



Online Dialogue is the Netherlands' leading agency when it comes to conversion optimization and evidence-based growth. Since our founding in 2009, we've been

successfully combining data science and psychology, resulting in ever-growing conversions and game-changing insights. We are always excited to try out new innovations in the field, and therefore carried out the following proof of concept with Euroflorist and Evolv.

## Massively Multivariate Testing with Evolv's AI Software

### EVOLUTIONARY ALGORITHMS IN CRO: THE FIRST CASE STUDY IN THE NETHERLANDS

The conversion optimization market is evolving every minute. New technologies are redefining conversion optimization programs and helping CRO teams find more ways to increase their conversions faster and more efficiently. Evolv is one of these new disruptive tools that uses AI and evolutionary algorithms to make breakthroughs in the world of CRO solutions.

### AI FOR HIGH EFFICIENCY

Evolv is an automated conversion optimization system that uses evolutionary algorithms and massive multivariate testing to find the best combination of ideas in the shortest time. Unlike traditional A/B testing, which only finds the best combination of two different variations, Evolv's massively multivariate testing tool finds the best combination of *thousands* of variants, optionally throughout an entire funnel. And unlike ordinary multivariate testing, which simultaneously tests multiple variants of multiple elements on a page, Evolv divides its experiments into generations, choosing to test only the best variants from every generation to find the most optimal combination.

## CASE STUDY



### BY THE NUMBERS

Online Dialogue used Evolv to optimize conversions on Euroflorist's product pages. The value of AI-powered CRO was evident after only one test:

**11** week experiment

**4** generations of testing

**256** possible candidates

**4.3%** conversion increase for best test



Each new generation teaches the system which combinations of which elements achieved the highest score. The winning element variants are combined in the next generation and reassessed (“mutations” are also tested to ensure the system finds the global—not just a local—maximum). With this method, the system finds the best combination of element variants as quickly as possible by eliminating the losers and testing combinations of the winners in order to find the best overall combination.

Each generation, therefore, improves the average efficiency of the combinations and does so almost immediately from the get-go. With this platform of “parallel interactive evolution,” Evolv offers a new method to perform large-scale multivariate tests and decreases time lost during experiments, analysis and the start of a new experiment.

## HOW DO EVOLV EXPERIMENTS DIFFER FROM A/B TESTS?

Conversion optimization is an evidence-based (re)design of customer journeys aiming to increase a specific human action, such as registering for a service or buying products. To understand which hypotheses are related to which effects, online experiments are conducted. Teams most often conduct A/B tests to gain insight on which adjustments improve the site’s performance.

Setting up an A/B test, however, is time-consuming and requires a very large audience in order to gather reliable test results in a timely manner. The number of experiments that can be performed is dependent on the stream of visitors of the site as well as on the time resources available. In practice, we have noticed that the elements that are able to be tested are limited and prioritization is key for conducting the right experiments.

With this in mind, we believe Evolv has disruptive power in this market. The platform performs best with many inputs contributing to a large “search space” of potential design combinations. The larger the search space, the greater potential for the global maximum—and the Evolv system is designed to find this maximum as quickly as possible.

That doesn’t mean you should test everything that comes to mind. You still want to be thoughtful about which elements and variations you test to yield the best results, as testing more “mediocre” ideas means the system may take longer to weed those ideas out. Hence, prioritization of the hypotheses is key for maximum efficiency.

The number of combinations that will be tested within a generation are limited by the number of visitors available, but will surely, in comparison with A/B test, make full use of your test potential. Evolv maximizes the number of tests that you are able to conduct within your digital domain.

## PROOF OF CONCEPT AT EUROFLORIST

It didn't take long to find a suitable CRO team who shared our same enthusiasm and wanted to give Evolv a try. Euroflorist, one of Europe's leading flower delivery companies, was a perfect candidate. Early pioneers in the world of ecommerce, Euroflorist was one of the first hundred online web shops worldwide, selling their first online bouquet in 1995. Today they sell an average of two million bouquets annually from 19 locations across 11 countries, and most of their revenue comes from online sales. Furthermore, they have more than ten million monthly unique visitors on their sites and have achieved conversion rates above 25%!.

In the spring of 2017, we—Euroflorist and Online Dialogue—embraced Evolv and took up the challenge and to launch a proof of concept (POC) on the Euroflorist's Swedish desktop site.

