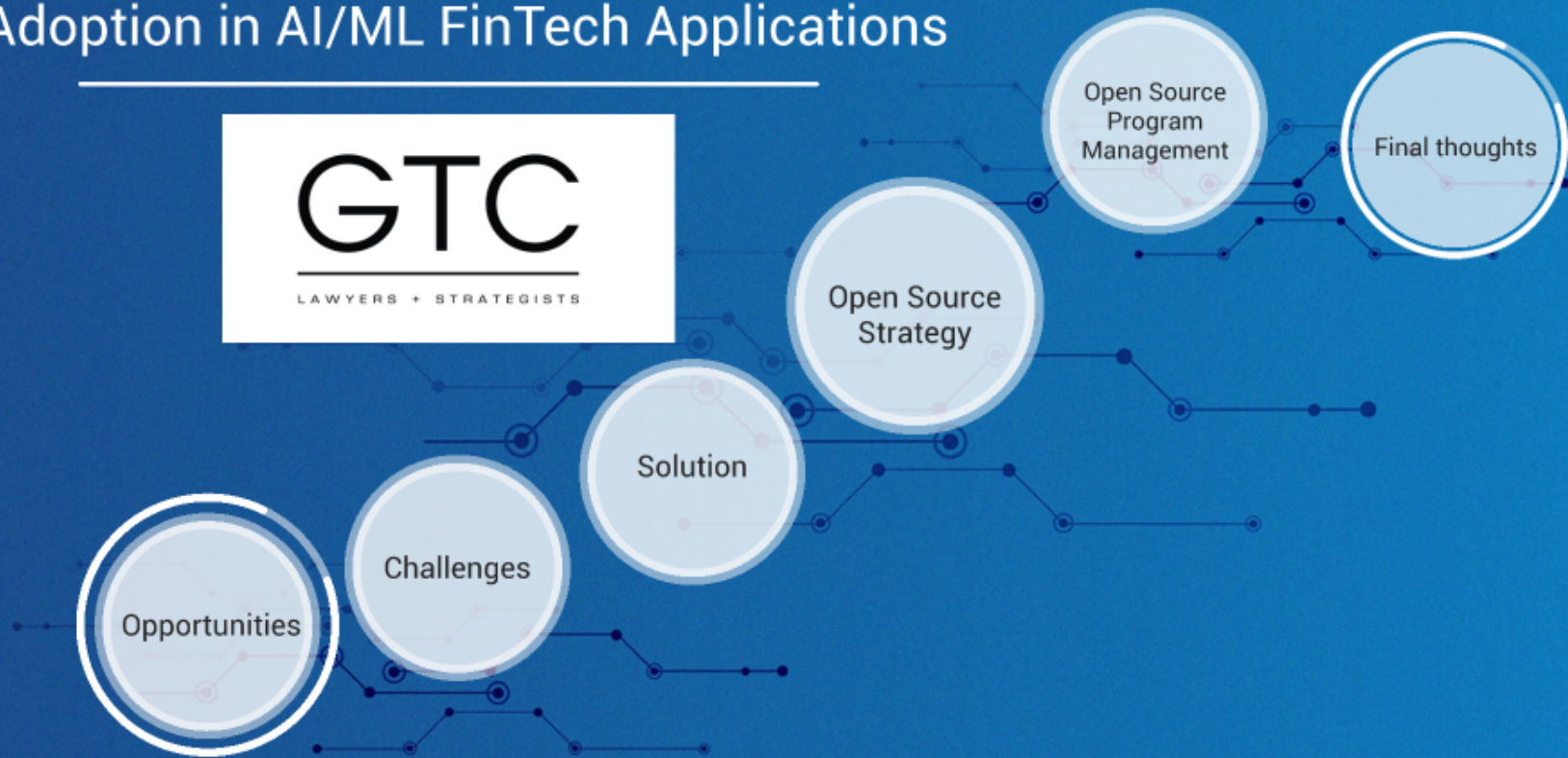
IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM

Smart Money Bets on Open Source Adoption in AI/ML FinTech Applications

GTC

LAWYERS + STRATEGISTS



LAILA PASZTI / 10.11.18

IP STRATEGY • MERGERS & ACQUISITIONS • BUSINESS & TECHNOLOGY TRANSACTIONS • DATA PRIVACY & SECURITY

GTCLAWGROUP.COM