Google Site Search Notes

Benefits:

* **Reduced load and dependency** on our **SQL Server**
* **Seamless integration with** **Google Analytics**, which may provide more data collection and insight than the current model.
* **Misspell** / **Typo** logic for searches
  + Entering “Pole**b**ault” in Extron search yielded 0 results and a “Not Found” message. Entering “Pole**b**ault” within a Google search will yield the normal results for “Polevault”.
* **Relatively** **low cost:**
  + $100 year / 20,000 queries per year
  + $250 year / 50,000 queries per year
  + $750 year / 150,000 queries per year
  + $2000 year / 500,000 queries per year
* **Integrated features** for “Advanced” search that current Extron search does not have
  + Label refinements
  + Image search
* **User Intent** and **Relevance**. Google claims to have superior understanding of User Intent which is difficult to recreate without a significant engineering effort.
* **The peace of mind of using a Brand Name system** that adheres to standards and industry best practices.
* **Easy to get started** with the “out of the box” package that Google provides.
* **Automatically generated thumbnails**, bringing reduced load on file system.
* **Increased search speed**: Currently running a search on Extron.com custom search took on average 5.3 seconds. A Google search of the same string took on average 1.7 seconds

Drawbacks:

* **Results will NOT be clean and filtered** right out of the box.
  + It will take some massaging of the Google Indexing system to get the same type of controlled results as our own search.
  + Babysitting of the entire system may be needed the first few months to ensure a proper user experience.
* **Security and UX issues** may arise if Google’s bots picks up a result that you don’t want seen. You must submit an “On demand” ticket to Google to have these removed. Which takes extra time.
* **Passive updates take a while**: around 1-2 weeks
  + “On demand” indexing features may require extra fees, and still takes around 24hrs.
  + “On demand” feature has strict quotas based on account level.
* **No Role/Login based search results**. Since you just get back Googles public indexed search results. You will not be able to show specific result sets based on Role/login.
* **The Ranking logic is based on “Google Magic”** and non-inspect able or modifiable.
* **Seems finicky and difficult to customize page values and ranks**. Articles have claimed that a common frustration is dealing with inferior pages listed above the important ones.
* **No direct feed to the database**: some pages w/ dynamic database driven content may not be indexed properly
* **Some results may never appear**: Google’s indexing is completely automated. This could imply that some of our pages may never get picked up and we would never notice. We would need to predetermine every single desired result, and confirm that it was indexed properly and searchable.
  + Even then, it would require reverse engineering of Google’s rules to fit the missing pages back into their automated procedure.
* **The results page is only a themed Google page with some limited customizability.** Out of the box, the GSS result page is underwhelming and will require programming and API manipulation to render the results seamlessly with the main design of Extron.

Considerations:

* **GSS is not an Enterprise level solution**. If Extron grows beyond GSS’s intended scope, it may be inefficient to start over with a new system.
  + For example, if EXTRON will need to start indexing:
    - .xls
    - .pdf
    - .dwg
    - .ppt
    - .msg
    - .wmf
    - This is not offered with GSS by automatic means
* **Relinquishment of full control**: Currently, we could manifest new search results instantly without being governed by an external party.
  + If timeliness is not a concern, and automation with the convenience of a 3rd party handling our search indexing was more desirable, Google Site Search would be the superior option.
* **It will never be 100% fully automated** someone should always verify periodically that Google’s bots are doing the job correctly.
* **Google outdates their APIs often** so you may need to redo work when this occurs. There is also always a level of uncertainty of when they decide to do this.
* **Google may drop the product from their line.** Resulting in having to find another option or completely reverting to the old system.
  + Confirmed happening: 4/1/2018

Sources:

<http://artsy.github.io/blog/2014/10/23/how-we-customized-google-site-search-at-artsy/>

<http://fortune.com/2017/02/21/google-site-search-discontinued/>

<https://enterprise.google.com/search/products/gss.html>

<http://www.magikcommerce.com/blog/how-to-install-google-site-search-in-magento-google-search-for-magento-stores/>