This is what we wanted to learn from the POC:

1. Could Evolv and its method prove that it improves conversion rates? Would it live up to the promise?
2. Would the winning combination be one we were expecting or not?
3. Realistically, when would you use this solution?

## SETTING UP THE EXPERIMENT

We decided to focus the experiment on Euroflorist's product detail pages—a key part of the purchase funnel that presents the user with multiple options for their bouquet. Clarity here is key. Conversions would be defined as purchases.

The POC started with a comprehensive hypothesis session. Based on Online Dialogue's "Evidence-Based Growth" principal, we first conducted research on the behavioral determinants, using data such as former (A/B) tests, surveys, heatmaps, and scientific research on flower sales. During the hypothesis session, hypotheses were prioritized based on the odds of success (backed up by available research data), and test variants were designed.

Secondly, we carried out a bandwidth calculation to determine the amount of experiments we can statistically run. It proved that with the visitors of Euroflorist in Sweden we could test eight combinations per generation. Therefore, we selected eight different elements to test, each with one variant per element. In practice, it is possible to create more variants per element, but for this POC, we opted for a simple setup.

The eight elements are listed below, with the variation communicated at a high level (though not the in-depth hypotheses for each).

Want more information on how this testing method can help your organization? Please contact:

### Valentina Djoemai

CRO STRATEGIST & PROOF-OF-CONCEPT LEAD, [ONLINE DIALOGUE](#)

[valentina.djoemai@onlinedialogue.com](mailto:valentina.djoemai@onlinedialogue.com)

+3130 4100 177

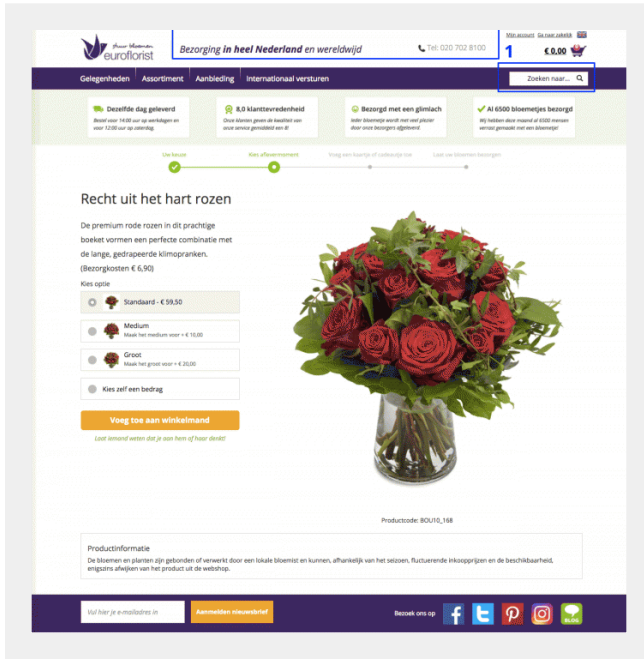
### Michael Versluis

SALES DIRECTOR,  
EVOLV AMSTERDAM

[michael.versluis@evolv.ai](mailto:michael.versluis@evolv.ai)

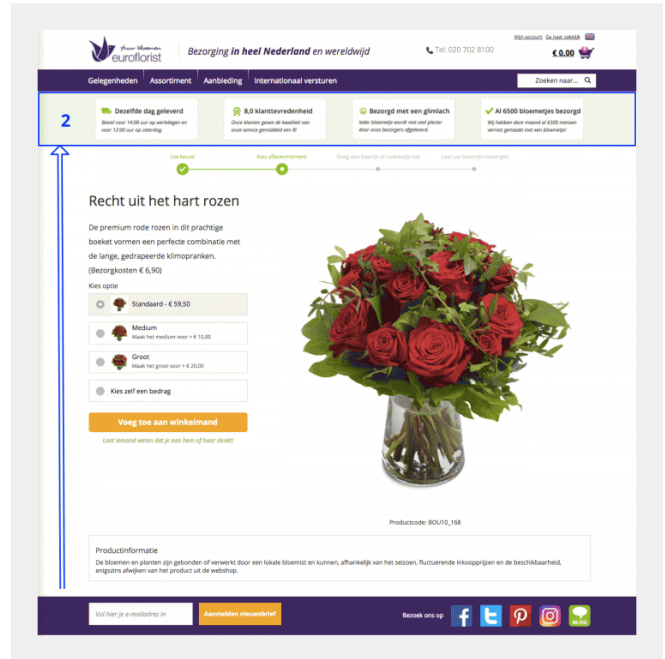
+31 6 5188 4180

# EVOLV: EUROFLORIST



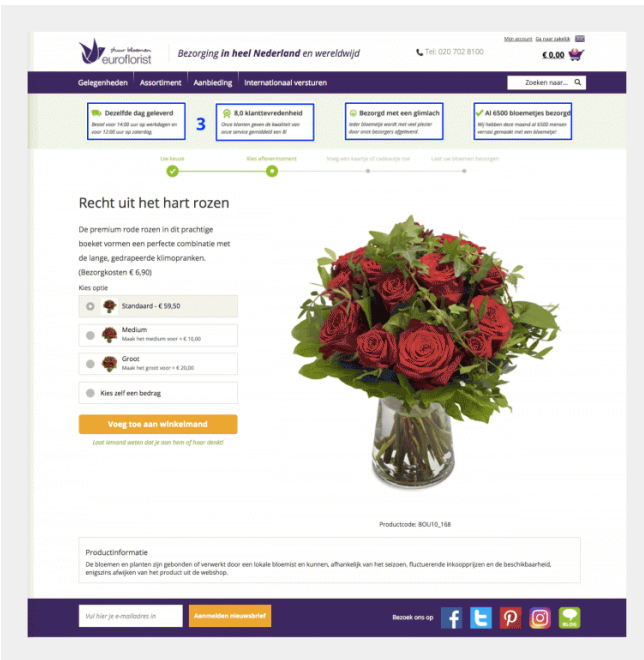
## 1. HEADER

The control header is quite crowded; removing this might mean less distraction and more focus on where the visitors should be drawn.



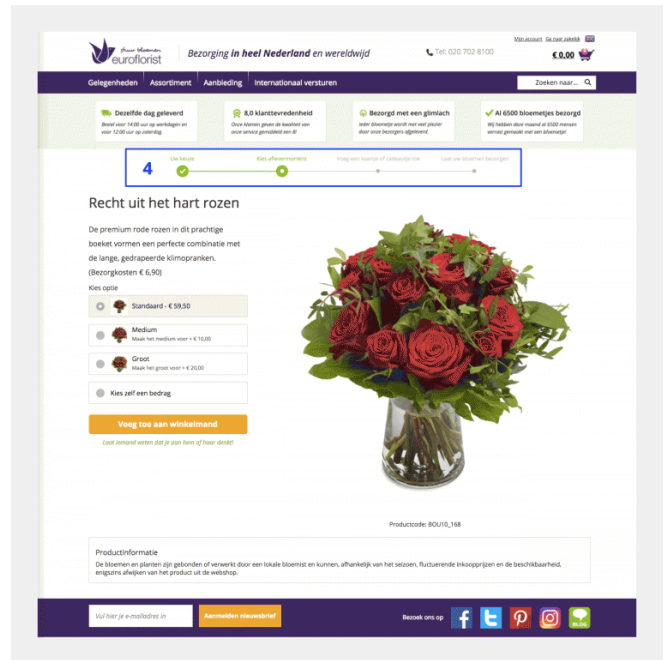
## 2. USP BAR POSITION

The visibility of the unique selling proposition (USP) bar—at the bottom of the control page—is not ideal. The variant elevates it to the top.



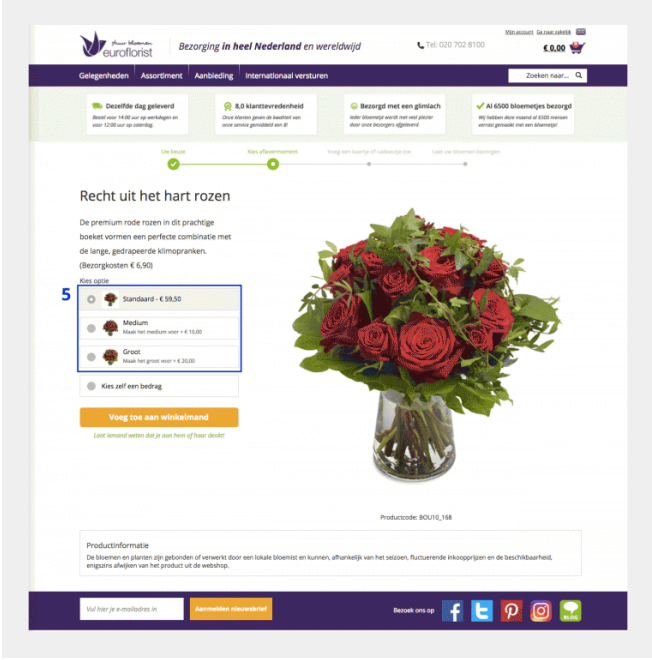
## 3. USP BAR CONTENT

The variant's motivating content bar displays new information, intended to further motivate Euroflorist shoppers.



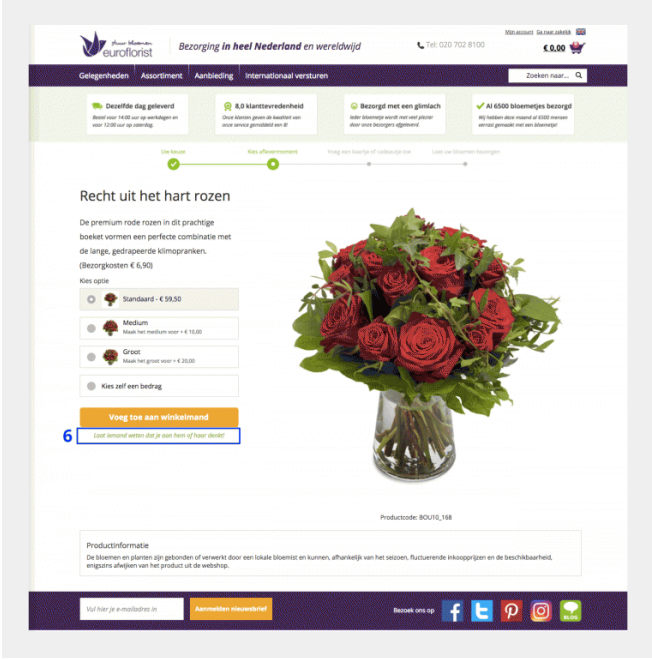
## 4. PROGRESS BAR

In the control, visitors were not informed what happens when they decide to buy, which can cause confusion and decreased sales.



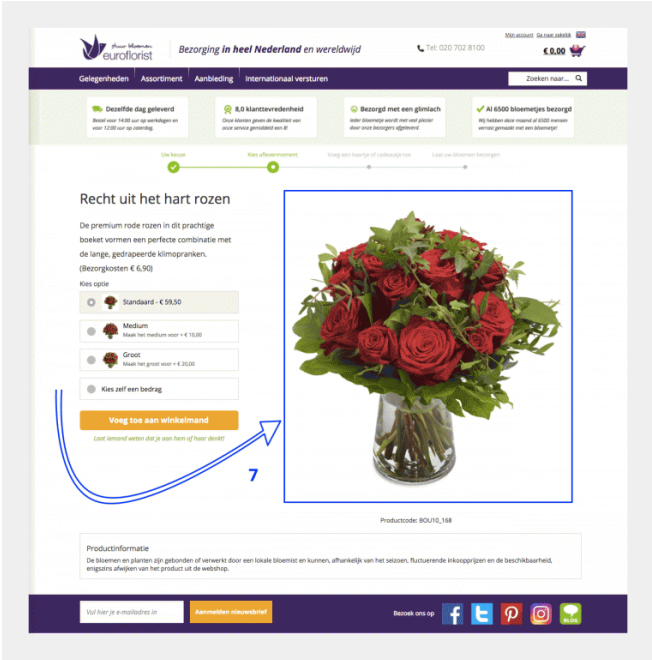
5. PRICING DISPLAY

The pricing of bigger bouquets was altered to show the extra costs on top of the smallest bouquet instead of the total price of the bouquet.



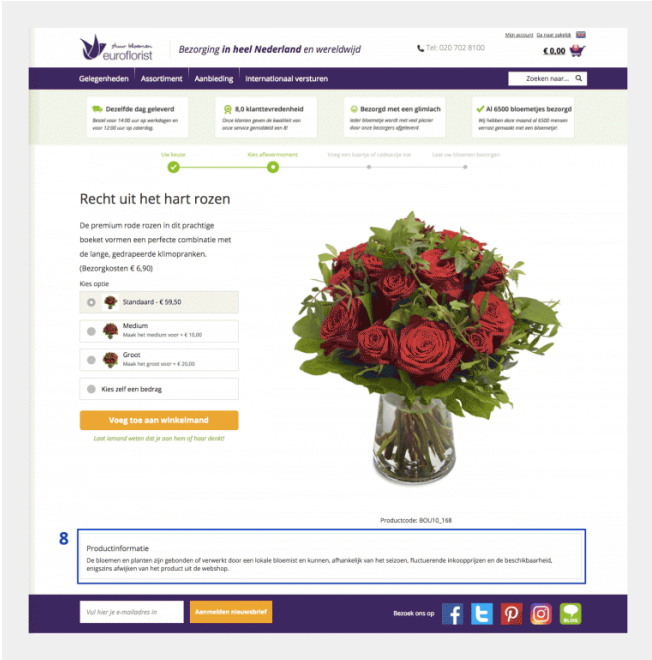
6. SOCIAL SENTENCE UNDER CTA

Visitors assess the call-to-action (CTA) before proceeding. To ensure trust, a social proof sentence was tested, encouraging the shopper to proceed with the purchase.



7. POSITION OF THE PRODUCT PHOTO AND CTA

In Western society, text is read from left to right. We recommend switching the photo block on the left with the block of product information and price choices on the right. With the variant layout, theoretically, the visitor would see the different price ranges first and the photo second.



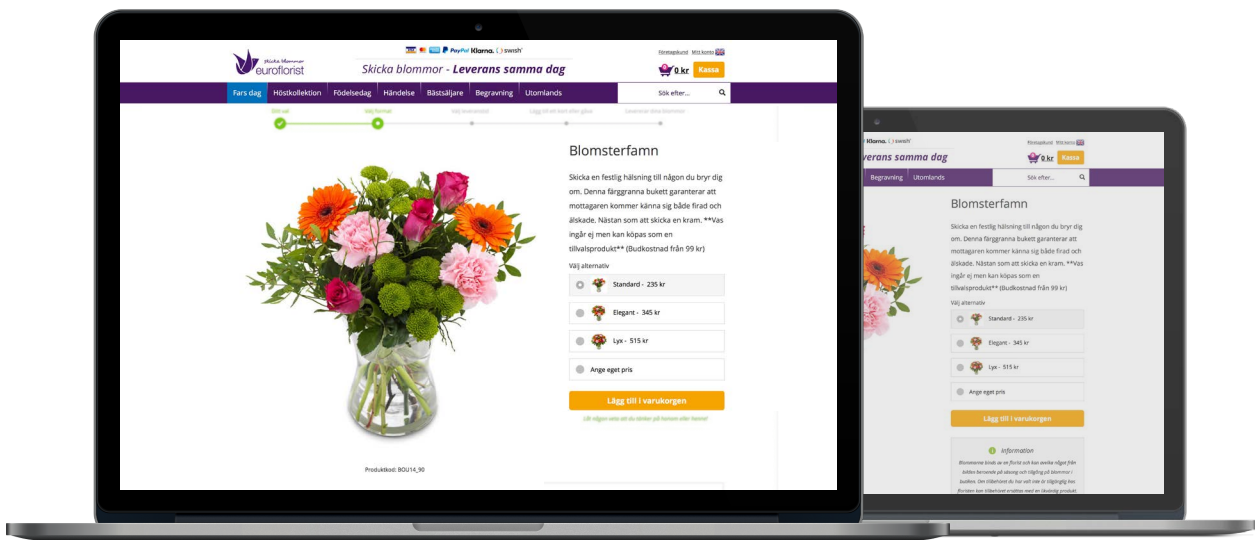
8. PRODUCT INFORMATION/WARNING

Directly under CTA, there is a disclaimer regarding the possibility of receiving a slightly different-looking bouquet than the picture shows. While this message is important, its proximity to the CTA could potentially distract a shopper from their purchase. Therefore, the variant displayed it near the bottom of the page.



## SURVIVAL OF THE FITTEST, TEST RESULTS, AND CONFIRMATION

The whole experiment on the Swedish website took place within eleven weeks, of which four generations were fully executed. In the third generation, we saw the first significant positive uplifts, and the fourth generation had even more combinations with a significant increase in conversion. We observed that each generation achieved a higher average conversion rate (all active combinations combined) than the average conversion rate of the previous generation. Step by step, the system grew toward a significant winner.



**The best-performing combination after 4 generations increased conversions by 4.3%.**

We were interested to see which changes—and, equally important, which combination of changes—did and did not result in a higher conversion rate. In the best-performing combination, shown above, the following aspects have been adjusted compared to the original:

1. Progress indicator at the top of the page
2. Price perception adjustment
3. Social sentence under the CTA
4. Product warning information further removed from CTA
5. New USPs in the USP bar

We do not know for certain that this is the global maximum. Further generations would give us more confidence, but due to time constraints and the significant

results shown by Evolv, we decided to continue with the next step in the validation of the Evolv system: a reassessment of the winning combination with an A/B test to learn if this combination would beat the original.

The new test, done over the course of two weeks with an independent A/B testing tool, led again to the fact that the winning Evolv combination makes a significant impact when compared to the default. The Evolv system has proven capable to find winning combinations, and ones which are different than the potential winners we would have guessed—i.e., all elements active, potentially without changing position of the product photo and the CTA block (which is indeed contrary to existing usability of general e-commerce website). Our header adjustments and the combination of progress indicator and USP bar at the top did not lead to positive impact.

## HOW WOULD WE USE EVOLV IN A CRO PROGRAM?

For us, and A/B test on its own is an exploration in itself. We are an advocate of making full use of your testing potential by upscaling the quantity first and working on the quality after this step. Fifty percent of our learnings of client behavior come from determinant studies (prior research), and the other fifty percent from analysis of A/B tests. The test analyses give us a lot of information on changing behaviors.

Due to the multivariate setup, the Evolv experiment gives us less detailed insights on behavioral change, but *more* insights on interactive effects and the right composition of elements. For us, the implementation of Evolv would be ideal to optimize customer journeys (e.g. website, or a set of pages) when research and A/B tests have shown what works for you and what does not. If you know which hypotheses are winning, you could use this for an experiment in which you want to give a boost to find the ideal combination of elements.

Normally, after a streak of winning experiments we make a “re-align” of the optimized flow; it seems natural to use Evolv here. In addition, Evolv is working on a segmentation update with AI: the independent recognition of corresponding behavioral segments. With this we learn which composition of elements works best. When this becomes successful, “always-on” testing will be closer. The current setup will still be a Evolv experiment bound by time with a winner that is going to be implemented. In the future, this will not be needed: there will be a continuous experiment running to improve the conversion, with new inputs added when certain elements have proven to have no impact.

## COSTS (TIME) VS. BENEFITS (RESULTS)

An Evolv experiment takes a lot more time than an A/B test. It takes about the same amount of time to set up and conduct one generation of Evolv testing as it does to set up and run a single A/B test. However, While A/B tests show positive effects after an average of 3 to 4 tests, each single Evolv experiment generates significant

results—provided that the right input has been used with the system. In our opinion, the right input with Evolv is extra-important to show the positive impact on the business case. If this is done properly, Evolv would be a better fit to find the global maximum, compared to a long streak of A/B tests. If *not* done properly, you are likely to end up wasting your time testing elements that have no impact on conversion, making Evolv potentially less effective than a long streak of A/B tests.

Although the cost of the tool is greater than many A/B testing solutions, the promise of growth from local maxima to global maxima should more than compensate the cost. We believe that Evolv will be a suitable addition for conversion optimization clients who have already been learning about their customers' behaviors and are ready to get the most out of these learnings! Therefore, along with being a Lighthouse CRO Agency for Google Netherlands, an ABTasty premium partner, and an Optimizely partner, we are also a proud Evolv Authorized Solutions Partner.

**Want more information on how this testing method can help your European organization? Please contact:**

**Valentina Djoemai**

CRO STRATEGIST & PROOF-OF-CONCEPT  
LEAD, [ONLINE DIALOGUE](#)

[valentina.djoemai@  
onlinedialogue.com](mailto:valentina.djoemai@onlinedialogue.com)

+3130 4100 177

**Michael Versluis**

SALES DIRECTOR, [EVOLV TECHNOLOGIES](#)  
[AMSTERDAM](#)

[michael.versluis@evolv.ai](mailto:michael.versluis@evolv.ai)

+31 6 5188 4180



P +1 415.444.9040  
E [info@evolv.ai](mailto:info@evolv.ai)  
W [evolv.ai](http://evolv.ai)

**Learn more at [evolv.ai](http://evolv.ai)**